Sahba S. Jahromi

SINGLE PHOTON AVALANCHE DETECTOR DEVICES AND CIRCUITS FOR MINIATURIZED 3D IMAGERS
SAHBA S. JAHROMI

SINGLE PHOTON AVALANCHE DETECTOR DEVICES AND CIRCUITS FOR MINIATURIZED 3D IMAGERS

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Information Technology and Electrical Engineering of the University of Oulu for public defence in Auditorium L9, Linnanmaa, on 5 June 2020, at 12 noon

UNIVERSITY OF OULU, OULU 2020
Jahromi, Sahba S., Single photon avalanche detector devices and circuits for miniaturized 3D imagers.

University of Oulu Graduate School; University of Oulu, Faculty of Information Technology and Electrical Engineering; Infotech Oulu

Acta Univ. Oul. C 752, 2020
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Abstract

In this thesis, a solid-state 3D imager receiver architecture based on the direct time-of-flight (dTOF) technique is discussed, implemented and tested. The distinctive feature of this work is the combination of a unique laser diode (LD) transmitter operating in enhanced gain-switching mode and a highly integrated single-chip SPAD/TDC array receiver. The LD transmitter is capable of producing short/energetic (~100-200 ps FWHM/~1-5 nJ) laser pulses at pulsing frequencies of some hundreds of kHz with a rather simple driver structure and small size. Another distinction of this work is the receiver IC architecture strategy of separating the SPAD array from all the on-chip electronics in a dedicated chip area, with the aim of achieving a high fill factor.

To evaluate the above-mentioned architecture, a receiver IC based on SPAD/TDC arrays was developed in a 0.35 μm HVCMOS technology. This IC, which included 9x9 SPADs and 9+1 TDCs, was paired with an LD transmitter and tested in a 3D imager prototype. The verification measurements of circuit/system-level performance confirmed the possibility of a compact low-cost solid-state 3D imager with spatial resolution of a few kilo pixels, cm-level depth precision and a frame rate of tens of fps based on the given architecture.

Working from the results obtained with the first prototype, a second IC was designed and manufactured in the same HVCMOS technology, which was a combination of a 128x32 SPAD array and a 256+1 TDC array. Two ICs were used to form a receiver with 8-kilo pixel spatial resolution and combined with a LD transmitter into a compact USB-powered 3D imager (total size 5x7x4 cm³). The LD transmitter provided an average optical power of 1 mW at a wavelength of 810 nm, producing 150 ps pulses at a pulsing frequency of 250 kHz and targeting a field-of-view of ~42°x21° with flood-pulsed illumination by means of simple optics. The 3D imager demonstrated frame rates of 5-10 fps with cm-level precision in the case of Lambertian targets within a range of 5 m.

Keywords: 3D imaging, CMOS SPAD, direct time-of-flight, focal plane imaging, laser diode, TDC
Jahromi, Sahba S., Yksittäisten fotonien havaitsemiseen perustuvan detektori- ja piiritekniikan kehittäminen 3D-kuvantamiseen.
Oulun yliopiston tutkijakoulu; Oulun yliopisto, Tieto- ja sähkötekniikan tiedekunta; Infotech Oulu
Acta Univ. Oul. C 752, 2020
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Tiivistelmä
Työssä on suunniteltu, toteutettu ja testattu vastaanotinarkkiettiuuri 3D-etäisyyskuvantamiseen. Mittaus perustuu tietyssä toimintamuodossa olevan, tarkoitusta varten suunnitellun laserdiodin tuottamien pulssien kulkuajojen mittaan toimii korkean integraatioasteen omaavalla, yksittäisiä fotoneja ilmaisemaan kykenevällä SPAD/TDC IC-piirillä. Laserdiodilähetin tuottaa lyhyitä ja energisiä (~100-200 ps FWHM/~1-5 nJ) laserpulsseja satojen kHz:n pulssitaajuudella pieneneen tilaan toteutetun elektroniikan ohjaamana. Eräs työn tärkeä piirre on valittu vastaanotinarkkitehtuuri, jossa fotoni-ilmaisimatriisi (SPAD-matriisi) on erotettu muusta vastaanotinelektroniikasta korkean ilmaisuhyötysuhteen aikaansamiseksi.

Arkitehtuurievaluointia varten työn alussa kehitettiin 0,35 µm:n HVC莫斯-teknologiassa toteutettu vastaanotin IC-piiri, joka sisältää 9x9 fotoni-ilmaisinta ja 9+1 aika-digitaali-muunninta samalla sirulla. 3D-testimittaukset varmensivat arkitehtuurin toimivuuden ja osoittivat, että siihen perustuen on mahdollista kehittää edullinen, ilman liikkuvia osia toimiva 3D-kuvannintechnologia, jolla voidaan saavuttaa cm-luokan etäisyysmittaustarkkuus ja muutaman kilopikselin spatiaalinen resoluutio sekunnin murto-osien mittausnopeudella.

Näihin tuloksiin perustuen työn seuraavassa vaiheessa suunniteltiin laajempi vastaanotin IC, jossa on 32x128 pikseliä sisältävä fotoni-ilmaisimatriisi ja 256+1 aika-digitaali-muunninta samalla sirulla. Piiri toteutettiin 0,35 µm:n HVC莫斯-teknologiassa. Kahta tällaista vastaanotinta käytetään toteutettiin kompakti (koko 5x7x4 cm³ laserlähettimen kanssa), USB-väylästä virtaavalla 8 kilopikselin 3D-kuvannit tuottavalla kytkennänä. Laserdiodilähetin keskimääräinen optinen teho on n. 1 mW ja se toimii 810 nm:n aallonpituudella. Lähetin tuottaa 150 ps:n pulssia 250 kHz:n pulssitaajuudella kuvantimen ~42°x21° mitausavaruuteen. Toteutettu 3D-kuvannin tuottaa etäisyyskuviot n. 5 m:n etäisyysdistansissa kohteista cm-luokan tarkkuudella 5-10 kuvaa sekunnissa mittausnopeudella.

Asiasanat: 3D-kuvantaminen, CMOS SPAD, pulssin kulkuaikeittaus, TDC, vastaanotinmatriisi
Acknowledgements

I would like to thank Professor Juha Kostamovaara for supervising this thesis. He has been a source of motivation and inspiration in research and life, and I am truly grateful for all I have learnt from him during these years. I would like to thank Docent Jussi-Pekka Jansson, with whom I collaborated closely in my thesis research. Working alongside and learning from him has been a pleasure all along. Thanks also to Associate Professor Ilkka Nissinen and Docent Jan Nissinen for their help and patient discussions, especially during the first years of my research. Similarly, thanks to Dr. Lauri Hallman, Dr. Pekka Keränen and Jaakko Huikari for their cooperation on different projects and to Matti Polojärvi for providing technical support. I would like to thank my roommates Dr. Henna Ruokamo and Aram Baharmast for sharing the experience of PhD research and for the interesting discussions about SPADs and TOF. I also wish to thank all my other former and current colleagues at CAS research unit (former Electronics Lab.) for creating an excellent research atmosphere.

I would like to thank the pre-examiners of this thesis, Professor Horst Zimmermann from Vienna University of Technology and Associate Professor Lucio Pancheri from University of Trento, for the time and effort they put into reviewing this thesis. I am also grateful to Malcolm Hicks for revising the English language of the manuscript and most of the publications.

I would like to gratefully acknowledge the financial support of Infotech Oulu Graduate School, Academy of Finland, Finnish Funding Agency for Technology and Innovation (TEKES), Tauno Tönning Foundation, Finnish Fund for Technology Promotion (TES), and KAUTE Foundation.

Finally, I would like to thank my friends and family for their support and encouragement. Above all, I would like to express my deep gratitude to my husband, Arash, for his love, patience and willingness to always listen to, discuss and help with research questions.

