Automotive Three-Dimensional Vision Through a Single-Photon Counting SPAD Camera

Danilo Bronzi, *Member, IEEE*, Yu Zou, Federica Villa, *Member, IEEE*, Simone Tisa, Alberto Tosi, *Member, IEEE*, and Franco Zappa, *Senior Member, IEEE*

***Abstract*—We present an optical 3-D ranging camera for automotive applications that is able to provide a centimeter depth resolution over a 40° × 20° field of view up to 45 m with just 1.5 W of active illumination at 808 nm. The enabling technology we developed is based on a CMOS imager chip of 64 × 32 pixels, each with a single-photon avalanche diode (SPAD) and three 9-bit digital counters, able to perform lock-in time-of-flight calculation of individual photons emitted by a laser illuminator, reflected by the objects in the scene, and eventually detected by the camera. Due to the SPAD single-photon sensitivity and the smart in-pixel processing, the camera provides state-of-the-art performance at both high frame rates and very low light levels without the need for scanning and with global shutter benefits. Furthermore, the CMOS process is automotive certified.**

***Index Terms*—3-D ranging, 2-D imaging, single-photon avalanche diodes (SPADs), time-of-flight (ToF), range-finding.**

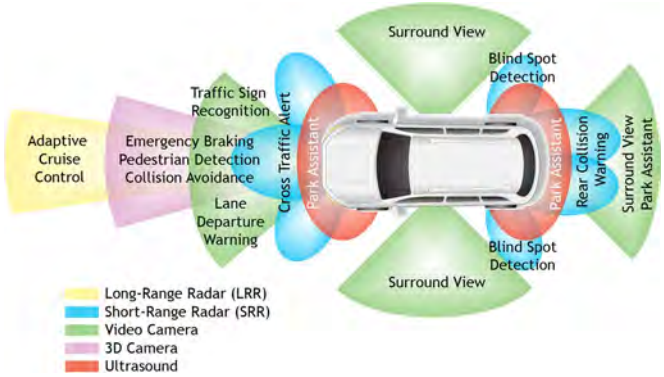

Fig. 1. Vehicle functions and technologies likely to be used for advanced driver assistance systems.

I. Introduction

THREE-DIMENSIONAL (3-D) vision systems and object detection are more and more requested in the automotive field for decreasing road accidents and providing better driving experience [1]. As a result, in the last decade, global automotive industry has seen the proliferation of new control strategies—such as electronic stability control, rear-view facing cameras, vision-based pedestrian detection systems, lane departure warning systems, night vision systems, electronic parking assistance, blind spot detection, 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 all aspects of driving (including safety, drivability, and fuel economy) called advanced driver assistance systems (ADAS).

The most advanced ADAS technologies to fight road accidents are the collision avoidance systems (CAS), which typically employ either radar-, lidar-, ultrasonic- or camera-based depth sensors. 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. For instance, radars have a long detection range and operate slightly better under bad weather conditions. However, radar's field of view is much narrower than a camera-based system's, which is much more effective at detecting and differentiating between moving and stationary objects. Ref. [2] gives an overview of typical strengths and weaknesses for automotive sensors available today: strengths of different sensor types (e.g. radar and camera systems) can be combined to end up with an improved sensor system. This represents the concept of multi-sensor data fusion. Concerning camera-based 3-D vision systems, we can consider two main categories based on different principles: stereo-vision (SV) and time-of-flight (TOF). SV systems require neither moving parts nor active illumination and provides high spatial resolution at low power consumption. Moreover, in daylight condition they usually have better performance—compared to TOF systems—since strong sunlight enhance image contrast, thus helping in the 3-D reconstruction, whereas in TOF this usually leads to detector's saturation. Unfortunately, in SV systems, the projection of a 3-D scene onto two-dimensional sensors causes a compression of spatial information and the need for high-contrast scenes to identify projections of point pairs (correspondence problem). Additionally, SV suffers from problems resulting from shadows and moving objects, and depth resolution is determined by optical arrangements of the two cameras and cannot be changed [3]. Intensive processing is also required to extract the correct distance information, thus implying a very low frame-rate. Eventually, SV-based systems deliver ambiguous distance measurement in difficult lighting conditions and do not represent

Manuscript received March 20, 2015; revised July 3, 2015 and September 19, 2015; accepted September 24, 2015. This work was supported by the European Commission FP7-ICT Framework through the "MiSPiA" Project under G.A. 257646. The Associate Editor for this paper was P. Cerri.

D. Bronzi, Y. Zou, F. Villa, A. Tosi, and F. Zappa are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy (e-mail: danilo.bronzi@polimi.it).

S. Tisa is with Micro Photon Devices SRL, 39100 Bolzano, Italy.

the best choice in automotive applications, especially when fast (higher than standard video-rate) 3-D ranging is demanded.

Active triangulation methods are not affected by shadows and moving object, but the major drawback is the need of a costly and cumbersome source to generate structured light; in addition, a mechanical solution is required to project light patterns on the scene, thus making the system very sensitive to vibrations and difficult to use in automotive applications [4].

Such limitations are overcome by time-of-flight (TOF) cameras that, in the last decade, have gained attention thanks to their attractive characteristics: TOF-vision systems are made of a single camera, which can include the imager and a light source, and can reach higher frame-rates than SV-systems because depth measurement is straightforward since no correspondence problem has to be solved and in-pixel pre-processing can be performed.

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 light source to an object and then back to a photodetector. The measured delay time is then converted to distance, as the speed of light is given. Instead of analog dTOF sensors, our idea was to employ a "digital" single-photon detector able to exploit the ultimate sensitivity, by detecting individual quanta of light, and to provide picosecond timing resolution for achieving millimeter accuracy over a long range. Pioneer multi-pixel rangefinders based on Single-Photon Avalanche Diodes (SPADs), able to extract distance information from photons' time-of-flight, were developed in [5] by using a custom integration process for bonding a back-thinned SPAD wafer on the readout and processing electronics, consisting of an array of 32 $\times$ 32 time-to-digital converters (TDCs). Since then, other approaches were developed to fabricate SPAD arrays with on-chip TDCs [6]–[10] or time-to-amplitude converters (TACs) [11] using cost-effective single-chip standard CMOS processes. Nonetheless, TDCs/TACs produce a consistent amount of raw data and require large computational effort and bandwidth to handle massive data throughput. Moreover, typical laser sources for dTOF (with low jitter and sub-nanosecond pulse width) are bulky and expensive and are not suitable for rugged automotive environments.

We considered an alternative solution, represented by indirect time-of-flight (iTOF) estimation, where distance information is extracted from the phase-delay between a pulsed-light (PL) [12] or continuous-wave (CW) [13]–[15] excitation shone toward the target and its back-reflected echo, such as in heterodyne or homodyne demodulation. With respect to CMOS/CCD [16] and CMOS iTOF rangefinders [17], a SPAD array has lower fill-factor but inherently better timing resolution (dominated by SPAD timing jitter, typically below one hundred picoseconds), higher accuracy (impaired only by photon shot-noise), and better linearity (amplification and quantization non-idealities are not present) [13]. Moreover, in-pixel demodulation allows not only to compute the depth image, but also to retrieve information about the actual intensity of light reflected by the scene; the latter information can be used to enhance image segmentation and to remove errors caused by phase-wrapping, thus extending the sensor's distance range [18]. In this sense, iTOF measurements provide more information per photon, compared to dTOF technique; furthermore, acquisition speed is higher because in-pixel demodulation reduces output data throughput and external computational effort, thus fulfilling requirements on minimal and transparent data transmission for sensors fusion. Additionally, neither high bandwidth electronics nor short pulse width lasers are required, hence allowing the development of really cost-effective systems.

In this paper, we present a complete automotive-oriented camera for optical 3-D ranging using an indirect time-of-flight approach, based on a 64 $\times$ 32 CMOS SPAD imager able not only to deliver two-dimensional (2-D) intensity information through free-running photon-counting, but also to perform smart light demodulation with in-pixel background suppression, thus enabling three-dimensional (3-D) depth-resolved mapping of objects in the scene. We conceived the camera for providing simultaneous 2-D and 3-D videos of rapidly changing (e.g., in shape, intensity, distance, etc.) scenes in light-starved environments. We validated the system both indoor and in real traffic scenarios, yielding 110 dB dynamic-range, high-speed (100 fps, frames per second) depth measurements, with better than 60 cm precision at 40 m distance.

II. Pulsed-Light vs. Continuous-Wave Indirect TOF

As previously stated, two different techniques can be exploited for iTOF measurements: continuous-wave iTOF (CW-iTOF) and pulsed-light iTOF (PL-iTOF). In this section, we discuss first the two techniques separately, in order to provide a mathematical analysis for theoretical precisions over the whole distance range. Finally, we compare them for identifying the best performing technique which will drive design and optimization of the system.

A. Pulsed-Light Indirect TOF

In PL-iTOF systems, a laser source emits light pulses with amplitude A and duration T_P, which sets the maximum distance range to $d_{\mathrm{MAX}} = T_P \cdot c/2$, given the speed of light c. The reflected signal, together with background light and detector noise, are integrated within three distinct time slots (see Fig. 2). In [19] we demonstrated how to improve precision by means of the double sampling technique (DST), where a first window W_0, synchronous with the laser pulse, and a second window W_1, in quadrature with the laser signal, accumulate two portions of the reflected photons, each one proportional to object's distance; whilst a third window W_B, enabled when no light pulse is emitted, collects only background photons. If C_0, C_1, and C_B are the counts accumulated in W_0, W_1, and W_B, respectively, then the object's distance d, the received active-light intensity A_R and the background B are given by [19]:

$$d = \frac{c \cdot T_P}{2} \cdot \left(\frac{C_1 - C_B}{C_0 + C_1 - 2 \cdot C_B} \right) \tag{1}$$

$$A_R = \frac{C_0 + C_1 - 2 \cdot C_B}{T_P} \tag{2}$$

$$B = \frac{C_B}{T_P}. \tag{3}$$

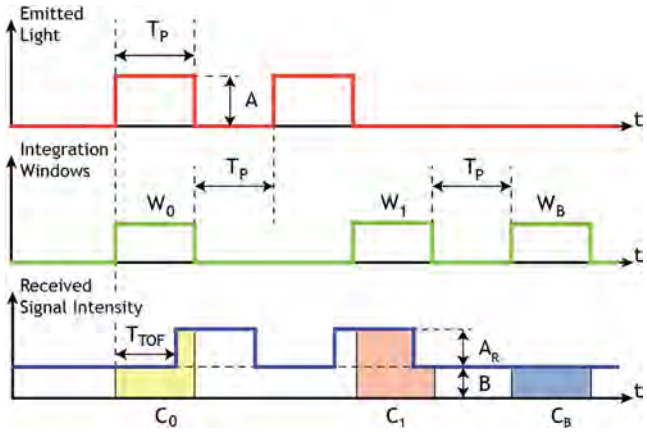

Fig. 2. Emitted light pulses, integration time windows, and reflected signal for PL-iTOF. W_0 and W_1 allow one to identify the time delay (T_{TOF}), proportional to the object's distance d, while W_B is used for background suppression.