April 2020  

Sahba S. Jahromi
# Abbreviations and symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>one-dimensional</td>
</tr>
<tr>
<td>2D</td>
<td>two-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>APD</td>
<td>avalanche photodiode</td>
</tr>
<tr>
<td>CCD</td>
<td>charge-coupled device</td>
</tr>
<tr>
<td>CMOS</td>
<td>complementary metal oxide semiconductor</td>
</tr>
<tr>
<td>cps</td>
<td>counts per second</td>
</tr>
<tr>
<td>CW</td>
<td>continuous wave</td>
</tr>
<tr>
<td>DCR</td>
<td>dark count rate</td>
</tr>
<tr>
<td>DH</td>
<td>double heterostructure</td>
</tr>
<tr>
<td>DLL</td>
<td>delay locked loop</td>
</tr>
<tr>
<td>dTOF</td>
<td>direct time-of-flight</td>
</tr>
<tr>
<td>FF</td>
<td>fill factor</td>
</tr>
<tr>
<td>FOV</td>
<td>field of view</td>
</tr>
<tr>
<td>FPGA</td>
<td>field-programmable gate array</td>
</tr>
<tr>
<td>fps</td>
<td>frames per second</td>
</tr>
<tr>
<td>FWHM</td>
<td>full-width-at-half-maximum</td>
</tr>
<tr>
<td>HV</td>
<td>high voltage</td>
</tr>
<tr>
<td>GaAs</td>
<td>gallium arsenide</td>
</tr>
<tr>
<td>GUI</td>
<td>graphical user interface</td>
</tr>
<tr>
<td>IC</td>
<td>integrated circuit</td>
</tr>
<tr>
<td>InGaAs</td>
<td>indium gallium arsenide</td>
</tr>
<tr>
<td>INL</td>
<td>integral nonlinearity</td>
</tr>
<tr>
<td>INL-LUT</td>
<td>integral nonlinearity lookup table</td>
</tr>
<tr>
<td>iTOF</td>
<td>indirect time-of-flight</td>
</tr>
<tr>
<td>LD</td>
<td>laser diode</td>
</tr>
<tr>
<td>LiDAR</td>
<td>Light detector and ranging</td>
</tr>
<tr>
<td>MEMS</td>
<td>microelectromechanical systems</td>
</tr>
<tr>
<td>MOS</td>
<td>metal–oxide semiconductor</td>
</tr>
<tr>
<td>NBOF</td>
<td>narrow band optical filter</td>
</tr>
<tr>
<td>PCB</td>
<td>printed circuit board</td>
</tr>
<tr>
<td>PDE</td>
<td>photon detection efficiency</td>
</tr>
<tr>
<td>PDP</td>
<td>photon detection probability</td>
</tr>
<tr>
<td>pn</td>
<td>p-type/n-type</td>
</tr>
<tr>
<td>PVT</td>
<td>process, voltage and temperature</td>
</tr>
<tr>
<td>RLC</td>
<td>resistor-inductor-capacitor</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>RMS</td>
<td>root mean square</td>
</tr>
<tr>
<td>SNR</td>
<td>signal-to-noise ratio, defined here as the peak ratio of signal amplitude to RMS noise</td>
</tr>
<tr>
<td>SPAD</td>
<td>single photon avalanche diode</td>
</tr>
<tr>
<td>TAC</td>
<td>time-to-amplitude converter</td>
</tr>
<tr>
<td>TDC</td>
<td>time-to-digital converter</td>
</tr>
<tr>
<td>TOF</td>
<td>time-of-flight</td>
</tr>
<tr>
<td>QW</td>
<td>quantum well</td>
</tr>
<tr>
<td>VCSEL</td>
<td>vertical-cavity surface-emitting laser</td>
</tr>
<tr>
<td>$A_{\text{rec}}$</td>
<td>area of the receiver lens</td>
</tr>
<tr>
<td>$B_{\text{op}}$</td>
<td>optical bandwidth of the receiver</td>
</tr>
<tr>
<td>$c$</td>
<td>speed of light</td>
</tr>
<tr>
<td>$E_{\text{ph}}$</td>
<td>photon energy</td>
</tr>
<tr>
<td>$E_{\text{tr}}$</td>
<td>transmitted pulse energy</td>
</tr>
<tr>
<td>$F_{\text{FFSPAD}}$</td>
<td>fill factor of a single SPAD</td>
</tr>
<tr>
<td>$f_{\text{optics}}$</td>
<td>focal length of the receiver lens system</td>
</tr>
<tr>
<td>$\text{FOV}_{\text{SPAD}}$</td>
<td>field of view of a single SPAD</td>
</tr>
<tr>
<td>$h$</td>
<td>Planck constant</td>
</tr>
<tr>
<td>$I_{\text{s}}$</td>
<td>spectral irradiance</td>
</tr>
<tr>
<td>$N_{\text{BG}}$</td>
<td>average number of background detections</td>
</tr>
<tr>
<td>$N_{\text{meas}}$</td>
<td>number of measurements considered</td>
</tr>
<tr>
<td>$N_{\text{signal}}$</td>
<td>average number of signal photons seen by a SPAD</td>
</tr>
<tr>
<td>$N_{\text{SPAD}}$</td>
<td>number of SPADs in a detector array</td>
</tr>
<tr>
<td>$P_{\text{B}}$</td>
<td>background power seen by a SPAD</td>
</tr>
<tr>
<td>$R$</td>
<td>target distance</td>
</tr>
<tr>
<td>$R_{\text{max}}$</td>
<td>maximum range of 3D imager</td>
</tr>
<tr>
<td>$V_{\text{br}}$</td>
<td>breakdown voltage of SPAD</td>
</tr>
<tr>
<td>$V_{\text{ex}}$</td>
<td>excess bias voltage of SPAD</td>
</tr>
<tr>
<td>$\Delta t_n$</td>
<td>time interval between start and stop pulses for the $n^{th}$ detector</td>
</tr>
<tr>
<td>$\Delta t_{\text{pulse}}$</td>
<td>optical pulse time length</td>
</tr>
<tr>
<td>$\varepsilon_{\text{opt}}$</td>
<td>efficiency of the optics</td>
</tr>
<tr>
<td>$\rho_{\text{target}}$</td>
<td>target reflectivity</td>
</tr>
<tr>
<td>$\lambda_{\text{tr}}$</td>
<td>central wavelength of the laser</td>
</tr>
<tr>
<td>$\sigma_N$</td>
<td>measurement precision after averaging $N$ signal counts</td>
</tr>
<tr>
<td>$\sigma_{\text{LD}}$</td>
<td>standard deviation of the laser pulse width</td>
</tr>
<tr>
<td>$\sigma_{\text{SPAD}}$</td>
<td>timing jitter of SPAD</td>
</tr>
<tr>
<td>$\sigma_{\text{TDC}}$</td>
<td>time interval measurement uncertainty of the TDC</td>
</tr>
<tr>
<td>$\tau_{\text{BG}}$</td>
<td>mean time between background detections</td>
</tr>
</tbody>
</table>
List of original publications

This thesis is based on the following publications, which are referred to throughout the text by their Roman numerals:


Papers I–V were written by the author of this thesis with the help of co-authors. The author of this thesis designed the detector arrays and respective electronics for the two receiver ICs presented in papers I–V, co-designed the measurement systems and conducted all the measurements. The TDC arrays of the receiver ICs were designed and the parts of paper I–V regarding these written by the co-author Dr. Jussi-Pekka Jansson. The laser diode gain-switching principle, which was the basis for the LD transmitters used in papers I–V, was invented by Dr. Boris Ryvkin, Dr. Eugene Avrutin and Prof. Juha Kostamovaara. Some previously unpublished measurement results are also presented in this thesis.
1 Introduction

1.1 Motivation

Optical three-dimensional (3D) range imaging has long had applications in surveying, mapping and geodesy (Lemmens 2011; Liang et al. 2016), but there are now many new, upcoming applications such as driverless cars, consumer electronics (hand-held devices and games), gesture control and face recognition, farming and forestry, smart homes, flying drones and the control of heavy machinery (Schwarz 2010; Malik 2011; Coffey 2014; Link & Baraba 2014; Wang, Liu & Chan 2015). These applications would benefit from a solid-state solution that can provide scalability in terms of frame rate, spatial resolution and depth accuracy according to the needs of the specific application while implemented in small sizes and at low cost.

Unlike traditional 3D scanners, such a solid-state 3D imager needs to be electronic (without any mechanical moving parts), so that it can utilize the scaling and customization properties of modern IC technologies. Customizability is of especially great importance since the system-level performance parameters required for the above applications vary considerably, especially with regard to the maximum measurement range. The required range might be one metre (e.g. face recognition), a few tens of metres (e.g. autonomous machines) or up to ~200 m (e.g. traffic applications), while a measurement speed of a few tens of fps, and in some cases measurement in bright sunlight, might typically be required, which is challenging due to high levels of background noise.

Many optical 3D range imagers have been developed in recent years based on a variety of approaches such as stereoscopic vision, structured light and time of flight (TOF), each with inherent advantages and drawbacks (a short review of these approaches will be presented in the following chapter). The aim of this thesis is to propose and verify a specific architecture that can potentially pave the way for the development of compact solid-state 3D imaging devices capable of being tailored to the requirements of different applications.

1.2 Content of the thesis

The method presented here is based on direct time of flight (dTOF) with the basis of detecting and measuring the travel time of single photons contained in short,
intensive laser impulses. As a result, a digital approach can be achieved, as opposed to the analogue detector approach (Shcherbakova et al. 2013). The presented dTOF-based 3D imager incorporates a high-energy/short (1-5 nJ/100-200 ps) pulsed laser transmitter and a receiver, which includes arrays of single photon detectors and photon arrival-time measurement circuits in one silicon chip.

As part of the present research, two complementary metal oxide semiconductor (CMOS) single photon receiver ICs were developed in a 0.35 µm high voltage (HV) CMOS technology, both single-chip integrations of a dense single photon avalanche diode (SPAD) array and a time-to-digital-converter (TDC) array. The first design included 9×9 SPAD/9+1 TDC arrays, which, paired with the high energy/short pulsed transmitter, served as a successful evaluation of the 3D imaging concept mentioned above. The second design was a high-resolution successor of the former, with 128×32 SPAD/256+1 TDC arrays. Based on the developed IC, a compact USB-powered 3D imager was implemented with 8-kilo spatial resolution, a single shot range precision better than +/-2 cm, and a field of view (FOV) of ~42°×21°. A typical frame rate of 5–10 fps was achieved while measuring non-cooperative targets within a 10 m range. Both of the 3D imaging systems were evaluated by means of measurements using specially designed FPGA interfaces and MATLAB/Qt-based graphical user interfaces (GUIs), allowing real-time graphical presentation of the measured data and 3D videos.

The important features of this work are the use of a custom laser diode (LD) transmitter with unique properties, a specific receiver implementation strategy of separate on-chip SPAD/TDC arrays (as opposed to the mainstream smart pixel approach), miniaturized prototypes with application-oriented specifications and detailed circuit and system-level tests. The designing of the SPAD arrays and front-end electronics in both of the above receiver chips and the assessment of the circuit and system-level performance of the solid-state 3D imagers by means of measurements constitute the focus points and main contributions of this thesis. This thesis is based on the following five publications:

Papers I and II: The circuit-level design of the first IC (9x9 SPAD/9+1 TDC arrays) is presented and the strategy of separate on-chip SPAD/TDC arrays is introduced. Measurement results verifying the IC in a dTOF environment are presented and discussed.

Paper III: The 3D imaging concept based on dTOF focal-plane imaging using a short/high energy pulsed LD is suggested and validated through a 3D imager prototype based on the first IC and a LD transmitter (~100 ps/1 nJ/100 kHz/870 nm).
Paper IV: Crosstalk, as a probable effective disturbance in a dense SPAD array, is discussed and measured in detail (using the first IC). The temporal correlation of crosstalk events, crosstalk probability, its dependence on illumination wavelength and intensity and the distance between SPADs are the focus points of the measurements.

Paper V: The circuit-level design and measurement results of the second developed IC (128×32 SPAD/256+1 TDC arrays) are presented. The IC was paired with a LD transmitter (~150 ps/~3.8 nJ/250 kHz/810 nm) to form a solid-state 3D range imager with flood-pulsed illumination. The 8-kilo pixel 3D range imager was implemented in a 5×4×7 cm³ volume covering a ~42°×21° FOV with frame rates of 5–10 fps and cm-level precision within a range of 5 m. The idea of matching a solid-state scan of illumination, e.g. using an addressable LD bar or a vertical-cavity surface-emitting laser (VCSEL) array, with sections of the detector array is presented, and discussed by comparison with the mainstream flood illumination approach.

This work has also contributed to published research other than the above-mentioned papers. The receiver ICs have also been used in research into one-dimensional (1D) rangefinding (Paper II; Huikari et al. 2017; Huikari et al. 2018), two-dimensional (2D) line profiling (Hallman et al. 2018), time-resolved diffused-optics imaging (Nissinen et al. 2018), time interval measurement with high-performance TDCs (Jansson et al. 2019) and characterization of SPAD properties (Nissinen et al. 2017). However, since the focus of this thesis is on dTOF 3D imaging, the above-mentioned five publications were chosen for inclusion in this thesis.

In the following chapters, an introduction to the approaches and techniques involved in optical 3D imaging is given first, in Chapter 2, then the dTOF 3D range imaging technique, which formed the basis of the present work, is discussed in more detail in Chapter 3, where the system-level structure and circuit-level design concepts and architecture, which are common to both ICs and 3D imager prototypes, are also explained in detail. Chapter 4 presents the design of the 9×9 SPAD array/9+1 TDC receiver IC and a solid-state laser 3D imager prototype based on it, together with measurements, while Chapter 5 presents the design and measurements of the 128×32 SPAD array/256+1 TDC receiver IC and the solid-state 3D imager based on it. A discussion of the measured results and a short comparison with the state of the art are presented in Chapter 6, where possible future improvements and particularly the idea of a laser illumination concentrated in both time (short sub-ns pulses) and space (targeting only the active rows of the
SPAD array) are also discussed. Finally, a summary of the thesis is given in Chapter 7.
2 3D range imaging approaches

Where a conventional camera measures the intensity of the radiant excitation of points on a 2D image surface within its FOV, a 3D imager can measure the spatial coordinates of points of a surface. In other words, a 3D imager can determine the distance of each point within the image surface with respect to a reference (Fig. 1).

Fig. 1. 3D vs. 2D imaging (reprinted by permission from Paper III © 2016 The Optical Society).

3D imagers can be implemented with different optical approaches, each utilizing unique technologies and techniques, such as stereoscopic vision (Lazaros, Sirakoulis & Gasteratos 2008; Link & Baraba 2014), structured light (Li et al. 2017) and time of flight (direct or indirect) (Haroud et al. 2016).

Stereoscopic 3D imaging is a solid-state approach that creates depth images in a way similar to the human eyes. The same image is captured by two (or more) cameras (on the same image plane) and the slightly different image fields are superimposed to measure distance using triangulation (Fig. 2(a)). These 3D imagers use well-established low-cost hardware (2D cameras) and do not require active illumination. On the other hand, a slight camera misalignment can cause significant errors in the calculation of the distance, and if the target is moving, precise synchronization between cameras is required. If the surface texture is uniform or repetitive, the measurement accuracy is limited, since detecting the correspondence pairs between images becomes difficult (Li et al. 2017). Above all, since stereoscopic vision relies on ambient light, only measurements under bright illumination conditions are possible.

Structured light is another solid-state optical 3D imaging approach, in which the target is actively illuminated (via laser, LED or lamp) by projecting a fixed pattern or series of programmable patterns using diffractive optics, and the resulting
image is recorded with an image sensor. Given the fixed distance between the illuminator and the sensor, it is possible to locate and measure the depth of specific points with triangulation algorithms (Fig. 2(b)) (Link & Baraba 2014).

One of the best-known examples of structured light was Microsoft Kinect I (PrimeSense) (Zhang 2012), which was mostly developed for inferring human body positions in electronic games and used a pre-defined unique dot distribution pattern. Although such solid-state implementations based on structured light have gained substantial success, one important issue is the large size (width of several tens of centimetres), which arises from the measurement principle that is used, which is sensitive to the distance between the cameras. The speed requirement for the sensor (usually a conventional CMOS or CCD camera), power dissipation, the complexity of the software algorithms and the low signal-to-noise ratio (SNR) (in daylight) are some of the other challenges (Niclass, Besse & Charbon 2005; Link & Baraba 2014).

Active optical illumination of the FOV (e.g. laser, LED …) and measurement of the travel time of light to the target is another popular approach to 3D imager construction. Optical TOF 3D imagers can be categorized in terms of their TOF measurement technique, i.e. indirect (iTOF) or direct (dTOF).
A modulated continuous wave (CW) laser is mostly used in the iTOF approach, the distance to the target being deduced from the phase difference between a modulated laser source and the reflected wave (Fig. 3). The transmitted light signal is usually precisely sinusoidally modulated in a frequency range of some tens of megahertz. The received signal is phase-shifted and attenuated during the round trip to and from the target with an amplitude offset due to the background illumination. In another variation of iTOF, square-shaped light pulses (tens to hundreds of ns) are used to illuminate the targeted area (instead of the usual CW) and the back-scattered signal is integrated within defined time windows to obtain the shift in time domain (Stoppa et al. 2007; Bellisai et al. 2013; Yamada et al. 2018; Davidovic et al. 2010; Zach, Davidovic & Zimmermann 2010). Each pixel of the receiver independently demodulates the received signal, measuring both its phase delay, amplitude and offset (Lange 2006; Walker, Richardson & Henderson 2011; Shcherbakova et al. 2013; Mitev & Pollini 2015; Bronzi et al. 2016). The pixel can perform phase comparison either by photo-generated charge integration using a CMOS Active Pixel Sensor (APS) (Mitev & Pollini 2015) or a CMOS/CCD pixel (Oggier et al. 2004), or by photon counting using a CMOS SPAD array (Bronzi et al. 2016).

Fig. 3. Indirect TOF principle.

This 3D imaging approach can be solid-state and has achieved commercial success, e.g. Microsoft Kinect II (Canesta Technology) (Bamji et al. 2015; Payne et al. 2014). Indirect TOF 3D imagers offer a high spatial resolution, but the distance range is limited to only a few metres. If a modulation frequency of 30 MHz is considered, the unambiguous range of the camera is 5 m. A decrease in modulation frequency
can reduce the phase-wrapping ambiguity, but it will also detract from the accuracy. As a result, complex schemes, e.g. using multiple modulation frequencies or repetition rates, are needed to increase the range capacity (Droeschel, Holz & Behnke 2010; Liang et al. 2014). In addition, the transmitter will typically use a CW illumination power of several hundreds of mWs (e.g. with several parallel sources), even in short range applications, if high spatial resolution is required (Mitev & Pollini 2015; Bamji et al. 2015; Bronzi et al. 2016), making miniaturization difficult and raising eye-safety issues (Bamji et al. 2015; Hsu et al. 2018). In high background situations the average optical power needs to be even higher.

Direct TOF is based on the emission of short laser pulses towards the target and directly measuring the transit time to the target and back to the detector. This technique has traditionally been used in laser radars based on linear optical detectors (e.g. APD), and has now become relatively popular for use with single photon detectors. Compared with iTOF, direct TOF has a potential for lower average illumination power, especially under high background noise conditions, since it concentrates the available laser power in short, intensive pulses (Koskinen, Kostamovaara & Myllyläe 1992; Kostamovaara et al. 2015).

The traditional dTOF solution to the 3D perception problem are 3D scanners, in which the beam (a single plane or multiple parallel planes) is mechanically scanned over the target volume using rotating components. These systems offer high performance, but are bulky, expensive and power hungry, e.g. (Niclass et al. 2013; Horaud et al. 2016; Zhang et al. 2018). Single photon detectors in dTOF 3D scanners have been implemented with different system-level approaches. In a few such systems a MEMS scanning mirror (Niclass et al. 2012; Ito et al. 2013) or a spinning mirror (Niclass et al. 2014) has been used for realization of the targeted FOV, leading to very accurate systems, although they do not fulfil the solid-state requirements or power, cost and size targets of many applications because of the mechanically rotating components.

A step forward in developing new generations of dTOF 3D imagers was the replacing of the mechanical scanner with a fully electronic solid-state imaging approach. Such solid-state 3D imagers have attracted a lot of attention in recent years (Bellisai et al. 2013; Bronzi et al. 2016; Buller & Wallace 2007; Henderson et al. 2019; Lussana et al. 2015; Oggier et al. 2004; Payne et al. 2014; Perenzoni, M., Perenzoni, D. & Stoppa 2017; Vornicu, Carmona-Galán & Rodríguez-Vázquez 2017; Zhang et al. 2018). The focus in the above studies was on the development of powerful single photon-based dTOF receivers for use with pulsed flood
illumination, where the receiver architecture of implementing CMOS arrays of smart pixels (combining optical detection functionality with signal processing power in each pixel) has received considerable attention. However, although excellent circuit-level results have been achieved, not much attention has been paid to the realization of the system (i.e. integrating the receiver with an illuminator), and hence there have not been many comprehensive realizations with convincing 3D range image results.

This thesis attempts to propose one possible technique for implementing a solid-state dTOF 3D imager. The solid-state laser 3D imaging concept studied in here is based on the dTOF focal plane imaging principle, as presented in Fig. 4. An optical pulse (burst of photons), usually from a laser diode, illuminates a target cone within the FOV of the system, the desired FOV being achieved by means of suitable optics (e.g. cross-coupled cylindrical lenses or an engineered diffuser).

A portion of the transmitted photons will be reflected or backscattered from the target to hit the detector array. The receiver consists of a 2D array of optical detectors (e.g. linear or single photon detectors) located approximately at the focal plane of a positive lens (or lens system). The FOV of the detector array can be matched with that of the transmitter by considering the dimensions of the 2D detector array (typically a few mm²) and adjusting the focal length of the receiver optics accordingly, e.g. small-sized paraxial optics with matched divergence in the transmitter and receiver.

For each detector within the array, the time interval between the laser pulse emission and detection is measured. These time intervals are the travel times of the laser pulse from the transmitter to the target and back to the detectors ($\Delta t_n$: for the $n^{th}$ detector). Since each detector sees only a small cone within the imager’s FOV (defined by its relative position within the array), a 3D range image can be produced. The $x$ and $y$-coordinates of the 3D image points (perpendicular to the optical axis) are defined by the location of the individual detector in the array and the $z$-coordinate (distance along the optical axis) is calculated from the measured travel time of the optical pulse, which is indicated by the corresponding detector. The distance is simply $z_n = (\Delta t_n \cdot c)/2$, where $c$ is the speed of light.
The distinct features of this present work will be discussed in more detail in the following chapter. It is recognized that the performance of the system will be improved if the available average optical energy is concentrated in short, intensive laser pulses. Thus, sub-ns laser pulses with high energy at the level of a few nJ is used, which is substantially higher than in previous studies. Another feature is that, instead of using the conventional smart-pixel architecture in the design of the receiver circuits, an approach is adopted that tries to make better use of the characteristic features of SPAD-based dTOF techniques and optimize the use of the available average illumination power.
3 Direct TOF solid-state focal plane 3D imaging

The intention here was to implement the detector array in the above dTOF architecture by means of CMOS Single Photon Avalanche Diodes (SPADs). With SPADs as detectors, the detection of the returning pulse can be implemented as a low-jitter (~100 ps) digital process, and the time interval measurement circuit can be integrated on the same detector chip. The CMOS SPAD will be briefly introduced and its properties discussed in Section 3.1 in order to provide a starting point for system-level design of the dTOF 3D imager.