If the distance measurement is repeated N times, by applying the error propagation rule to Eq. (1), the distance precision in PL-iTOF ($\sigma_{d,\mathrm{PL}}$) is given by the following equation [19]:

$$\sigma_{d,\mathrm{PL}} = \frac{d_{\mathrm{MAX}}}{A_R} \cdot \sqrt{\frac{A_R \cdot k_1(d) + B \cdot k_2(d)}{N}} \tag{4}$$

where the relationship $d_{\mathrm{MAX}} = c \cdot T_P/2$ was used to define $k_1(d) = d/d_{\mathrm{MAX}} - (d/d_{\mathrm{MAX}})^2$ and $k_2(d) = 1 - 3 \cdot k_1(d)$ coefficients.

B. Continuous-Wave Indirect TOF

In CW-iTOF cameras, a sinusoidal-modulated light source, with modulation period T_P, illuminates the scene and reflected light reaches back the detector, phase-shifted by an amount $\Delta\varphi$. Object's distance d is calculated by:

$$d = \frac{c \cdot T_P}{2} \cdot \frac{\Delta\varphi}{2\pi} = d_{\mathrm{MAX}} \cdot \frac{\Delta\varphi}{2\pi}. \tag{5}$$

To retrieve phase-shift information, the reflected wave is synchronously sampled by four integration windows of same duration T_{TAP}, thus providing C_0, C_1, C_2 and C_3 samples, as shown in Fig. 3. Through Discrete Fourier Transform, phase delay $\Delta\varphi$, reflected light intensity A_R and background B are given by [16]:

$$\Delta\varphi = \arctan \frac{C_3 - C_1}{C_0 - C_2} \tag{6}$$

$$A_R = \frac{\sqrt{(C_3 - C_1)^2 + (C_0 - C_2)^2}}{T_{\mathrm{TAP}} \cdot \mathrm{sinc}(\pi \cdot T_{\mathrm{TAP}}/T_P)} \tag{7}$$

$$B = \frac{C_0 + C_1 + C_2 + C_3}{4 \cdot T_{\mathrm{TAP}}} - \frac{A_R}{2}. \tag{8}$$

In order to compare commensurable quantities, B and A_R are equally defined for both PL-iTOF and CW-iTOF: B is the intensity associated only to signal uncorrelated with modulation; A_R is the peak-to-peak intensity of modulated signal impinging onto the detector. Now we can apply the error propagation rule to Eq. (5) and (6), thus obtaining the CW-iTOF precision, $\sigma_{d,\mathrm{CW}}$:

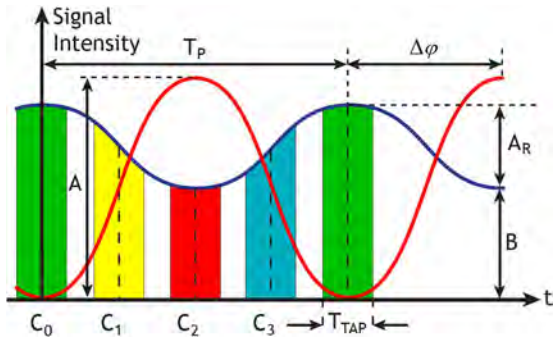

Fig. 3. Emitted modulated light (red curve) and reflected signal (blue curve) for CW-iTOF. The reflected light amplitude is less than the emitted one due to attenuation and it shows an offset due to background. The algorithm exploits four samples (C_0, C_1, C_2, C_3) to compute the phase-delay between the two signals.

$$\sigma_{d,\mathrm{CW}} = \frac{d_{\mathrm{MAX}}}{A_R} \cdot \sqrt{\frac{A_R + 2 \cdot B}{N}} \cdot \frac{1}{2\pi \cdot F(x)} \tag{9}$$

$$x = \frac{T_{\mathrm{TAP}}}{T_P}, F(x) = \sqrt{x} \cdot \mathrm{sinc}(x) \tag{10}$$

where N is the number of repeated measurements and $F(x)$ is a factor that brings into account the influence of the duration of the integration window on the CW-iTOF precision. This can be intuitively explained considering that a long sampling time T_{TAP} causes an averaging of the four samples, resulting in an attenuation of the received signal and thus in a lowering of the measurement precision; at the same time, a wide T_{TAP} allows to acquire more photons, thus improving precision. Therefore, an optimum T_{TAP}/T_P ratio, which maximizes $F(x)$ and precision, can be found. This optimal point is reached when T_{TAP} is about 40% of the modulation period, as shown in Fig. 4. We would like to underline here that in many other papers on the topic (e.g. [13], [16]) the formulation of Eq. (9) does not take into account the factor $F(x)$, thus leading to error underestimation.

C. Comparison of iTOF Techniques

Eq. (4) and Eq. (9) show that, for both PL-iTOF and CW-iTOF techniques, depth-precision depends on distance range (d_{MAX}), received light intensity (A_R) and background noise (B). However, by assuming a fixed A_R (i.e. by disregarding attenuation of the reflected signal due to solid angle and object's reflectivity) and considering a constant background, the CW-iTOF precision is independent of distance, while PL-iTOF depends on it. This is consistent with the fact that in CW-iTOF the whole echoed light is collected, regardless of reflecting target's distance; conversely, in PL-iTOF the collected signal is a function of object's distance. In fact, at short distance, W_1 collects almost no signal; consequently, SNR is very poor. This is however compensated by the signal collected in W_0, which is almost the whole reflected light and therefore shows higher SNR, thus allowing precise distance measurement. Similarly, at far distances, the large signal acquired by W_0 compensates for the poor information collected by W_1. At intermediate

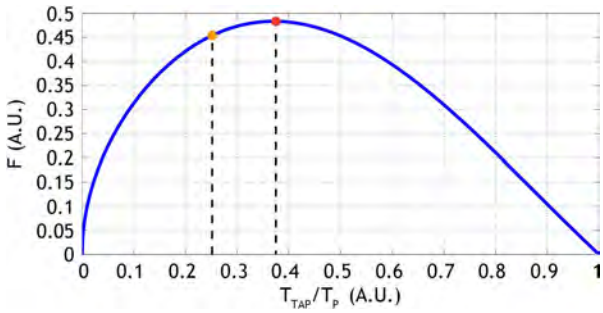

Fig. 4. To minimize distance error, the factor F reported in Eq. (9) must be maximized. This happens when the ratio between T_{TAP}/T_P is almost 40%. For the reasons clarified in the paper, we chose to work with a ratio of 25%.

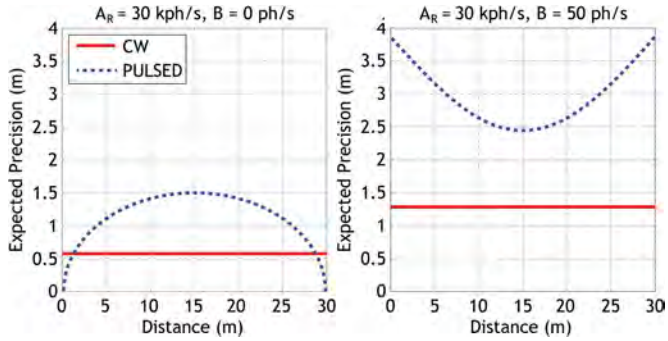

Fig. 5. Precision as a function of the distance for PL-iTOF and CW-iTOF techniques, with and without background, computed for a received photon rate of 30 kHz and 10 ms integration time. T_{TAP} for CW was set equal to 25% T_P.

distances, instead, the signal is equally collected by the two windows, thus resulting in a moderate SNR equal for both signals, so no compensation can occur and distance error is larger. On the other hand, in presence of strong background, its large variance dominates at both ends of distance range and the lowest error is achieved when the signal-to-background ratio (SBR) [13] is the same for both integration windows (i.e. at intermediate distances).

Fig. 5 shows another difference: while CW-iTOF employs only one period (T_P) for a complete measurement, PL-iTOF requires more than a period to compute the distance. Therefore, for same average power, PL-iTOF requires higher peak power, or longer integration time. Nonetheless, even at the same average power, PL-iTOF precision is not only worse than CW-iTOF precision, but it is also far more sensitive to background and dark counts, as proved in Fig. 5. For all these reasons, we designed a SPAD imager able to process at pixel-level 2-D intensity data and 3-D depth-ranging information through the CW-iTOF technique.

III. 3-D Vision System

Our 3-D vision system is based on a CMOS SPAD imager, an FPGA board for settings, data readout, and data upload to a PC, and a laser-diodes illuminator, as discussed in the following.

A. iTOF Pixel

Fig. 6 shows the sensor's pixel, presented in detail in [20], consisting of a SPAD detector, a quenching circuit, shaping electronics, three 9-bit counters and their respective storage memories, and output buffers for driving the column data bus. Each counter's control signal is driven by the external FPGA to properly drive counters in either interleaved mode, for lock-in iTOF demodulation, or independently for other gated-imaging applications (e.g. FLIM, FCS [21], or gated STED [22]). This is achieved through a proper gating scheme; indeed, as shown by Eq. (6) and (7), for computing phase shift $\Delta\varphi$ and received signal A_R, it is not required to know all four different samples (C_0, C_1, C_2 and C_3), but only their differences ($C_3 - C_1$ and $C_0 - C_2$). Furthermore, the computation of background intensity B requires only the sum of the four samples $(C_0 + C_1 + C_2 + C_3)$. As a matter of fact, one of the three counters is always enabled and integrates the background light, thus storing 2-D intensity information; the two remaining 9-bit bidirectional counters are alternately enabled (via *Direct* pin) to perform up/down counting and accumulate differential counts.

In our implementation, we set the T_{TAP}/T_P ratio to 25%: although this value represents a sub-optimal choice (Fig. 4), it still gives a precision very close to the best value, with the main advantage of allowing four separated integration windows within the same period and to accumulate the four samples in the same frame without missing any incoming photons, thus increasing acquisition speed. Furthermore, this choice benefits by a simplified in-pixel electronics and timing management, since both counters are controlled by only one main clock derivation. The described implementation favorably compares to the results reported in [15], where separate gate integrations, each 50% T_P long, were performed in four separated frames with a consequent loss of photons and longer overall acquisition time. In terms of speed and precision our approach is also advantageous as compared to [13], whose in-pixel digital circuitry is conceived with a 2-to-1 multiplexer and two 8-bit counters operating in interlaced counting, requiring two readout frames. Finally, thanks to up-down counting, we attain in-pixel real-time background suppression, thus further extending counting range.