The fact that the whole system of a SPAD-based dTOF 3D imager is digital in nature (impulse-like light pulses sent, single photons detected) enables a relatively simple and robust realization. Such an implementation has the potential to serve as a generic 3D range imaging principle, because it can be scaled with respect to the performance parameters needed in particular applications (e.g. range, spatial resolution, accuracy, FOV, etc.). The general characteristics of SPAD-based dTOF 3D range imaging are discussed in Section 3.2., but it should be remembered that the performance of the 3D imager depends on the architecture and specifications of its building blocks, most importantly the transmitter and the receiver. The design strategy and architecture of the transmitter and receiver, which are specific to the 3D imager designs put forward in this thesis, are presented in Sections 3.3.1 and 3.3.2.

3.1 CMOS single photon avalanche diode (SPAD)

A SPAD is a p-n junction (Fig. 5(a)) which is reverse-biased above its breakdown voltage ($V_{br}$) and left floating, so that any stimulus (triggering) leading to the creation of an electron-hole pair can trigger a self-sustaining avalanche (breakdown). If the primary carrier is photon-generated, the leading edge of the avalanche pulse marks the arrival time of the detected photon (Cova et al. 1996). Therefore, the avalanche breakdown results in a digital logic-level signal transition (detection). As a result, SPADs do not require analogue signal processing, contrary to that needed with a linear APD (Avalanche Photo-Diode) detector (Bronzi et al. 2012; Burri et al. 2014; Peranzoni, Pancheri & Stoppa 2016; Richardson et al. 2011; Rochas et al. 2002; Villa et al. 2014). SPADs are usually reverse-biased above breakdown by an additional voltage called an excess bias voltage ($V_{ex}$). After the breakdown, the avalanche current flowing through the SPAD is stopped by active or passive quenching and the SPAD needs to be re-biased above breakdown again.
by means of a dedicated circuit to be prepared for another detection (Fig. 5(b)). CMOS SPADs can be implemented with a variety of structures in different CMOS technologies, the most common structure being a shallow junction between a highly-doped diffusion layer and a well with opposite doping, (e.g. p+-nwell as shown in Fig. 5(a)). The lightly-doped region surrounding the junction (p in Fig 5(a)) is to avoid premature breakdown by reducing the electric field near the edges to ensure a high avalanche initiation probability towards the centre of the multiplication region and create a planar multiplication region.

**Fig. 5. a) Cross-section, and b) I-V curve of a CMOS SPAD.**

The probability of detecting a photon hitting the surface of a SPAD (i.e. active area) is known as the photon detection probability (PDP). With the CMOS SPAD shown in Fig. 5 (with a shallow junction), the wavelength-dependent PDP has its peak (~30%) in the visible wavelength spectrum (400–500 nm) and decreases in the near infrared region to ~2–5%, corresponding to 870–810 nm (Nissinen et al. 2011; Pancheri, Stoppa & Dalla Betta 2014; Richardson et al. 2011; Stoppa et al. 2009). The PDP is especially important in the near-infrared range (800–900 nm), as most powerful lasers operate in this range, the illumination is invisible to human eye and an Si detector can still be used. Reflection of a portion of the incoming photons from the surface of the SPAD, the wavelength dependent absorption probability, the statistical nature of the avalanche build-up (i.e. the avalanche initiation probability) together contribute to the above-mentioned detection probability (Niclass, Besse & Charbon 2005; Rochas et al. 2003). The strength of the electric field in the multiplication region (dependent on $V_{ex}$) can also affect the PDP.

Another important property of a CMOS SPAD is the low timing jitter of its photon detection (i.e. the statistical fluctuation of the time interval between a
photon arriving at the surface of the SPAD and the leading edge of the avalanche current). Photons, which are absorbed directly in the depletion region of the junction, produce an electron-hole pair which may start an avalanche. Since the avalanche build-up is a random process, it causes randomness in the response time of the SPAD, which in turn leads to a Gaussian-like timing distribution of the SPAD’s detections. On the other hand, photons that are absorbed in different vertical locations within the well might introduce electron-hole pairs, from which the minority carriers might diffuse towards the depleted area. This would lead to a tail in the timing distribution of SPAD’s detections (Rochas et al. 2003; Cova et al. 2004). CMOS SPADs usually have a timing jitter of 50–100 ps (full-width-at-half-maximum, FWHM).

The main source of uncorrelated noise in CMOS SPADs is the thermal or tunnelling-assisted generation of carriers, which might lead to unwanted avalanches (often regarded as dark noise). The dark count rate (DCR) is a measure of the mean number of dark counts (dark noise-induced breakdowns) per second and is considered to follow the Poisson distribution. DCR is dependent on the structure, size and temperature of the detector and varies from a few to hundreds of cps/µm² (Bronzi et al. 2012; Nissinen et al. 2011; Pancheri & Stoppa, 2007; Perenzoni, Pancheri & Stoppa 2016; Richardson et al. 2011). The DCR might be a problem in some SPAD applications, e.g. when measuring low intensity optical signals under dark conditions, but it is not usually very important in 3D imaging applications, since the presence of a much stronger uncorrelated noise (due to background illumination) dominates.

3.2 SPAD-based dTOF in 3D imaging applications

In most practical 3D imaging applications SPADs need to be measured many times to produce a 3D range image, i.e. more than one measurement is required for each SPAD to get a valid detection from the target. This is because the number of photons seen by the receiver array is typically low, since the received optical energy is proportional to the target reflectivity and $1/R^2$ ($R$ being the target distance), on top of the low PDP of CMOS SPADs in the near-infrared range. Therefore, when measuring Lambertian targets with distances above 1 m, usually only a few SPADs within the array detect photons returning from the target per measurement and since the detection process is spatially random throughout the array, multiple measurements of the array are needed to make sure all the SPADs have at least one valid detection of a photon returning from the target.
In addition, reliable measurement of the target distance usually requires more than one valid photon detection. This is because of the presence of noise sources, namely the quantum noise of the optical signal, the internal noise of the SPAD (i.e. DCR), the external background noise and jitter of the detection and time interval measurement operations. A SPAD DCR of 100–1M cps corresponds to a mean time of 0.1 s–1 µs between dark triggers and can usually be ignored in the presence of background radiation from the sun (especially in outdoor applications). At a 70 klux background illumination level, e.g. on a sunny day, the mean time between background photons hitting the SPADs can be as low as only one to few nanoseconds (depending on the measurement arrangements, especially the optics) (Kostamovaara et al., 2015). With a PDP of <5% at near-infrared wavelengths, this corresponds to one background-induced triggering per a few tens of nanoseconds. If a target within a range of a few tens of metres is considered (6.7 ns ~ 1 m) under such background conditions, the target may even be blocked due to random background photons occurring between the laser pulse emission and the arrival of photons backscattering from the target.

Neglecting the effect of the noise (i.e. assuming that the only limitation affecting distance measurement is the number of valid signal detections, Fig. 6(a)), three main factors contribute to the overall timing (distance measurement) uncertainty of a single valid detection (assuming a detection probability of <1 per pulse): the timing jitter of the detection (\(\sigma_{SPAD}\)), the time interval measurement uncertainty (\(\sigma_{TDC}\)) and the width of the optical signal (\(\sigma_{LD}\)). Since all these factors are independent when operating in single photon mode (single photon incidents causing avalanche events in detectors), the overall uncertainty can be estimated by \(\sqrt{\sigma_{SPAD}^2 + \sigma_{TDC}^2 + \sigma_{LD}^2}\). The precision can then be improved by building and analysing histograms of multiple detections (Fig. 6(a)). Consideration of successive detections of optical signals would improve the precision by a factor of \(\sqrt{N_{signal}}\) when considering a histogram composed of \(N_{signal}\) valid detections. With an intrinsic timing jitter of a CMOS SPAD of 50–100 ps FWHM, if the transmitter and time interval measurement circuit are designed to have an uncertainty of the same order, e.g. a laser pulse FWHM of ~150 ps (\(\sigma_{LD} = \sim 65\) ps) and a time interval measurement precision of ~100 ps FWHM, a single shot precision of ~200 ps (3 cm) FWHM can be achieved. The envelope of signal detections in the histogram usually has an exponential tail as the result of the diffusion tail of the SPAD detection distribution (and in some cases the shape of the optical pulse).
Fig. 6. a) Detection distribution without background noise, b) with the presence of relatively strong dark or background noise and c) with gating.

In the presence of high background noise, however, the number of laser shots needed is higher due to the reduced SNR resulting from the increased number of noise detections. As mentioned, the SPAD might even be blocked from detecting target photons (Fig. 6(b)). Many solutions have been proposed for background suppression problem in single photon detection devices, e.g. attenuation of the PDP of the detectors, photon coincidence detection, and enabling time interval measurement only in the presence of a certain time correlation of SPAD triggerings. Using predefined adjustable gate windows (gating) can also be an effective way of suppressing high background illumination (Fig. 6(c)). Gating is simply activation of the SPADs only within a time window around the expected returning laser pulse. Various algorithms can be developed for controlling the placement and width of the gate window (in each measurement) depending on environmental conditions such as target movement, background illumination level, etc., in order to reach the highest possible image update frequency. Against a high background, for example, where short gate windows (a few tens of nanoseconds) might be needed for background noise cancelation, quick scans can be used to perform the first measurements using long gate windows simply to locate the whereabouts of the target in the distance range. Then, short gate windows can be used to resolve the exact position of the target.

Gating can suppress background noise to some extent by preventing the SPADs from getting blocked by unwanted photons, but it cannot completely solve the background problem. Even in the case of ideal gate window placement, random background detections occurring within the duration of the backscattered laser pulse are still unaffected by gating. If building a histogram from detections of \( N_{\text{meas}} \) measurements (individual laser shots) is considered, the average number of signal photons seen by a single SPAD, \( N_{\text{signal}} \) (statistical variation follows the Poisson distribution), can be estimated by
\[ N_{signal} = \frac{E_{tr} \varepsilon_{opt} \rho_{target} A_{rec}}{\pi R^2 E_{ph}} \times \frac{PDP \times FF_{SPAD}}{N_{SPAD}} \times N_{meas} \]  

where \( E_{tr} \) is transmitted pulse energy (J), \( \varepsilon_{opt} \) is efficiency of the optics, \( \rho_{target} \) is target reflectivity, \( A_{rec} \) is area of the receiver lens (m\(^2\)), \( R \) is target distance (m), \( E_{ph} \) is photon energy \( (hc/\lambda_{tr} = 2.3 \times 10^{-19} \text{ J}) \), \( N_{SPAD} \) is the number of SPADs of the array, \( PDP \) is the photon detection probability of the SPAD, and \( FF_{SPAD} \) is fill factor of a single SPAD. The background power \( P_B \) (in units of W) seen by a SPAD is given by

\[ P_B = I_S A_{rec} \varepsilon_{opt} \rho_{target} \left( \frac{FOV_{SPAD} \times FF_{SPAD}}{2} \right)^2 BW_{opt}, \]

where \( FOV_{SPAD} \) is field of view of a single SPAD, \( BW_{opt} \) is the optical bandwidth of the receiver (nm) and \( I_S \) is the spectral irradiance (Wm\(^{-2}\)nm\(^{-1}\)). For example, an illumination level of 70 klux (bright sunlight) corresponds approximately to \( \sim 0.6 \) Wm\(^{-2}\)nm\(^{-1}\) at 810 nm.

Considering \( N_{meas} \) measurements, the average number of background detections \( (N_{BG}) \) within the duration of the returning laser pulse can be estimated using equation

\[ N_{BG} = N_{meas} \frac{\Delta t_{pulse}}{\tau_{BG}}, \]

where

\[ \tau_{BG} = \frac{1}{PDP} \left( \frac{P_B}{E_{ph}} \right)^{-1} \]

is the mean time between background detections, and \( \Delta t_{pulse} \) is the optical pulse duration. If we assume that the suppression effect of the background on the signal can be avoided by proper gate window adjustment, the SNR (due to background noise) then can be described as the average number of pulse detections divided by the standard deviation of the total number of noise detections during the received pulse envelope follows

\[ SNR = \frac{N_{signal}}{\sqrt{N_{BG}}} \propto \frac{E_{tr}/N_{meas}}{\sqrt{P_B \Delta t_{pulse}}}. \]

Eq. (4) suggests that, in principle, concentrating the available illumination energy in short, intensive laser pulses can improve the available SNR. On the other hand, assuming a fixed laser pulse energy, the SNR and maximum range \( (R_{max}) \) will increase by a factor of \( \sqrt{P_B} \) and \( \sqrt[4]{P_B} \), respectively, as the background illumination power decreases.
3.3 Design principles

3.3.1 Transmitter

The critical parameters of a pulsed laser transmitter intended for dTOF 3D imaging are the energy and width of its optical pulse, the pulsing rate, the complexity of its architecture and its size. Many applications require simultaneously a reasonable measurement FOV and range (e.g. 20°–50° in 5–50 m) to non-cooperative (Lambertian) targets with a frame rate of 10–25 fps and cm-level distance measurement precision (a 1 cm change in the target distance corresponds to ~67 ps in the measured travel time of the laser pulse).

High-power ns-scale pulsed illumination (e.g. 3–5 ns pulses/~40 W peak optical power/sub-MHz pulsing frequency) provided by a high-power laser diode has traditionally been used in dTOF (usually paired with APD and a highly sensitive analogue receiver channel to achieve cm-level precision). This type of transmitter can provide high average illumination power resulting in a larger range coverage (Niclass et al. 2012; Ito et al. 2013; Niclass et al. 2014). Since the detector in the SPAD-based dTOF 3D imager with such a transmitter produces ~100 ps jitter, the achievable single shot timing precision (uncertainty of measurement based on a single detected laser pulse) is directly limited by the pulse width, given that in a typical measurement situation (detection probability <<1) there is no prior knowledge concerning the position of the detected photon within the laser pulse envelope. In addition, at a higher detection level (detection probability ~1), e.g. while measuring shorter ranges, the detection probability is higher in the front part of the pulse, which also introduces a systematic timing walk error into the detection (Kostamovaara et al. 2015).

Another transmitter approach, which is especially appealing for single photon-based detection, is to use low power pulses at a high pulsing rate (e.g. 100–200 ps/~100 mW peak optical power/50 MHz pulsing frequency) (Niclass et al. 2005). This method provides a better single-shot precision, and a high pulsing rate is in any case advantageous for obtaining a higher frame rate, since single photon detectors are not very efficient devices. This would add to the complexity of both the transmitter and receiver electronics, however, and combined with the lower power optical pulse would increases the sensitivity to noise. On the other hand, an unambiguous measurement requires that a new laser pulse should not be sent before the echo from the previous pulse has been received. This sets a range-based limit on how high the pulsing rate can be (e.g. a maximum permitted pulsing rate of ~3
MHz for a 50 m range), unless a more complex scheme for resolving the ambiguity is employed (Krichel, McCarthy & Buller 2010).

Ideally, concentrating a high illumination energy (e.g. 1–10 nJ) in short pulses (100–200 ps) at a moderate pulsing rate (0.1–1 MHz) would provide a combination of the advantages of both transmitter architectures mentioned above, considering that the probability of detection per pulse in a single photon array-based dTOF device is <<1. In the 3D imagers presented here, short high-energy (~100–200 ps, ~10 W) laser pulses of this kind are produced by a customized LD transmitter operating on the enhanced gain switching principle (Lanz et al. 2013), which can be realized in any LD, e.g. bulk, quantum well (QW) or VCSEL. In the 3D imagers presented in the following chapters, two such LDs, bulk and QW, are used as pulsed illuminators, with corresponding central wavelengths of 870 nm and 810 nm, respectively. In such an LD transmitter, the drive current pulse width can be considerably longer than the optical pulse width (e.g. 1.5 ns compared with 100 ps), thus simplifying the design of the driver electronics. A combination of an LCR transient-based pulse shape control circuit and a high-speed ON-type switch was used to produce the necessary transient current, which can easily be realized with MOS/CMOS technologies (Nissinen & Kostamovaara 2013; Hallman, Huikari & Kostamovaara 2014; Huikari et al. 2015; Nissinen & Kostamovaara 2016). The LD driver schematics of such an LD transmitter is shown in Fig. 7. With a current pulse peak of ~10 A and width of 1–2 ns, we can expect an optical pulse of up to a few nJ in energy and as short as ~100 ps. If less energy is needed, the drive current can be scaled accordingly. The use of a semiconductor LD also leads to the possibility of a less complex system and miniaturization of the 3D imager as a whole. Since the current pulses driven through the laser diode are quite short (1–1.5 ns), the average current at a pulsing frequency of 1 MHz is only ~10 mA, for example. Thus, such an LD driver can achieve pulsing frequencies of ~0.1–1 MHz.

![Fig. 7. LD driver schematics.](image-url)
3.3.2 Receiver

Fig. 8(a) shows a general block diagram of the receiver architecture. The main parts of the receiver IC floorplan are the SPAD array, the array of interfacing electronics and the TDC array. Since both receiver ICs presented in the following chapters possess the same general architecture, the shared basic design concepts will be presented briefly here. Fig. 8(b) and (c) show how the three main arrays are implemented in separate IC areas in both designs. The ICs were implemented in a 0.35 µm HVCMOS technology, which provides acceptable performance for the SPAD and was chosen based on previous assessments of the SPADs in this technology undertaken by the research group.

Fig. 8. a) General IC architecture of the receiver, and micrographs of b) 9×9 SPAD/ 9+1 TDC array IC and c) 32×128 SPAD/256+1 TDC array IC (reprinted by permission from Paper I © 2015 IEEE and under CC BY from Paper V © 2020 Authors).
The structure of the SPAD (as part of a SPAD array), which was adopted in this thesis, is shown in Fig. 9. The SPADs have square-shaped active areas with rounded corners, a shared deep n-well cathode and a p+ anode. The shared deep n-well cathode leads to the possibility of achieving a high fill-factor (FF), and as a result, any interfacing electronics should be placed outside the SPAD array (no in-pixel electronics). Rounded corners together with a p-well guard ring around the active area (i.e. a lightly doped region compared to the p+ anode) prevent premature breakdown around the edges. In the structure presented here, a maximum $V_{cx}$ of 3.3 V can be reached (more detail in section 3.3.2.4).

**Fig. 9.** a) 3D model of a SPAD as part of an array, and b) cross-section of a SPAD in a 0.35 μm HVCMOS (reprinted [adapted] by permission from Paper IV © 2018 The Optical Society and under CC BY from Paper V © 2020 Authors).

SPAD arrays (with up to tens of thousands of detectors) can be realized in CMOS technologies with various active area sizes, structures and densities, but achieving high FF is difficult once such a large array is considered, mainly because of two limiting factors: the dark noise of the SPADs and the non-photosensitive area within the pixel pitch. Dark noise increases super-linearly with any increase in the active area, and thus limits the maximum possible size of the active area. In addition, in most SPAD array implementations with in-pixel electronics, the trade-off between the performance of the pixel and the FF leads to an active area-to-pitch ratio as low as a few percent. Instead of this mainstream strategy of smart pixels, the SPAD array in this work is implemented separately and routed to the interfacing electronics. Having the detectors and electronic arrays on the same chip but placing the time measurement and front-end electronics outside the SPAD array can simplify the design and effectively ensure a high FF. With this strategy a large array
of a few kilo pixels can be implemented in 0.35 µm HVC莫斯 technology with a FF of 30–40%, which is quite high compared with architectures employing in-pixel electronics, which might require the use of complementary FF recovery techniques, e.g. microlens arrays (Burri et al. 2014; Pavia, Wolf & Charbon 2014). If this strategy is used in a deep sub-micron CMOS technology (<0.25 µm), with an increased number of metal layers for routing and less limitations on the layout design rules, the active area of the SPAD can account for an even higher proportion of the pixel pitch. Another recent trend in manufacturing SPAD arrays has been 3D-stacked arrays to separate the detector array and the electronics, thereby achieving a high FF (Charbon, Bruschini & Lee 2018). Although very promising, these technologies are not mature yet.