B. Array Chip

The chip is composed by a 64 $\times$ 32 array of the aforementioned pixels, row and column access circuitry, pipelined multiplexers, and global electronics [20] (see Fig. 6 for the schematic and working operation). Access circuitry consists of shift registers, which allow sequential addressing of pixel data, multiplexers scan the column bit-lines, and global electronics handles clock management, data readout, and array initialization.

Readout is performed on both upper and lower end of the pixel array to reduce the minimum frame integration time and the column bus capacitance. Thanks to in-pixel memories, the array works in a fully parallel fashion: at the end of each frame, the samples accumulated by the counters are stored in three in-pixel latches and a new frame can be acquired while the previous one is read out (global shutter readout). In this way, in case of fast scenes, the acquired image does not undergo deformation (*jello effect*) or motion artifacts, even in case of fast moving objects.

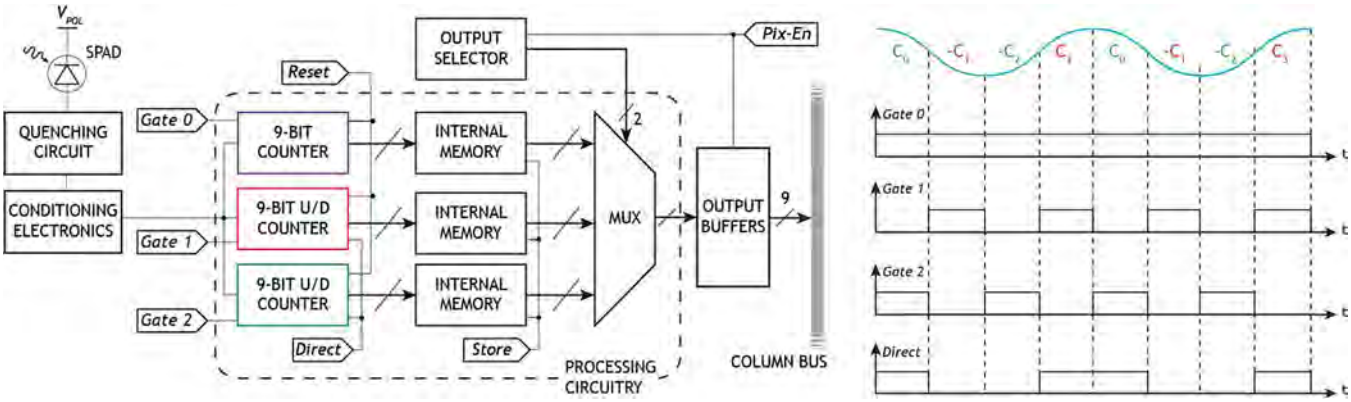

Fig. 6. Schematic block diagram (left) and working operation (right) of the 3-D iTOF pixel: all external signals come from an FPGA, so that counters are independently driven for easily implementing iTOF demodulation as well as gated imaging (for biological applications like FLIM, FCS [21], or gated STED [22]).

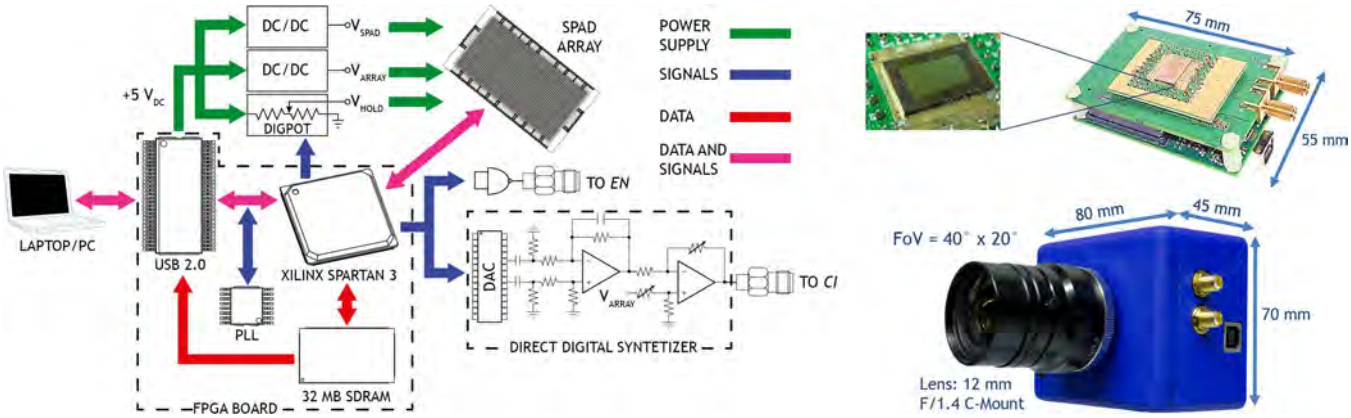

Fig. 7. Block diagram of the SPAD Camera (left). The EN and CI output signals are connected to the laser drivers (see Fig. 8). On the right, chip-on-board assembly and electronic board, and manufactured camera with optics and actual dimensions.

C. Camera

In order to operate the SPAD sensor chip, we developed a complete high-speed camera module, employing the SPAD array chip together with programmable electronics, optics, and software interface, as shown in Fig. 7.

The system electronics include a small form-factor ($75\text{ mm} \times 50\text{ mm} \times 16\text{ mm}$) XEM3010 board by OpalKelly [23] featuring an FPGA module by Xilinx (Spartan-3, XC3S1500-4FG320) that manages I/O timings and processes data coming from the chip, a high-speed USB 2.0 interface (Cypress FX2LP—CY68013A) enabling fast data transfer of pre-processed data to a remote PC, and a 32 MiB 16-bit wide SDRAM (Micron MT48LC1-6M16) for high-speed (100,000 fps) imaging. The imager chip is directly bonded on a second back-mounted board, through a chip-on-board (COB) assembly, which provides several related benefits, such as greater design flexibility, simpler manufacturing processes, more efficient heat dissipation, and a smaller board space with respect to standard or custom packaging, which would also raise costs. The second board accommodates two DC/DC converters—which generate the required power supply for the on-chip electronics (V_{ARRAY}, 3.3 V) and for the SPAD detectors (V_{SPAD}, 31 V) from the USB link—and a digital potentiometer to set the quenching hold-off time (V_{HOLD}) [20].

We also developed a mixed analog/digital circuitry—comprising a digital-to-analog converter (DAC), a current-to-voltage converter, a differential to single-ended amplifier and a variable-gain stage—to implement a direct digital synthesizer (DDS), thus providing arbitrary analog waveform modulation to the light source.

The camera is housed in a solid aluminum case supporting a 12 mm f/1.4 C-mount imaging lens, whose field-of-view is approximately $40^\circ \times 20^\circ$(H × V). The whole system is very rugged and compact, with dimensions of 80 mm × 70 mm × 45 mm and consumes about 1 W, mostly dissipated by the FPGA board (240 mA), with negligible contribution from the SPAD imager (10 mA).

A MATLAB interface is used for setting parameters (e.g., frame duration, number of frames to be acquired, modulation frequency) and for data acquisition and post-processing.

D. Laser Diodes Illuminator

As shown by Eq. (9), the distance error is strictly dependent on the active-light intensity; concurrently, eye safety hazard, cost, and consumption set other constraints. Therefore the optical power must be accurately balanced between all these factors. In our case, we designed a low power consumption

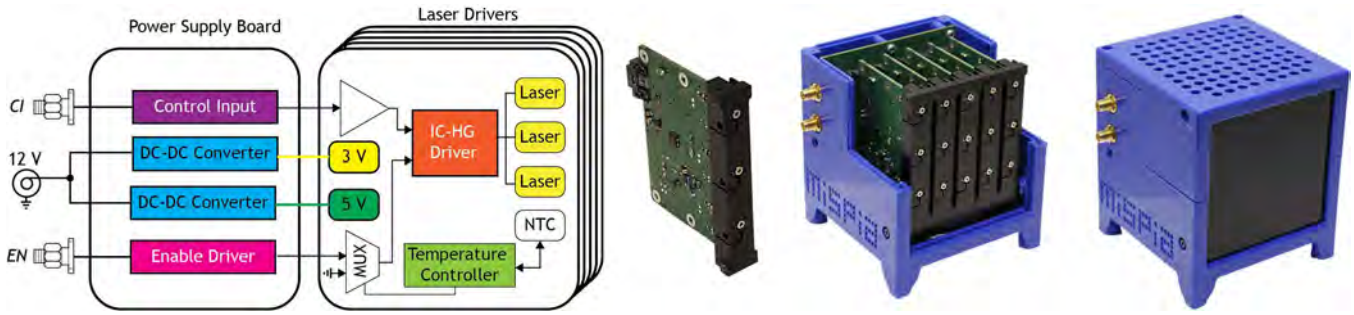

Fig. 8. Block diagram of the illumination source (left) and final assembly (right).

illumination source with 1.5 W optical power, that—thanks to the ultimate single-photon sensitivity of our SPAD imager—is able to provide good performance under average ambient conditions. Indeed, as proved in Appendix, an average optical power of 1.5 W suffices to range targets up to 40 m, with a precision better than 1 m, ensuring also eye safety (Class 1). Of course, a more powerful illuminator or a custom layout of laser sources (e.g. along the radiator grill, inside the headlights, or along the front bumper of a vehicle) can be customized to provide even longer distance ranges and higher precision, even under more demanding environments (e.g. under direct sunlight or with very opaque far-away objects).

The illuminator (Fig. 8) has a modular design based on a power supply board and five laser driver cards, each one mounting 3 laser diodes (LDs). The selected laser model (ADL-80Y04TZ-1) has relatively low threshold current (45 mA), high slope efficiency, and a peak CW power of 200 mW at 808 nm wavelength, and thus fifteen diodes achieve a total optical peak power of 3 W. Moreover, the selected LDs have a high operating temperature (50 °C) and a small TO-18 package, which reduces parasitic capacitance and improves modulation performance. A driver circuit by IC-Haus (IC-HG) was chosen to efficiently modulate lasers emission. The driver enables the switching of the lasers 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 Input (CI) is an analog voltage signal to control the LD current.

A temperature protection circuit, heat sinkers, and a fan are also used to lower and stabilize the temperature of the overall illuminator. Since the optical divergence of the laser diodes $(40^\circ \times 8^\circ)$ was not matched to the requested field-of-view (FOV) of about $40^\circ \times 20^\circ$, we used a beam diffusing sheet from Luminit, namely an LSD (Light Shaping Diffuser) with a circular 25° divergence angle, to achieve a $47^\circ \times 26^\circ$ FOV.

IV. System Characterization

A. Optical Components

We performed preliminary characterization of the illuminator. Each laser diode was operated at a maximum current of 225 mA, corresponding to an emitted peak power of 200 mW each. A low distortion sine wave was created by modulating LDs in the linear range of their I-V curve. In fact, the harmonic content of the light source is of great concern because, while CW technique rejects even harmonics, odd harmonics cause nonlinearity errors [16].

In order to experimentally evaluate the quality of the illumination waveform, a time-correlated single-photon counting (TCSPC) acquisition was performed at two frequencies (5 MHz and 8.333 MHz, as it will be explained later). Fig. 9 shows the FFT of the measured waveforms and the relative power spectral density: for the lower modulation frequency (5 MHz), the 3rd harmonic power is 40 dB lower than the fundamental frequency; this value is slightly higher (35 dB) for the 8.333 MHz modulation, whose spectral content shows also 5th and 7th harmonics, nonetheless with power at least 50 dB lower than the fundamental.

The plots in Fig. 9 show that optical waveforms have a nearly unitary modulation contrast (c_m), defined for a pure sinusoidal function as the ratio between the DC value and the fundamental frequency amplitude [24]:

$$c_m = A_0/A_{\mathrm{DC}}. \tag{11}$$

To understand the importance of this parameter, let us consider the ideal situation with no background light: if the modulation contrast is not unitary, a background equal to $A_{\mathrm{DC}} - A_0 = A_{\mathrm{DC}} \cdot (1 - c_m)$ is created at the source. Since the received signal is a replica of the emitted light, scaled by a factor k dependent on the object's reflectivity and distance [16], Eq. (9) can be written as:

$$\begin{aligned}\sigma_{d,\mathrm{CW}} &= \frac{d_{\mathrm{MAX}}}{k \cdot A_0} \cdot \sqrt{\frac{k \cdot A_0 + k \cdot A_0/c_m}{2 \cdot N}} \cdot \frac{1}{2\pi \cdot F(x)} \\ &= \sigma_{d,ID} \cdot \sqrt{\frac{1 + c_m}{2 \cdot c_m}}.\end{aligned} \tag{12}$$

Therefore, even in the ideal case of no background light, the precision of the distance measurement lowers as the modulation contrast decreases. The modulation contrast can be measured by taking into account the relationship between the power of an ideal sine wave $(P_{ID} = -3\,\mathrm{dB})$ and the measured power of the fundamental frequency $(P_{M,0})$, both normalized with respect to the power of the DC component:

$$P_{M,0} = P_{ID} \cdot c_m^2 \Rightarrow c_m = \sqrt{P_{M,0}/P_{ID}} = 10^{\left(\frac{PSD - 3\mathrm{dB}}{20\mathrm{dB}}\right)}. \tag{13}$$

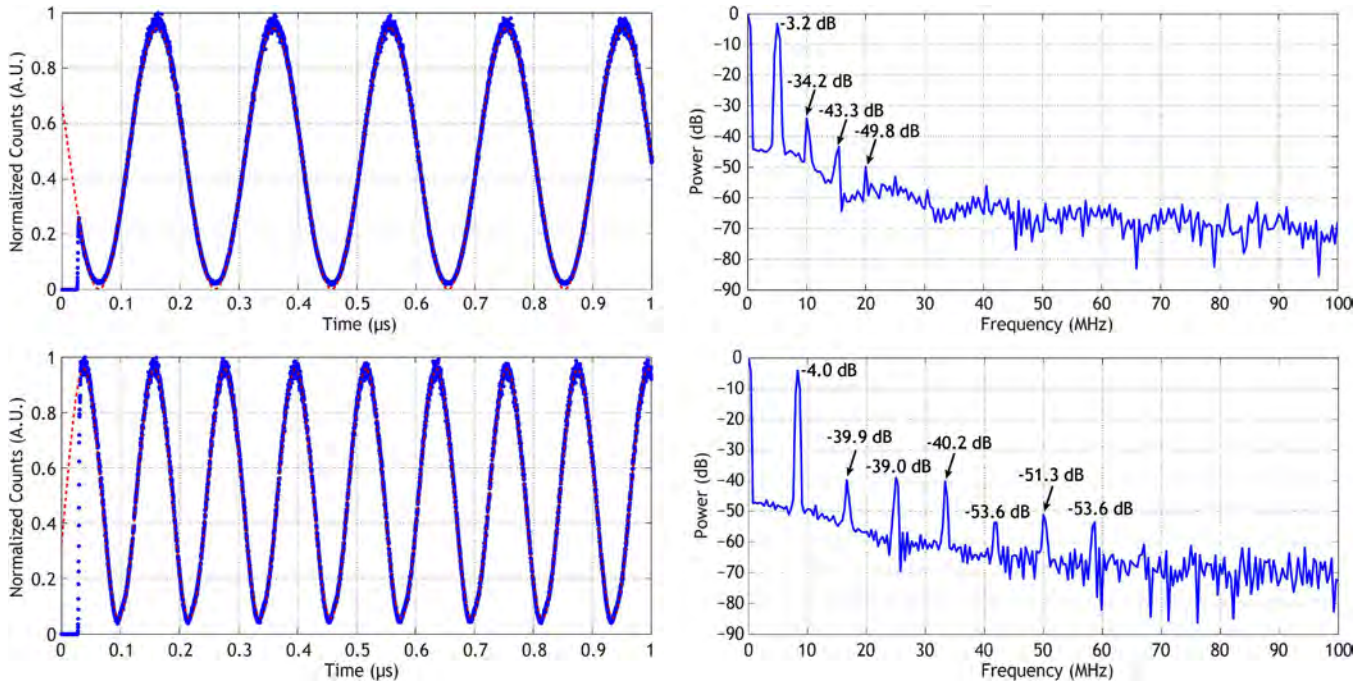

Fig. 9. Optical waveform (blue) and sine fitting (dotted red) of the illumination emission, as acquired through a TCSPC technique (left). The Fourier analysis shows a very neat spectral content for the 5 MHz waveform (top), while the 8.333 MHz one (bottom) shows more harmonics, but better ($>$ 35 dB) spurious-free dynamic range.

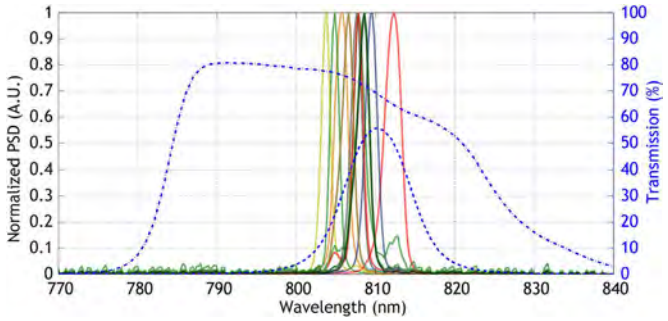

Fig. 10. Spectra of the 15 laser diodes (bold lines) and transmission curves for the two filters considered for the system (dashed lines).

Thus, we get $c_m = 97.7\%$ at 5 MHz and $c_m = 89.1\%$ at 8.333 MHz, corresponding to a percent precision loss of 0.6% and 3% respectively, compared to the ideal case of unity c_m.

Finally, we measured the emission of each laser by means of a spectrometer (HR400). Illumination spectra are plotted in Fig. 10 together with transmission curves of two band-pass filters (by Thorlabs): one filter with center wavelength at 800 nm and 40 nm width (FB800-40) and the other one with center wavelength at 810 nm and 10 nm width (FB810-10). The former filter has a high transmission efficiency ($>$ 70%) over the signal wavelength range but even higher transmission efficiency (80%) between 770 nm and 800 nm, where only background light is collected; conversely, the FB810-10 is more effective in blocking background light, but simultaneously it filters out a grater amount of laser power. As a matter of fact, none of the considered filters represents the best choice and a large margin of improvement in signal collection and background rejection still exists. Nonetheless, we decided to use the FB800-40, since it gave highest SNR in our evaluation setup.

B. Precision and Accuracy of the Rangefinder

As shown by Eq. (9), when the distance range increases, the standard deviation increases proportionally, so that precision is reduced. Therefore, in order to achieve the required 40 m range without degrading system performance, we adopted a double-frequency continuous wave (DFCW) modulation [25], [26], where two frames are acquired using two different modulation frequencies, thus extending the maximum non-ambiguous range (d_{MAX}) to [25]:

$$d_{\mathrm{MAX}} = \frac{c}{2 \cdot |f_1 - f_2|}. \tag{14}$$

Basically, each frequency gives a set of possible object locations (aliasing), but only at one location the two are in agreement. At that point, relative contributions can be weighted by relative modulation frequencies and signal magnitudes [26]:

$$d_{\mathrm{final}} = d_1 \cdot \frac{A_1 \cdot f_1}{A_1 \cdot f_1 + A_2 \cdot F_2} + d_2 \cdot \frac{A_2 \cdot f_2}{A_1 \cdot f_1 + A_2 \cdot f_2}. \tag{15}$$

The key benefit here is the possibility to use higher modulation frequencies, implying higher precision, but still achieving long distance ranges. In our case, the distance range required was 40 m, so we implemented the DFCW technique by using 8.333 MHz (18 m range) and 5 MHz (30 m range) modulation frequencies to achieve a distance range of 45 m.

We performed static tests by varying the distance (2 – 40 m) between the camera and a targeted panel (80% reflectivity). Measurement accuracy is shown in Fig. 11: a maximum non-linearity of 1 m is achieved for the higher frequency; whereas the lower modulation—having an overall lower harmonic content (Fig. 9)—exhibits a maximum non-linearity of about 0.8 m.