**Correlated noise in dense SPAD arrays**

On top of the uncorrelated dark and background noise, correlated noise in SPAD arrays can also be limiting for 3D imaging applications. The two known sources of correlated noise in SPADs are afterpulsing and crosstalk. Afterpulsing is caused by some carriers becoming trapped at energy levels near the edges of the bandgap during each avalanche. These carriers are subsequently released and can cause another avalanche, generating afterpulses that are correlated with the previous avalanche of the same SPAD (Cova et al. 2004). The release time constant of these traps is of the order of few nanoseconds. As a result, afterpulsing is not an important issue in DTOF applications unless the pulsing rate is some hundreds of MHz. CMOS SPADs have an afterpulsing probability of the order of 0.5–2% (Pancheri & Stoppa 2007; Rochas et al. 2002).

Crosstalk noise is the occurrence of unwanted avalanches in one SPAD (detector of crosstalk noise), which are caused by an avalanche in another SPAD of the array (emitter of crosstalk noise), leading to a false detection. Crosstalk may limit the dynamic range of the SPAD detector array and lead to a blurred image in imaging applications (e.g. in the case of two adjacent SPADs with a considerable difference in incident light intensity, crosstalk detections from the neighbouring SPAD might dominate the target photon detections in the less-illuminated SPAD, leading to a false measurement).

Separating the SPAD array from the interfacing electronics (as implemented here) creates a dense detector array in which crosstalk might become a potential issue. This noise source was studied thoroughly in this research (Paper IV) (especially from the point of view of the temporal correlation of crosstalk events
and its practical effects in dTOF systems). The mechanisms that can lead to crosstalk in two neighbouring SPADs are shown in Fig. 10. For simplicity, the devices are shown as p-n junctions, and the third dimension is not presented. The SPADs share a common terminal, as is the case in high-density CMOS SPAD arrays (Paper IV; Niclass et al. 2008). During an avalanche in one SPAD, photons are emitted owing to the electroluminescence effect as a result of the relaxation of hot-carriers generated in response to the passage of a large current in a strong electric field. A portion of these photons find their way to neighbouring SPADs, travelling along direct and indirect optical paths, thus leading to optical crosstalk (Paper IV; Kindt, Zeijl & Middelhoek 1998). The crosstalk may also be electrical, i.e. some of the carriers generated during the avalanche may exit the depletion region of the avalanching SPAD, diffuse laterally and eventually reach the depletion region of a neighbouring SPAD (Seitz & Theuwissen 2011; Vila et al. 2014). A combination of optical and electrical crosstalk is also possible; a secondary photon generated as a result of an avalanche in one SPAD may travel into the vicinity of a neighbouring SPAD and be absorbed, generating a minority carrier that might diffuse to the depletion region of the neighbouring device. The photons or carriers that reach the neighbouring SPADs may trigger an avalanche and yield incorrect measurement results.

![Fig. 10. 2D representation of direct (1) and indirect (2) optical crosstalk, and electrical crosstalk (3) mechanisms and their combination (4) (reprinted by permission from Paper IV © 2018 The Optical Society).](image)

The probability of crosstalk noise causing a false detection in a neighbouring device of an avalanching SPAD, known as the crosstalk probability, is directly dependent on the energy dissipated during an avalanche event in the emitter, since both the
electroluminescence intensity and the number of carriers diffusing out of the depletion region are related to the number of carriers in an avalanche pulse. Detailed measured results regarding crosstalk in the above-mentioned dense CMOS SPAD array are presented in Section 4.1.

**TDC array**

The use of CMOS SPADs opens up the possibility of integrating TOF measurement circuits on the same chip, leading to a simple, compact receiver architecture which in turn reduces the cost and paves the way for further miniaturization of the 3D imager. There are two approaches to on-chip measurement of TOF using short pulses: measuring the photon arrival time by means of a Time-to-Digital Converter (TDC) (Beer et al. Oct 2017; Keränen & Kostamovaara 2019; Niclass et al. 2014; Perenzoni, M., Perenzoni, D. and Stoppa 2017; Zhang et al. 2019) or a Time-to-Amplitude Converter (TAC) (Parmesan et al. 2014; Stoppa et al. 2009), and scanning of the sub-ns time windows over the range and recording the presence of photon arrivals within each (Ren et al. 2018; Ruokamo, Hallman & Kostamovaara 2019). It was the TDC approach that was chosen here, i.e. in addition to the 2D array of the SPADs, the receiver ICs also included an array of TDCs on the same die (as can be seen in Fig. 8).

As discussed, the TDC array is separated from the SPAD array in the following designs in order to improve the FF. There is a ratio $N>1$ between the number of SPADs and the number of TDCs ($N=N_{SPAD}/N_{TDC}$) in order to reduce the complexity of the receiver, i.e. every $N$ SPADs share one TDC. This structure means that not all the SPADs can be measured simultaneously and $N$ measurements are needed to cover the whole array (the drawbacks and benefits of this design choice are explained in detail in the following chapters). In addition to the $N_{TDC}$ TDCs, that measure the on-chip electrical ‘stop’ signals from their respective SPAD breakdowns, one extra TDC is added to measure the emission time of the laser pulse (‘start’ signal). The start signal, taken from the LD transmitter, marks the start of the travel time of the pulse ($t_{sharedstart}$), after which the laser pulse travels to the target and is reflected/backscattered to the receiver, generating a stop signal when detected in the optically sensitive active area of a SPAD, marking the end of the laser pulse’s flight ($t_{SPADstop}$). Thus, the travel time of the laser pulse is simply $t_{SPADstop}-t_{sharedstart}$. The TDC array then digitizes the travel time (i.e. the distance from the target) for all $N_{TDC}$ active SPADs. The TDC arrays of both chips which are presented in the following chapters are based on a delay-locked loop (DLL)
interpolation and a counter (Jansson, Mäntyniemi & Kostamovaara 2006; Jansson, Mäntyniemi & Kostamovaara 2009).

Having hundreds of TDCs on the same chip as a detector array can result in a temperature increase due to high power consumption in the IC, and as a consequence, an increase in the temperature-dependent DCR. In most cases involving 3D imaging applications, however, the background noise still dominates even over the increased DCR.

**Interface array**

In order to control the SPADs and connect them to the TDCs, each SPAD has an interfacing logic, a circuit schematic for which is shown in Fig. 11. The aim of this circuit is to implement controllable selection (activation/deactivation of the SPAD during a certain measurement), loading (biasing the SPAD above \( V_{br} \) ready for detection) and quenching (biasing the SPAD below \( V_{br} \)) both with programmable timing (i.e. gating) and with a minimal number of gates. The interface array is located beside the SPAD array.

Each pixel has its own high-active selection signal (Sel). When the SPAD is selected for measurement, it can be quenched and loaded, i.e. reverse-biased below and above \( V_{br} \), through \( M_q \) and \( M_l \), respectively. The load, quench and selection signals are produced by their dedicated logic based on the stored measurement setting and routed to each element of the array. Quench and Load signals are shared over the whole array.

Breakdowns (photon arrivals) are read from the anode, and the cathode is biased to a high positive voltage (\( V_{bias} \)). This \( V_{bias} \) has to be adjusted in a way that when the SPAD is quenched (the anode is biased to 3.3 V), it is reversed-biased slightly below the breakdown voltage. When the SPAD is loaded (anode at 0 V) and left floating for photon detection, this structure will lead to a nominal \( V_{ex} \) equal to the VDD of the technology (3.3 V).

A timing diagram including gating of the SPAD using interfacing electronics in one measurement cycle is shown in Fig. 12. At the beginning of each measurement the Sel signal is used to determine whether the SPAD is selected for measurement or not. After the arrival of Start, the chosen SPADs are loaded by means of a short Load pulse to make them ready for detecting incident photons during the measurement period. If no triggering occurs, the SPADs are actively quenched (via the zero active Quench pulse at the end of the gate window, Fig. 12,
SPAD 1, after a preselected and programmable period) and wait for the arrival of the next Start signal.

Upon detection of a photon, the current flowing through the SPAD charges the capacitance at the anode with a time constant dependent on the dynamic resistance of the diode and the overall capacitance at the anode. The SPAD will send stop signals to the time measurement circuit at the rising edge of the anode voltage. The inverter connected to the anode is designed to have a low threshold voltage to improve the switching speed in the case of photon detection. The tristate buffer controls whether the SPAD anode is connected to TDC or not. As the anode capacitance is charged, it gradually stops the current flow (the SPAD is self-quenched).

Fig. 11. Schematic diagram of the interfacing electronics of a SPAD (reprinted by permission from Paper I © 2015 IEEE).
Since a triggered SPAD remains quenched until the arrival of the next Start (transmission of the next laser pulse), only the first breakdown in each measurement cycle can be measured. Consequently, if a SPAD is triggered by dark or background noise before the arrival of the target photons, it will be blocked from detecting the returning laser photons. In order to prevent this, the Quench and Load signals (marking the beginning and end of the measurement time window) can be delayed with respect to Start (Fig. 12), so that the SPADs are operative only within the desired gate window (as explained in Section 3.2). The delay and width of the gate window can be adjusted with steps equal to the period of the reference clock (produced on or off-chip), e.g. TDC CLK. Fig. 12 also demonstrates how gating can be used to avoid background noise from blocking the SPADs, e.g. in case of SPAD 3 as opposed to SPAD 2. By defining a suitable gate window, the 3D measurement can be focused on a certain desired distance range, and the possibility of the SPAD being blocked by false triggering under bright ambient conditions or because of unwanted objects can be reduced.
4 Direct TOF 3D imager evaluation prototype

To test the performance of the dTOF focal-plane imaging concept, as described in Chapter 3, a receiver IC based on the architecture presented in Section 3.3.2 was designed and manufactured. The CMOS chip contains a 2D array of 9×9 SPADs with an array of 9+1TDCs with ~65 ps resolution. The IC was meant to serve as a receiver in a prototype to evaluate the possibility of combining the above-mentioned short high-energy pulsed LD transmitter and a high-resolution 2D SPAD array in a dTOF focal plane 3D imager in the later phases.

In this chapter the design and construction details of the SPAD receiver IC and its measured performance will be described (Section 4.1), the 3D imager prototype architecture (combining the receiver IC and an LD transmitter) will be explained and measurement results indicating its functionality and performance will be presented (Section 4.2).

4.1 9×9 SPAD/9+1TDC receiver IC

The receiver IC realized in a 0.35 µm HVCMOS technology consists of 81 SPADs in a 9×9 array and 9+1 TDCs ($N_{SPAD}/N_{TDC}=9$). During each measurement, only one of the 49 possible 3×3 sub-arrays is selected to detect photons, and this selection can be changed between measurements. Because of the SPAD/TDC ratio, nine measurements of 3×3 are needed in order to scan the whole 9×9 array. The on-chip TDC array has 10 identical channels: one start and nine stops (as described in section 3.3.2.3). A micrograph of a chip (size: 2.5 × 4 mm² including I/O pads) is shown in Fig. 13.

The SPAD array has a structure similar to that shown in Fig. 9. The size of the SPAD array is 330 µm × 330 µm, with the active area of each SPAD being 24 µm × 24 µm, resulting in a FF of 42.9%. The placement of the front-end electronics outside the SPAD array roughly doubles the capacitive load of the SPADs with the addition of the routing capacitance, and creates a slight wiring delay mismatch (SPAD anode to TDC input) across the array, which is less than 20 ps, and being static, can be easily compensated for by calibration.
The 9+1 TDCs are based on a counter and two interpolation levels. A multiplying delay-locked loop (MDLL) internally multiplies a 20 MHz input reference signal to 240 MHz, simultaneously acting as the first coarse interpolation level with a 4-bit result, generating 16 even-sized phases with identical successive delay elements (Jansson, Mäntyniemi & Kostamovaara 2009; Jansson, Mäntyniemi & Kostamovaara 2006). A 7-bit counter counts the rounds of the MDLL between the start and stop signals, providing a long ~80 m (2^7 × 4.2 ns) measurement range. The second fine interpolation level locates the stops within the resolution of the first interpolation level. The time interval between the asynchronous stop and the synchronized edge at the first interpolation level is solved with parallel load capacitor-scaled delay elements with a 5-bit resolution. The minor differences of ~8 ps between parallel elements are achieved by a small difference in their capacitive loads (Jansson, Mäntyniemi & Kostamovaara 2006).

The performance of the TDC was intentionally overdesigned in order to be able to measure the inherent timing properties of the system (e.g. characteristics of the SPAD array). However, in order to be in line with the concept presented above (matching TDC resolution with SPAD jitter and laser pulse width), the first three LSBs of the TDC results are ignored here, bringing the resolution to ~65 ps. The TDC has a power consumption of 150 mW (45 mA with a 3.3 V power supply) when operating at a clock frequency of 240 MHz, the main source of power consumption being the continuously circulating MDLL. The maximum pulsing rate the receiver chip can support covering the full range is about 1.1 MHz.
The receiver IC is capable of operating in both free running and gated mode, the principle of the latter having already been discussed in Fig. 12 (making SPADs operative for only a desired period synchronized with the transmitted laser pulse). At the beginning of each measurement and before the laser pulse transmission, one 3×3 sub-array is selected for measurement and the position and width of the gate window is defined with respect to the start signal (the laser pulse) with a ~4.2 ns resolution (the internal clock period of the TDC). After the arrival of the electrical start signal, its arrival time is measured by its dedicated TDC and the nine chosen SPADs are biased above breakdown for the predefined gate window. Upon photon detection, one or more SPADs send stop signals to their respective TDC at the rising edge of the anode voltage. The measured data are transferred out of the chip after the gate window is over in nine 16-bit words, with each word defining the interval measurement for one SPAD \((t_{\text{SPAD stop}} - t_{\text{shared start}})\). Each word is read out in two bytes (via a bidirectional 8-bit data path) with a maximum data transfer rate of 66 Mbytes/s. The gate window and sub-array selection settings can be changed after the transfer of the measurement data and before the arrival of the next start.

The DCR of the SPAD array was measured to be less than 150 kHz for 70% of the SPADs, with a few SPADs having DCRs higher than 1 MHz. The higher than expected DCR (Nissinen et al. 2011) was probably the result of a rise in the temperature of the chip owing to the TDC operation.

An important feature in the IC structure is a dense SPAD array implementation separate from the on-chip electronics. As mentioned, in such a tight implementation of the detector array, correlated noise (i.e. crosstalk) between the detectors might become an issue. The crosstalk probability and its dependence on various parameters such as distance between the devices, excess voltage, die thickness, etc. in CMOS SPAD arrays have been studied previously (Ficorella et al. 2016; Kindt, Zeijl & Middelhoek 1998; Lacaita et al. 1993; Nissinen et al. 2014; Rech et al. 2007; Rech et al. 2008; Sciacca et al. 2008; Vila et al. 2014), but the temporal correlation of the crosstalk coincidences has not been discussed extensively or has been reported with limited timing resolutions (Lubin et al. 2019; Mahmoudi et al. 2019; Nissinen et al. 2014; Sciacca et al. 2008; Vila et al. 2013; Wu et al. 2018). The temporal distribution of crosstalk events can become important specifically in dTOF 3D imaging, because it might be mistaken for the actual distribution of photon detections from the target, as discussed below.

The crosstalk and its characteristics in the first two immediate neighbours of an avalanching device were studied here, the focus being on measurement of the temporal behaviour of this noise source. The crosstalk was assessed through
coincidence measurements with a high-DCR SPAD acting as the crosstalk source (emitter). Then, to investigate the correlated noise in a practical dTOF application, coincidence measurements were repeated with the emitter triggered with pulsed laser illumination instead of its own noise (Paper IV).

The measurement results indicate both optical and electrical crosstalk when using a high-DCR pixel as the crosstalk emitter. The crosstalk probability in the first two adjacent pixels of the emitter was found to be 0.3% and 0.01%, with a distribution having FWHMs of 700 and 400 ps and tails (90% to 10%) of ~1.5 ns and ~1.2 ns, respectively (Fig. 14).

**Fig. 14.** Measured crosstalk distributions over time for (a) first and (b) second neighbouring detectors (with respect to the emitter breakdown, note the different scales in y axis) (reprinted by permission from Paper IV © 2018 The Optical Society).

The presence of a long tail in the crosstalk histograms and the wide FWHMs indicate electrical in addition to optical crosstalk. Both the direct and indirect optical paths from the emitter to any of the detectors have delays of <1 ps. Thus, in the case of only optical crosstalk, the distribution would represent the combined effects of the avalanche current build-up, spread and quenching in the emitter (i.e. the jitter of the emitter, ~100 ps FWHM), the relaxation time constant of the avalanche generated carriers (<1 ps), the added jitter of the detector (~100 ps FWHM), and the TDC uncertainty (~100 ps FWHM), which would lead to an FWHM of ~200 ps. The results suggest that the diffusion time of the minority carriers in the deep n-well due to electrical crosstalk contributes to the further widening of the distribution. On the other hand, the reduction in the crosstalk probability from the first to the second adjacent detector is greater than the theoretically expected rate of direct optical attenuation \((1/r^2)e^{-\alpha r}\), where \(r\) is the distance from the emitter, and \(\alpha\) is the absorption coefficient of silicon (Kindt et al.)
1998; Rech et al. 2007). This also suggests the presence of electrical crosstalk, which is attenuated with a higher rate proportional to $e^{-r^2}$ (Vila et al. 2014). The exponential shape of the crosstalk distribution in the first adjacent detector indicates that electrical crosstalk has significant involvement, whereas for the second adjacent detector, a distinct combination of a Gaussian distribution (due to optical crosstalk) and an exponential distribution (due to electrical crosstalk) is evident.

In dTOF measurements, when one SPAD is triggered with external short-pulsed (FWHM ~ 200 ps) illumination, extra correlated noise in the adjacent SPADs adds to the crosstalk noise, increasing the correlated noise in the first neighbouring pixel considerably. This additional noise is a secondary effect of the absorbed laser photons deep in the substrate of the SPAD under pulsed illumination.

The crosstalk measurements with active illumination were performed in the single-photon mode, but this only means that the avalanche initiation in the emitter of crosstalk would have a high probability of being caused by absorption of a single photon. On the other hand, there are multiple other photons from the laser pulse that might be absorbed in the deep n-well or substrate which do not lead to breakdown in the emitter but might trigger detectors through the aforementioned processes. This correlated noise is not directly related to an emitter avalanche event (rather, it is related to the time of arrival of the laser pulse at the emitter) but might be equally as important with regard to practical applications of pulsed-laser single-photon detection.

The correlated noise probability in the first neighbouring detector increased to 0.39% and 0.9% measuring with two laser pulses of wavelengths 630 nm and 870 nm, respectively. The correlated noise distribution in the first neighbouring detector in these two cases is shown in Fig. 15. The shorter penetration depth of the red laser photons (~3 µm) makes it more probable for the secondary laser photon absorption to affect the emitter rather than the detectors, hence leading to a smaller probability than the 870 nm laser case.
4.2 3D imager evaluation prototype

A graphical representation of the prototype 3D imager based on the 9×9 SPAD/9+1 TDC IC is shown together with a photograph of the physical implementation in Fig. 16(a) and (b), respectively. The measurement system uses a fibre and a collimator together with an engineered diffuser on the transmitter side to achieve a square-shaped homogeneous illumination with a divergence of 24.5 mrad and a positive lens on the receiver side with an effective diameter of 18 mm and focal length of 20 mm. Thus, the divergence of the receiver is approximately 15 mrad. A narrow-band interference optical filter (NBOF) is used to limit the optical bandwidth of the receiver to ~40 nm around the central wavelength of transmitted laser pulse in order to block most of the background at the same time, while accommodating the spectral drift of the laser pulse. An FPGA interface (OpalKelly XEM6001) is used to provide triggering signals for the LD driver and configuration data for the receiver IC, and to transfer the measurement data to a PC, where MATLAB GUI is used for real-time data analysis and 3D image construction.
The transmitter consisted of a bulk GaAs/GaAlAs DH LD (30 µm stripe width and 3 mm optical cavity length) operating in enhanced gain-switching mode at a wavelength of ~870 nm and a MOS current driver (as discussed in Section 3.3.1) producing a high-energy (~1 nJ), short (~125 ps FWHM) optical pulse (Ryvkin, Avrutin & Kostamovaara 2009). The peak current and pulse width of the drive current needed to produce the optical pulse were 6.5 A and ~1 ns, respectively. The typical shape of the driving current and corresponding optical output pulse are shown in Fig. 17.
The pulsing rate of the transmitter was 100 kHz in the measurement setup. In order to provide a low-jitter electrical start signal for time interval measurements, the current of the laser diode is sampled to eliminate the jitter of the LD driver as opposed to the FPGA, which provides the start signal. The whole LD transmitter electronics occupies a circuit board area of 2–3 cm² only, as shown in Fig. 16(b).