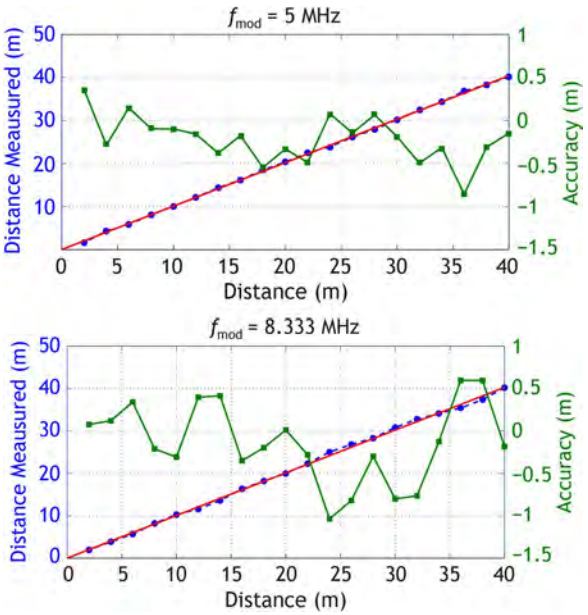

Fig. 11. Measured distance (blue dash line) and accuracy (green) computed as the difference between true (solid red) and measured distance.

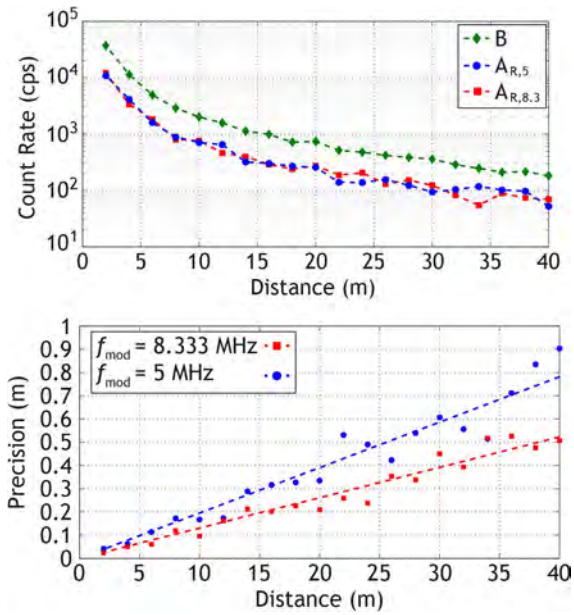

Fig. 12. Acquired reflected active-light (top) for both 5 MHz ($A_{R;5}$) and 8.333 MHz ($A_{R;8}$) modulations, together with background intensity (B) as a function of target distance. Precision vs. distance and linear fitting curves (bottom).

Fig. 12 illustrates the acquired reflected signal (A_R) and the background intensity (B), about 6 times higher than A_R. Nevertheless, precision at 40 m is 90 cm for the 5 MHz modulation and 50 cm for the 8.333 MHz one. Although the precision curve follows an almost linear trend, some perturbations are present, caused by multiple light reflections, due to concavities in the scene [27]. When the information from the two modulations is combined according to Eq. (15), the resulting precision gets better than 60 cm. This value is inferred from real measurements, by applying the propagation error rule to Eq. (15).

C. *Effects of Multiple Cameras*

We made some tests to study the interference between identical cameras, as it would happen in an ordinary traffic scenario. For these tests, we acquired 200 frames at 100 fps, with a second camera positioned within the scene as shown in Fig. 13. The "single camera" represents the reference condition, where only the primary 3-D camera is used; in case (a) both cameras image the scene from a similar position (as in the case of one vehicle overtaking the other one); in case (b) the two cameras almost face one another (as two front-side approaching vehicles).

Compared to reference situation, case (a) provides brighter background image, but active-light acquisition, distance measurement, and precision are comparable. On the other hand, in case (b), the direct exposure to disturbing active illumination produces a brighter background image and causes saturation where the disturbing illuminator is located (right-hand side). Nonetheless, active-light and distance images show negligible differences—besides clearly displaying the presence of the disturbing camera—and precision is just slightly impaired.

This rejection of disturbances from other cameras comes from the fact that different cameras' clocks, although running at same nominal frequencies, are not correlated. Therefore, the disturbing illumination contributes as a common-mode signal that is canceled through the homodyne demodulation performed by the up-down counters within each pixel. Since distance and amplitude information are related to the content of the up-down counters, the related data are not affected by inaccuracy (excluding those pixels where saturation prevents a correct demodulation, thus jeopardizing the measurement only therein). Finally, the slightly decreased precision of case (b) is explained by Eq. (9): in case (a) the disturbing illumination causes minor background increment, hence a negligible effect on the measurement repeatability. Conversely, in case (b), the disturbing illumination considerably increases the total background light, thus degrading precision.

V. Outdoor Tests

In order to assess the correct SPAD camera operations, the entire system (SPAD camera and illumination source) was installed on a car, as shown in Fig. 14, together with a standard action camera (Hero3 by GoPro) for co-registration. Internal installation was not possible since most of current windscreens include a thin film of reflective material that blocks infrared radiation to ensure cabin climate comfort. Of course, in a custom installation such film could be removed from a small optical window just in front of the illuminator and camera, as shown in [3]. All measurements were performed between November and January during afternoon and evening hours, also in adverse (foggy and light rain) weather conditions.

The first tested scenario was an underground parking lot, where we could easily verify if the illumination power was sufficient to properly light targets up to 40 m distance. A 3-D frame and an RGB image from this scenario are represented in Fig. 15, where it is shown that all the targets in the scene—a pedestrian (at 7 m), many pillars located at different distances

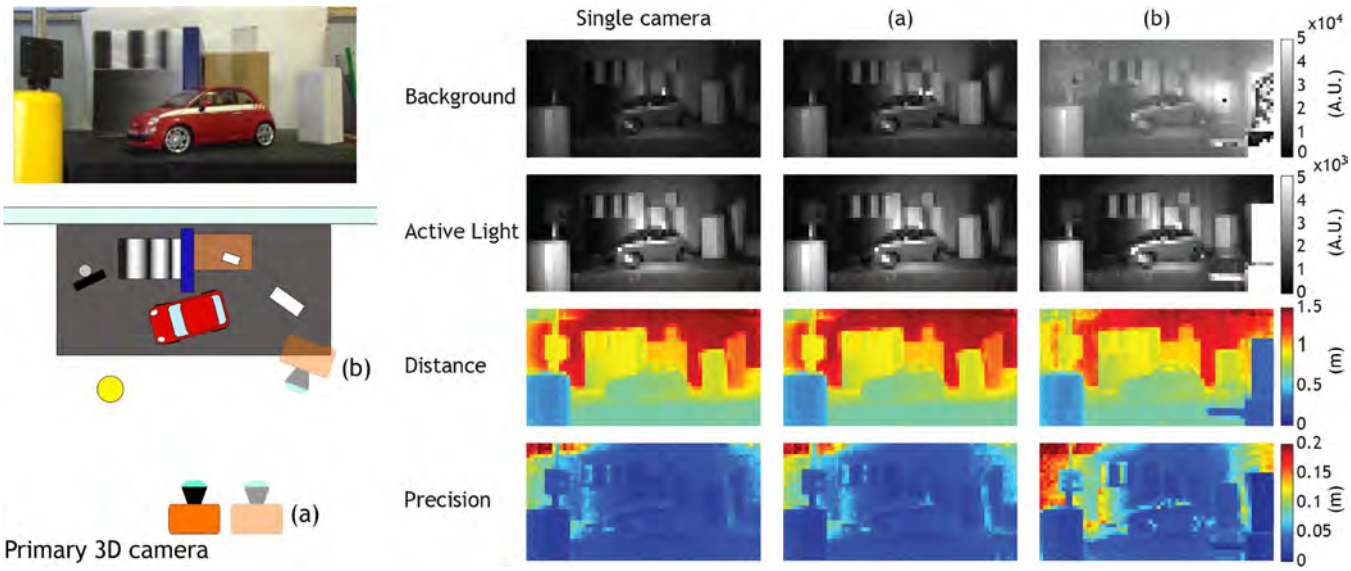

Fig. 13. Multi-camera scenarios: (a) two side-by-side cameras lit the same scene; (b) the second camera crosses the illumination of the primary one. Acquisitions are compared with those of the single camera scenario. Thanks to in-pixel demodulation, even in case of strong disturbance, the background is fully rejected (see the active light images) and the measured distance is not affected. Instead, precision degrades as predicted by Eq. (9).

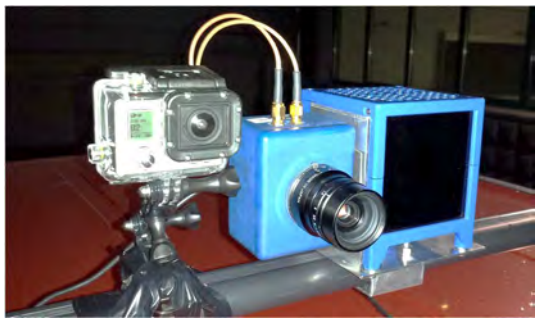

Fig. 14. 3-D SPAD camera and illuminator installed together with a standard action camera on a car, for outdoor test in real driving environments.

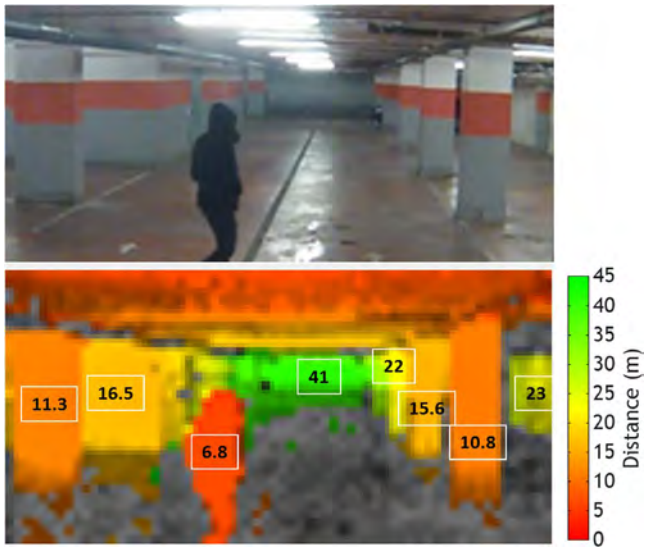

Fig. 15. Indoor scene (top) and 3-D frame (bottom) showing unaliased depth ranging up to 41 m, thanks to DFCW.

and a very far away wall at 41 m—are clearly and correctly ranged with no evident aliasing error. As a second step, we moved into a real traffic scenario: Fig. 16 displays the co-registered

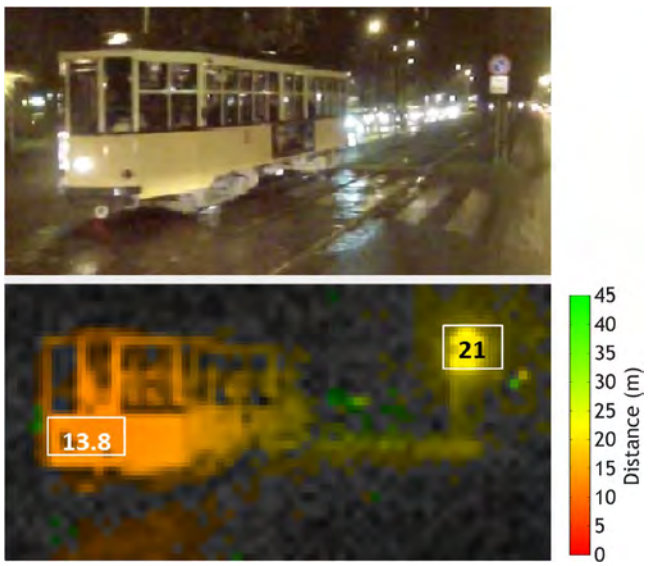

Fig. 16. Frame showing a tram passing by, while the test car is turning left.

real scene and the respective frame from a 3-D video acquired at 100 fps, i.e. 10 ms frame time, of a tram passing by, while the car turns left. The 2048 pixel resolution of the iTOF camera is enough to discriminate the tram from other vehicles (e.g., cars or trucks) and can be easily detected through further post-processing [18].

Fig. 17 shows another scene with different objects clearly detected and ranged by our 3-D system: a fast vehicle passing by very close to the camera (5 m); in the background, a tree (16 m), a concrete pillar (23 m) and nearby (22 m) a pedestrian in front of a parking boom gate. The speed of the crossing car is estimated (from its displacement between consecutive frames) to be 40 km/h. As the test car was traveling at urban speed limit (50 km/h) different videos were recorded.

In Fig. 18 a 3-D frame obtained with 10 ms integration time shows the detection of a van—proceeding 18 m ahead at the

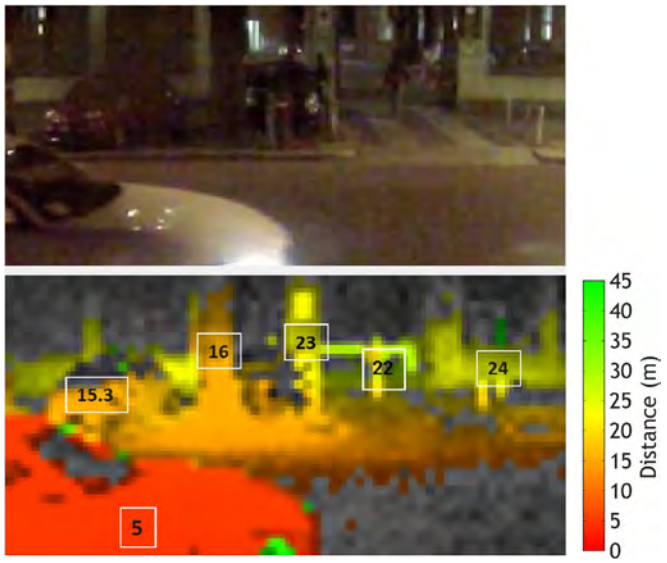

Fig. 17. Frames from a 3-D video where a vehicle is moving fast in front of the car. Note also pedestrian (at 22 m), a pillar (at 23 m), a tree (at 16 m) and a boom gate beyond the pedestrian.

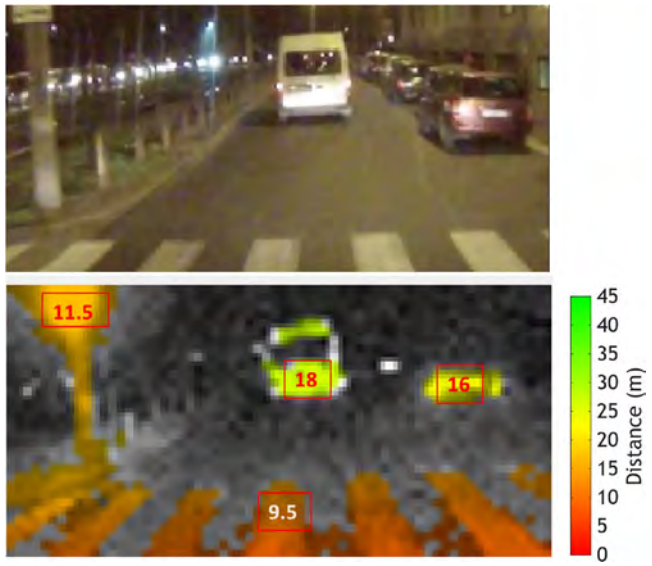

Fig. 18. Frame from a 100 fps 3-D video recorded when the car was following a van at 50 km/h. Note the ability of the camera to acquire the zebra crossing pattern at about 9.5 m distance and also the narrow pole at 11.5 m.

same speed of the test car—and the ability of the 3-D camera to acquire the distance of the zebra crossing, at about 9.5 m distance, and also the narrow (about 25 cm thick) lamp post at 11.5 m.

Eventually, Fig. 19 shows several frames from a 3-D movie at 100 fps: in the first frame a car (in red at 5 m) is driving into a lane where another car (in orange at 9.5 m) is travelling while a third car (in yellow-orange at 12 m) is parking. Other six frames were taken, from the same movie: two pedestrians cross the street in the direction of a parked car (top left frame); then a car passed by from left to right; finally a third car, moving from right to left, enters the field of view (bottom right frame). Since frames are taken one out of 100, the estimation of the second car's speed is 6 km/h.

From the previous pictures, it is clear that the illumination system is able to uniformly illuminate the scene without shadow zones and, although the camera resolution is limited to 2048 pixels, the integration of all information provided by the SPAD sensor permit to generate images rich of details, thus allowing to easily locate and recognize objects in the scene through real-time image processing. A video, showing all the potentialities of the automotive 3-D SPAD camera is available here [35].

VI. Comparisons With Other 3-D Rangefinders

Table I shows a comparison with current rangefinders on the market, whose datasheets are publicly accessible. Only the Continental product is suited for automotive application, while other vendors' cameras are mostly designed for indoor or gaming applications. Since distance precision depends on the amount of photons acquired within a frame, the performance of each camera depends on a high number of system parameters (number of pixels, illumination power, optics, background, distance range, detector efficiency, etc.) and a 1:1 comparison is not easy—as many parameters are unknown—or even unfair. With respect to other off-the-shelf imagers, the developed 3-D SPAD camera has lower pixel count, but other cameras can handle mostly indoor scenarios or short-range outdoor scenes, while the SPAD camera handles very different ambient light conditions, both indoor and outdoor. We also tested the camera under direct 33,000 lux: at 100 fps the SPAD camera is still able to locate a pedestrian at 5 m, while at longer distances the return active-light signal was overwhelmed by background (background photons were 250 times as much as reflected photons). As discussed in Section III-D and in Appendix, better performance—also in sunlight and harsher environmental conditions—can be achieved both by an optimized custom positioning of lasers on the vehicle and by increasing illumination power, according to given customer specifications and other automotive constraints.

The FOV of $40^\circ \times 20^\circ$ is comparable or lower than the others, but a fair judgment should take into account the full distance range, which in our case reaches 40 m in low ambient light condition, compared to 10 – 13.5 m (maximum) provided by other rangefinders.

Finally, the simultaneous acquisition of 2-D intensity movies together with 3-D video represents another major advantage and innovative aspect of our SPAD camera, together with the highest frame-rate (100 fps) and the achieved 60 cm precision at 40 m. The power consumption is one of the lowest reported, and this is particularly remarkable considering the operating frame-rate and the fairly wide $40^\circ \times 20^\circ$ field-of-view covering up to 40 m range. For average level light conditions, like those here reported, the power consumption of the overall camera is about 200 mA from the 5 V USB 2.0 connector (i.e. 1 W), plus other 3 W for the illuminator.

VII. Conclusion

We have presented a 3-D vision system based on a Single-Photon Avalanche Diode (SPAD) detector array chip, manufactured in a cost-effective 0.35 μm automotive-certified high-voltage CMOS technology. Since SPAD detectors are able to detect and count single-photons in the near-infrared

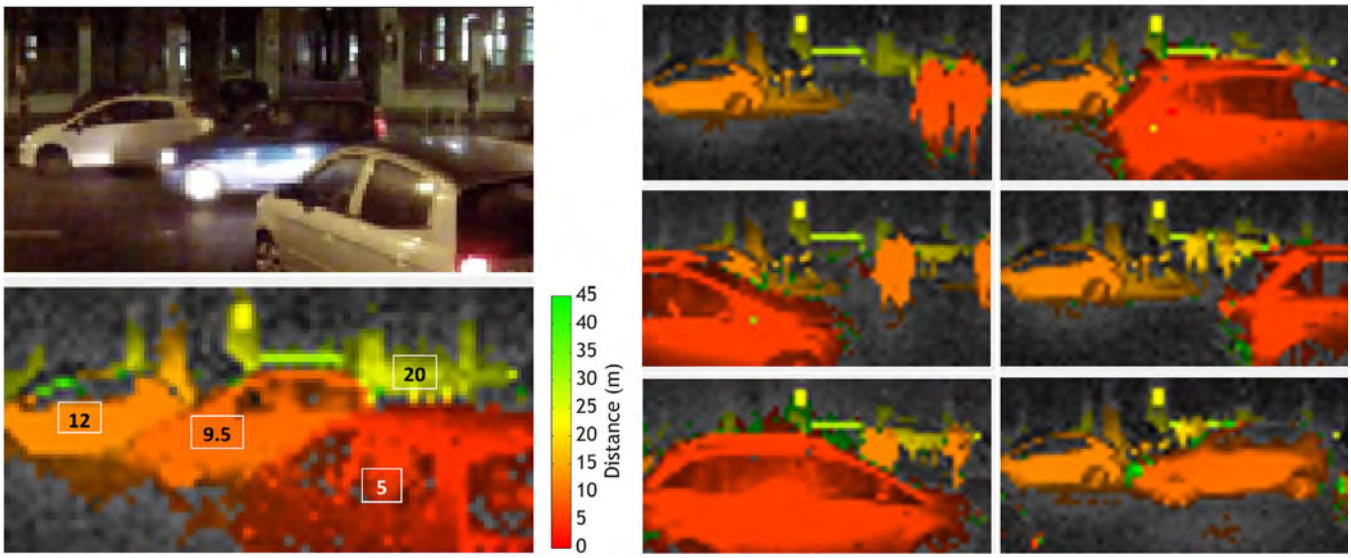

Fig. 19. Left: frames from a 3D movie at 100 fps: one car (in red at 5 m) is entering into a lane where another car (in orange at 9.5 m) is moving fast and a third car (in yellow-orange at 12 m) is parking. Right: frames from the same 3-D movie acquired in a chaotic situation, with traveling cars and pedestrians rushing to cross the road. Frames were taken at about 1 s from each other, i.e. about one frame out of 100.

TABLE I
COMPARISON WITH OFF-THE-SHELF 3-D CAMERAS

Parameters	Real. iZ-1K Vision system	PMD Cam Board Nano	Mesa Imaging SR4500	SoftKinetic DS325	Panasonic D-Imager	Continental SRL1	**This Work**
Ref.	[28]	[29]	[30]	[31]	[32]	[33]	**-**
Num. of Pixels	1.3 M	19200	25344	76800	19200	3	**2048**
Technique	dTOF	CW iTOF	CW iTOF	CW iTOF	CW iTOF	dTOF	**CW iTOF**
Pixel Size	14 µm × 14 µm	45 µm × 45 µm	40 µm × 40 µm	30 µm × 30 µm	-	-	**150 µm × 150 µm**
Sensor Process	CMOS	CMOS	CCD/CMOS	CMOS	CCD	-	**CMOS**
Wavelength	905 nm	850 nm	850 nm	850 nm	850 nm	905 nm	**808 nm**
Illumination	LD	LED	LED	diffused laser	LED	Laser	**LD**
Max Range	10 m	2 m	9 m	1 m	5 m	13.5 m	**40 m**
Precision	1 cm	5 mm	9 mm	1.4 cm	14 cm	5 mm	**60 cm**
Typical Frame Rate	30	90	30	60	30	100	**100**
Field of View	50° × 50°	90° × 68°	69° × 56°	74° × 58°	60° × 44°	27° × 11°	**40°×20°**
Power Consumption	< 15 W	2.5 W	12 W	<2.5 W	8.4 W	1.8 W	**4 W**
Use	In/Out	In/Out	In	In	In	Out	**In/Out**

wavelength range, the active illuminator can be designed to output very low power eye-safe light. Each pixel of the 64×32 SPAD imager can acquire 3-D information from the scene under observation, through indirect time-of-flight measurement, i.e., by counting photons in time-slots synchronized with the active illumination, which can be either pulsed (PL-iTOF) or continuous-wave (CW-iTOF), although we demonstrated that CW-iTOF exhibits much better performance.

We validated the 3-D camera in real outdoor automotive scenario under low ambient light condition and we acquired 3-D maps of objects located up to 41 m away, within a field-of-view of $40^\circ \times 20^\circ$, with better than 1 m precision at the farthest distance. We also used the 3-D system under direct sunlight and we could distinguish a pedestrian at short (about 5 m) distance. Therefore, with the presented low (1.5 W optical average) power illuminator, the camera proved to be a viable system for vehicle automation in indoor settings, such as robotic driven industrial carts, and also in outdoor applications at mild background (i.e., not full sun light exposure) levels and definitely at low ambient light. In order to extend its usefulness to all vehicular applications, proper customization of illuminator power, vision field-of-view, layout of lasers on the specific vehicle, and operating wavelength are required, to match the target application.

Finally, we compared the performance and the benefits of the SPAD-based 3-D system with respect to other commercially available product, most of them not suitable for automotive applications. For the future we envision further developments aimed at improving camera functionalities and performances. As an example, object recognition will be implemented within the FPGA in order to have a real standalone system able to communicate over CAN bus to fulfill functional requirements of a collisions mitigation system.

APPENDIX

During the design of a laser light source, the Maximum Permissible Exposure (MPE) must not be exceeded: this is the maximum irradiance or radiant exposure that may be incident upon the eye (or the skin) without causing biological damage. To determine the MPE for a pulsed laser source, the emission wavelength (λ), the pulse period (T_{PERIOD}), the

duration of a single pulse (T_{PULSE}), and the total exposure time (T_{EXPOSURE}) must be known. Moreover, three limit values must be checked to determine the lower MPE: the single pulse ($\mathrm{MPE}_{\mathrm{PULSE}}$), the repetitive pulse ($\mathrm{MPE}_{\mathrm{TRAIN}}$), and the average power ($\mathrm{MPE}_{\mathrm{AVERAGE}}$) limit. As discussed in Section IV-B in the text, our laser source has a minimum pulse duration (T_{PULSE}) of 60 ns, hence [34]:

$$\mathrm{MPE}_{\mathrm{PULSE}} = 5 \cdot 10^{-3} \cdot C_A \cdot C_E \cdot J/\mathrm{m}^2 \tag{16}$$

where C_A is a constant dependent on wavelength and C_E depends on the angle subtended by the laser source and are equal to [34]:

$$C_A = 10\frac{\lambda - 700\ \mathrm{nm}}{500\ \mathrm{nm}} = 1.64 \tag{17}$$

$$C_E = 100\ \mathrm{mrad}/1.5\ \mathrm{mrad} = 66.7 \tag{18}$$

Concerning the illumination wavelength, we selected $\lambda =$ 808 nm as a trade-off between eye safety constraints and CMOS SPAD detection efficiency. Therefore, from Eq. (16) we obtain [34]:

$$\mathrm{MPE}_{\mathrm{PULSE}} = 0.55 \cdot \mathrm{J}/\mathrm{m}^2. \tag{19}$$

The $\mathrm{MPE}_{\mathrm{TRAIN}}$ value can be obtained from Eq. (19), by knowing the number of pulses for a given exposure time [34]:

$$\mathrm{MPE}_{\mathrm{TRAIN}} = \mathrm{MPE}_{\mathrm{PULSE}} \cdot N_{\mathrm{PULSE}}^{-1/4} \tag{20}$$

where N_{PULSE} is [34]:

$$N_{\mathrm{PULSE}} = T_{\mathrm{EXPOSURE}}/T_{\mathrm{PERIOD}} \tag{21}$$

being T_{PERIOD} equal to 120 ns, in the worst case and considering a limiting exposure time of intentional viewing equal to 100 s, as suggested by [34], we get:

$$\mathrm{MPE}_{\mathrm{TRAIN}} = 3.2 \cdot \mathrm{mJ}/\mathrm{m}^2. \tag{22}$$

Eventually, to compute the average power limit, we consider the scene as lit by a single pulse with duration equal to the exposure time. The maximum irradiance for such a light pulse ($\mathrm{MPE}_{\mathrm{MAX}}$), at 808 nm emission and still a 100 s exposure time, is given by [34]:

$$\mathrm{MPE}_{\mathrm{MAX}} = 18 \cdot C_A \cdot C_E \cdot T_{\mathrm{EXPOSURE}}^{3/4} \cdot \mathrm{J}/\mathrm{m}^2 = 62.2 \cdot \mathrm{kJ}/\mathrm{m}^2. \tag{23}$$

From Eq. (23), $\mathrm{MPE}_{\mathrm{AVARAGE}}$ can be obtained as follows [34]:

$$\mathrm{MPE}_{\mathrm{AVERAGE}} = \frac{\mathrm{MPE}_{\mathrm{MAX}}}{N_{\mathrm{PULSE}}} = 74.6 \cdot \mu\mathrm{J}/\mathrm{m}^2 \tag{24}$$

which gives the limit for the illumination irradiance [34]:

$$E_{\mathrm{MAX}} = \frac{\mathrm{MPE}_{\mathrm{AVERAGE}}}{T_{\mathrm{PULSE}}} = 1.25 \cdot \mathrm{kW}/\mathrm{m}^2. \tag{25}$$

This value must be compared with the irradiance of the light source. To this purpose, we considered our illuminator as a point-like source, with $47^\circ \times 26^\circ$ divergence. The source irradiance is not constant, but depends on the overall lit area, which quadratically increases with distance. By referring to Fig. 20, we compute the irradiance as:

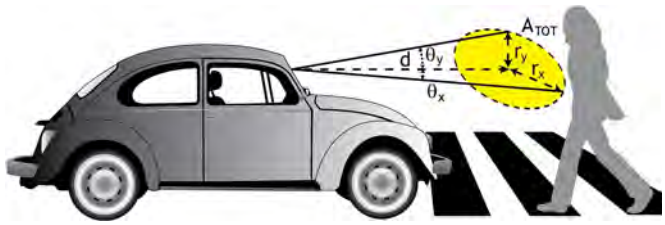

Fig. 20. Schematization of light beam exposure for a pedestrian crossing the road.

$$E_{\mathrm{SOURCE}} = \frac{P_{\mathrm{SOURCE}}}{A_{\mathrm{TOT}}} = \frac{P_{\mathrm{SOURCE}}}{\pi \cdot r_x \cdot r_y} = \frac{P_{\mathrm{SOURCE}}}{\pi \cdot \tan(\theta_x) \cdot \tan(\theta_y) \cdot d^2} \tag{26}$$

with $\theta_x = \mathrm{FOV}_x/2 = 23.5^\circ$ and $\theta_y = \mathrm{FOV}_y/2 = 13^\circ$. By comparing E_{MAX}, Eq. (25), and E_{SOURCE}, Eq. (26), we can compute the nominal hazardous zone (NHZ) that is the distance within which the irradiance of a beam is greater than the MPE. Since our 3-D system was mounted on the top of a vehicle near the windscreen (see Fig. 20), the NHZ should be less than the length of a car bonnet (about 0.8 m) so that pedestrians do not incur any damage resulting from intentional staring at the lasers for less than 100 s. Nonetheless, since we also tested the camera indoor, we chose an even more conservative value ($\mathrm{NHZ} < 0.1$ m) to ensure eye safety even in case of very close (10 cm) proximity to the laser source:

$$E_{\mathrm{SOURCE}} = \frac{P_{\mathrm{SOURCE}}}{\pi \cdot \tan(\theta_x) \cdot \tan(\theta_y) \cdot \mathrm{NHZ}^2} < E_{\mathrm{MAX}}. \tag{27}$$

Since $E_{\mathrm{MAX}} = 1.25 \cdot \mathrm{kW}/\mathrm{m}^2$, we get $P_{\mathrm{SOURCE}} < 4$ W.

From the computed source irradiance, we can compute the associated radiance L_{SOURCE}, assuming the worst-case case when the viewer is looking directly into the beam [34]:

$$L_{\mathrm{SOURCE}} = \frac{4E_{\mathrm{SOURCE}}}{\pi\theta_x\theta_y} = 13\frac{\mathrm{kW}}{\mathrm{m}^2\mathrm{sr}}. \tag{28}$$

The source radiance is lower than the maximum radiance of a diffused source for Class 1, which is $L_{\mathrm{CL},1} = 76,576\ \mathrm{W}/(\mathrm{m}^2 \cdot \mathrm{sr})$ for the considered wavelength range [34].

REFERENCES

[1] D. Cualain, M. Glavin, E. Jones, and P. Denny, "Distance detection systems for the automotive environment: A review," in *Proc. ISSC*, Derry, Ireland, Sep. 13/14, 2007, pp. 1–6.

[2] R. Rasshofer and K. Gresser, "Automotive radar and lidar systems for next generation driver assistance functions," *Adv. Radio Sci.*, vol. 3, no. 10, pp. 205–209, May 2005.

[3] S. Hussmann, T. Ringbeck, and B. Hagebeuker, "A performance review of 3D TOF vision systems in comparison to stereo vision systems," in *Proc. Stereo Vision*, Vienna, Austria, 2008, pp. 103–120.

[4] D. Piatti, F. Remondino, and D. Stoppa, "State-of-the-art of TOF range-imaging sensors," in *TOF Range-Imaging Cameras*. Berlin, Germany: Springer-Verlag, 2013, pp. 1–9, 2013.

[5] B. F. Aull *et al.*, "Geiger-mode avalanche photodiodes for three-dimensional imaging," *Lincoln Lab. J.*, vol. 13, no. 2, pp. 335–350, 2002.

[6] C. Niclass, C. Favi, T. Kluter, M. Gersbach, and E. Charbon, "A 128 × 128 single-photon image sensor with column-level 10-bit time-to-digital converter array," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2977–2989, Dec. 2008.
[7] C. Veerappan *et al.*, "A 160 × 128 single-photon image sensor with on-pixel 55ps 10b time-to-digital converter," in *Proc. ISSCC Dig. Tech. Papers*, 2011, pp. 312–314.
[8] M. Gersbach *et al.*, "A time-resolved, low-noise single-photon image sensor fabricated in deep-submicron CMOS technology," *IEEE J. Solid-State Circuits*, vol. 47, no. 6, pp. 1394–1407, Jun. 2012.
[9] C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, "A 100-m Range 10-Frame/s 340×96-pixel time-of-flight depth sensor in 0.18-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 48, no. 2, pp. 559–572, Feb. 2013.
[10] F. Villa *et al.*, "CMOS imager with 1024 SPADs and TDCs for single-photon timing and 3-D time-of-flight," *IEEE J. Sel. Topics Quantum Electron.*, vol. 20, no. 6, pp. 364–373, Nov. 2014.
[11] D. Stoppa *et al.*, "A 32 × 32-pixel array with in-pixel photon counting and arrival time measurement in the analog domain," in *Proc. 35th ESSCIRC*, 2009, pp. 204–207.
[12] D. Stoppa *et al.*, "A CMOS 3-D imager based on single photon avalanche diode," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 1, pp. 4–12, Jan. 2007.
[13] C. Niclass, C. Favi, T. Kluter, F. Monnier, and E. Charbon, "Single-photon synchronous detection," *IEEE J. Solid-State Circuits*, vol. 44, no. 7, pp. 1977–1989, Jul. 2009.
[14] R. J. Walker, J. A. Richardson, and R. K. Henderson, "A 128×96 pixel event-driven phase-domain $\Delta\Sigma$-based fully digital 3D camera in 0.13 μm CMOS imaging technology," in *Proc. Dig. Tech. Papers—IEEE Int. Solid-State Circuits Conf.*, 2011, pp. 410–411.
[15] S. Bellisai, F. Guerrieri, and S. Tisa, "3D ranging with a high speed imaging array," in *Proc. Conf. PRIME*, Jul. 18–21, 2010, pp. 1–4.
[16] R. Lange, "3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology," Ph.D. dissertation, Dept. Electr. Eng. Comput. Scie. Univ. Siegen, Siegen, Germany, 2000.
[17] D. Stoppa *et al.*, "A range image sensor based on 10 m lock-in pixels in 0.18 μm CMOS imaging technology," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 248–258, Jan. 2011.
[18] M. Hansard, S. Lee, O. Choi, and R. P. Horaud, *Time-of-Flight Cameras, Principles, Methods and Applications*, ser. Springer Briefs in Computer Science. Berlin, Germany: Springer-Verlag, 2013.
[19] S. Bellisai *et al.*, "Single-photon pulsed-light indirect time-of-flight 3D ranging," *Opt. Exp.*, vol. 21, no. 4, pp. 5086–5098, Feb. 2013.
[20] D. Bronzi *et al.*, "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. 2014.
[21] M. Vitali *et al.*, "A single-photon avalanche camera for fluorescence lifetime imaging microscopy and Correlation spectroscopy," *IEEE J. Sel. Topics Quantum Electron.*, vol. 20, no. 6, pp. 344–353, Nov. 2014.
[22] G. Vicidomini *et al.*, "Gated CW-STED microscopy: A versatile tool for biological nanometer scale investigation," *Methods*, vol. 66, no. 2, pp. 124–130, Mar. 2014.
[23] OpalKelly XEM3010. [Online]. Available: http://www.opalkelly.com/products/xem3010/
[24] N. S. Kopeika, "Modulation contrast function," in *A System Engineering Approach to Imaging*. Bellingham, WA, USA: SPIE Pub., 1998, ch. 9, sec. 2, pp. 315–317.
[25] A. D. Payne, A. P. P. Jongenelen, A. A. Dorrington, M. J. Cree, and D. A. Carnegie, "Multiple frequency range imaging to remove measurement ambiguity," in *Proc. Opt. 3-D Meas. Techn.*, Jul. 2009, pp. 139–148.
[26] A. P. P. Jongenelen, D. A. Carnegie, A. D. Payne, and A. A. Dorrington, "Maximizing precision over extended unambiguous range for TOF range imaging systems," in *Proc. IEEE I2MTC*, May 2010, pp. 1575–1580.
[27] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sens. J.*, vol. 11, no. 9, pp. 1917–1926, Sep. 2011.
[28] Odos imaging VS-1000 vision system datasheet, (accessed Mar. 2015). [Online]. Available: http://www.odos-imaging.com/products/vs-1000/
[29] Pmd[vision]® CamBoard nano. (accessed Mar. 2015). [Online]. Available: http://pmdtec.com/products_services/reference_design.php
[30] MESA Imaging SR4500 datasheet, (accessed Mar. 2015). [Online]. Available: http://downloads.mesa-imaging.ch/dlm.php?fname=pdf/SR4500_DataSheet.pdf
[31] SoftKinetic DS325 datasheet, (accessed Mar. 2015). [Online]. Available: http://www.softkinetic.com/Portals/0/Documents/PDF/WEB_20130527_SK_DS325_Datasheet_V4.0.pdf
[32] Panasonic D-Imager EKL3106 datasheet, (accessed Mar. 2015). [Online]. Available: http://www2.panasonic.biz/es/densetsu/device/3DImageSensor/en/pdf/EKL3106_E.pdf
[33] Continental SRL1 datasheet, (accessed Mar. 2015). [Online]. Available: http://www.conti-online.com/www/download/industrial_sensors_de_en/themes/download/srl1_datasheet_en.pdf
[34] *British and European Standard BS EN 60825-1:2007*, Int. Std. IEC 60825-1, 2007. (accessed Mar. 2015). [Online]. Available: http://www.bristol.ac.uk/nsqi-centre/healthandsafety/nsqi-laser/lasersafetybs.pdf
[35] D. Bronzi, SP-ADAS: High-Speed Single-Photon Camera for Advanced Driver Assistance Systems, 2014, (accessed Mar. 2015). [Online]. Available: https://www.youtube.com/watch?v=OCBfMEG6X8o

Danilo Bronzi (M'14) received the B.Sc. degree (*cum laude*) in biomedical engineering, the M.Sc. degree (*cum laude*) in electronic engineering, and the Ph.D. degree from Politecnico di Milano, Milan, Italy, in 2008, July 2011, and November 2014, respectively. He is currently with Department of Optics and Atomic Physics, Technische Universität Berlin. His activity focuses on the design and development of optical transceivers for low-cost optical data center interconnects at 100–400 Gb/s.

Yu Zou received the M.Sc. degree in electronic engineering from Politecnico di Milano, Milan, Italy, in December 2012. She is currently working toward the Ph.D. degree in electronic engineering with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano. Her research activity focuses on the design and development of novel CMOS single-photon sensors for 2-D imaging and 3-D ranging applications.

Federica Villa (M'15) received the B.Sc. degree (*cum laude*) in biomedical engineering and the M.Sc. degree (*cum laude*) in electronic engineering from Politecnico di Milano, Milan, Italy, in 2008 and 2010, respectively. She is a Postdoctoral Researcher with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, working on advanced technologies for photon-counting and photon-timing applications. Ms. Villa received an award for her work in time-to-digital converters for direct time-of-flight 3-D imaging in 2014.

Simone Tisa received the M.Sc. degree in electronic engineering and the Ph.D. degree from Politecnico di Milano, Milan, Italy, in 2001 and 2006, respectively. He is currently an R&D Manager for electronics design and system integration with Micro Photon Devices SRL, Bolzano, Italy. In 2008, he pioneered the first monolithic 2-D single-photon detector (SPAD) imager of 32 × 32 pixels. His main research interests include the field of single-photon imaging and single-photon timing of fast phenomena by means of integrated arrays of SPADs and associated microelectronics.

Alberto Tosi (M'07) was born in Borgomanero, Italy, in 1975. He received the master's degree in electronics engineering and the Ph.D. degree in information technology engineering from Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively. In 2004, he was a Student with IBM T. J. Watson Research Center, Yorktown Heights, NY, USA, working on optical testing of CMOS circuits. Since 2014, he has been an Associate Professor of electronics with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano. Currently, he works on silicon and InGaAs/InP single-photon avalanche diodes (SPADs). He is involved in research on arrays of silicon SPADs for 2-D and 3-D applications and on time-correlated single-photon counting electronics.

Franco Zappa (M'00–SM'07) was born in Milan, Italy, in 1965. In 2004, he cofounded "Micro Photon Devices SRL," which is a company focused on the production and deployment of detection systems based on single-photon detectors (SPADs) and cameras for 2-D imaging and 3-D depth ranging. Since 2011, he has been a Full Professor of electronics with Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy. He is a coauthor of about 130 papers, which are published in peer-reviewed journals and conference proceedings, and 9 textbooks on electronic design and electronic systems. He is the coholder of four international patents. His research deals with microelectronic circuitry for SPADs and CMOS SPAD imagers for high-sensitivity time-resolved measurements, 2-D imaging, and 3-D depth ranging. Mr. Zappa has been a Senior Member of the IEEE since 2007.