The single shot distribution of one SPAD measuring a flat, non-cooperative target at ~19 m is presented in Fig. 18(a). The measurement is carried out in the single photon detection mode, i.e. the received echo leads to a less than 10% detection rate in a SPAD, resulting in a relatively identical detection process independent of when the detection happens. The FWHM value of the distribution is ~150 ps, which matches well with the calculations in the case of a 125 ps FWHM laser pulse, ~100 ps FWHM SPAD jitter and ~70 ps FWHM TDC uncertainty (dominated by the quantization noise of the TDC). The asymmetry of the histogram (distribution tail) is caused by the diffusion of minority carriers in the SPAD and also the shape of the laser pulse (Fig. 17), accounting for nearly 30% of the detections.

In order to investigate the effect of histogram analysis on the precision achieved, TOF was resolved by considering \(N\) consecutive detections where \(N\) varied between 1 and 250. The measurements for each value of \(N\) were repeated \(10^4\) times and each time the histogram of \(N\) detections was formed and the TOF was calculated by weighted averaging. Then, a distribution of the measured TOFs was formed for each value of \(N\). The measurements were performed under laboratory conditions, so that the dominant source of noise was the dark count of the SPADs, and conditions (e.g. temperature, background illumination, target distance and reflectivity) were kept stable for the whole duration of the measurements.

Two cases, corresponding to presence or filtration of the tail, were considered in the interpretation of the histograms. In the case of filtering the tail a symmetrical histogram was assumed with respect to the peak and the detections that fell outside the time period of this ideal histogram were removed. The measured mean and sigma of the distributions versus the number of averaged detections \(N\) are shown in Fig. 18(b) for both cases. Filtering the tail obviously reduces the value of the mean. In both cases the uncertainty improves by a factor of \(\sqrt{N}\) considering \(N\) measurements with respect to the single shot sigma. Comparing the two cases, filtering the tail seems to improve the uncertainty by a factor of nearly 3, so that with the tail detections considered in averaging, around 350 detections of the target
are needed, for example, to achieve the same result as 50 detections with omission of the tail (15 out of the 50 are omitted detections of the tail).

Fig. 18. a) Single shot distribution of one SPAD operating in single photon mode with a flat target at ~19m, b) the measured mean and sigma of the distributions versus the number of averaged detections (N) in two cases corresponding to presence or filtration of the tail (inset: zoomed version for 1< N<50) (reprinted by permission from Paper III © 2016 The Optical Society).

If the received light power is increased so that it is higher (e.g. by a factor of ~10–10³) than in the single photon mode, multi-photon detection, where more than one photon is assumed to start the avalanche breakdown, becomes more probable. As a result, the build-up of the avalanche multiplication is faster, resulting in a walk error of up to ~100–200 ps (Paper II). This error, common to both linear and single photon optical detectors, is produced by the higher energy of the received optical echo, due to e.g. higher reflection properties or shorter distance of the target. In this
case the precision is improved compared with the single photon mode, since the
detection takes place mostly around the high power part and even the front edge of
the laser pulse. If the received light power is further increased to $10^7$ times higher
than the single photon mode (e.g. when measuring distances to a diamond-grade
target), a larger walk error of up to $500$ ps can be seen (Paper II). In this case, the
received echo is so strong that detections occur even before the steep rise in the
laser pulse, owing to spontaneous emission from the laser diode during the whole
current pulse and the effect of the super-luminescence at the onset of lasing.

3D images of three targets placed at distances of $19$–$20$ m, at which the FOV
of the detector array covers roughly an area of $27 \times 27$ cm$^2$ (i.e. each SPAD sees a
$3 \times 3$ cm$^2$ square) are shown in Fig. 19(a), (b) and (c). The first target (Fig. 19(a))
is a white pyramid with three steps, the second (Fig. 19(b)) a white paper ramp with
its surface at a 60-degree angle to the optical axis, one-third of which is covered
with a lower reflective material (brown carton), and the third target (Fig. 19(c)) a
flat plane of white paper with a cube placed on top, causing a $3$ cm step. The
distance for each SPAD is resolved by considering 50 detections, filtering out the
distribution tail detections, and averaging over the rest of the detections ($35$),
leading to a sigma value of $20$ ps ($3$ mm). The measured distances are calibrated
to compensate for the static delay difference between the detectors in the array. In
the measurements presented in Fig. 19, the detectors were operating in the single
photon mode. As a result, there is no need for walk error compensation schemes.
Fig. 19. a) 3D images of a pyramid with three steps, b) a ramp, one-third of which is covered with a low reflective material (brown cardboard), c) a white, flat plane with a cube placed on top, and d) a colour map representation of the detection probability for the ramp target (reprinted by permission from Paper III © 2016 The Optical Society).

Moreover, the characteristics of the histogram of hits can be further analysed for more information about the target area seen by one SPAD other than only a distance measure, e.g. the presence of a step within the FOV or target reflectivity. While operating in the single photon mode, the photon detection rate of each SPAD gives a measure of the reflectivity of the measured target spot, because the effects of the distance of the target spot and the illumination pattern are known and can be accounted for. A colour map representation of the photon detection rate of the SPADs for the ramp target after removing the dark noise is shown in Fig. 19(d). The average measured photon detection rate for white paper in this measurement was ~4.5%. The lower detection rates of the SPADs on the left side represent the lower reflectivity of left side of the target. In practical measurements, the prototype shown in Fig. 16 was able to follow in real time the movement of a match-box sized target (3×5×5 cm³) spinning on a wheel (diameter ~30 cm) with a speed of ~1 cycle per second at a distance of ~19 m.
The measurements presented above, at both the circuit and the system level, validated the proposed concept of a laser 3D imager based on the use of short, intensive laser pulses and separate configurable SPAD/TDC arrays as a viable solution for 3D imagers. This conclusion could be reached even though the prototype was not ideal in some aspects, e.g. the use of a fibre or diffuser instead of compact optics, the loss of almost half the illumination power due to unmatched transmitter/receiver divergence, non-optimum pulse energy, pulsing rate and wavelength (compared with achievable values of ~5 nJ/1 MHz/810 nm, for instance), and low spatial resolution.

Apart from the 3D imager prototype, the IC mentioned above was also used as a receiver for dTOF 1D rangefinding in paraxial LiDAR measurements (single optical axis). Paired with a high energy/short pulsed LD, the high-precision/long-range TDC (65 ps/80 m), the FF of ~50%, the use of parallelism and the ability to use short time gates of down to ~4.2 ns to reduce the blocking effect introduced by high background noise made the device well-suited for real-time laser ranging (Huikari et al. 2018; Huikari et al. 2017; Paper II). Use of a 2D array structure with a larger detector area (330 µm × 330 µm) alleviated the requirements with regard to the precision of the optics and the mechanical tolerance of the opto-mechanics. The received optical illumination (with a width of only some tens of µm) could be located anywhere within the frame of the 2D SPAD array, so that the corresponding 3×3 array (100 µm × 100 µm) could be selected for connection to the TDC array for measurement purposes. Movements of the target image at the detector surface (while measuring different target distances due to the paraxial optics) could be followed by means of selectable 3×3 sub-arrays within the 9×9 array.

In addition, the potential of the imaging concept for building a compact 2D line profiler for non-cooperative targets with a ~10 m distance range (Hallman et al. 2018) was evaluated using the developed IC. An LD transmitter (~1 nJ/140 ps) was combined with cylindrical lenses to illuminate a line-shaped ~45° FOV, and a similarly shaped FOV was realized with the IC at the receiver side by physically moving the IC to cover the FOV. 2D line-profiling can be advantageous in some applications, since it could potentially have less complicated electronics, higher detection efficiency and better tolerance of background noise. Demonstration measurements showed a line rate >15 Hz, a lateral resolution of about 5 mrad and cm-level single-shot precision at distances of more than 10 m from non-cooperative targets. This line of research was later continued in a separate study for the development of a fully-integrated SPAD-based line receiver and 2D laser line
profiler for solid-state LiDAR profiling applications (Keränen & Kostamovaara 2019a; 2019b).
5  The 8k-pixel dTOF solid-state 3D imager

To improve the dTOF laser 3D imager prototype in terms of spatial resolution, illumination power, optical efficiency, compactness, etc., a single-chip 32×128 SPAD/256+1 TDC receiver IC was realized in a 0.35 µm HVCMOS technology, and a solid-state 3D imager was build based on it and verified through measurements. The receiver was characterized using flood-pulsed illumination from an LD-based transmitter, which produced short, high-energy pulses (150 ps/3.8 nJ). Two detector/TDC ICs formed an 8-kilo pixel receiver, targeting a FOV of ~42° × 21° by means of simple optics. However, as briefly discussed in Chapter 6, block-based illumination, i.e. using the available average optical energy in blocks, is more efficient than flood illumination from the point of view of the system, and the receiver architecture was selected to match this type of illuminator as well.

In the following chapters, the design, implementation details and measured performance of the SPAD-based receiver IC will be explained in detail (Section 5.1), and the functionality and measured results of the 3D imaging system will be presented (Section 5.2).

5.1 32×128 SPAD/256+1TDC receiver IC

The IC consists of a 32×128 SPAD array and a 256+1 TDC array on the same die. The size of the chip (including I/O pads) is 6.6 mm × 5.5 mm. A micrograph of the IC and a zoomed version of a sample part of the SPAD array are shown in Fig. 20. The SPADs form a 32×128 rectangular array with each SPAD having a 40 µm × 40 µm pitch and a rectangular active area of 26 µm × 21 µm, leading to a FF of 35% for the light-sensitive part of the chip. The SPAD array structure is similar to that shown in Fig. 9, with the interfacing electronics placed in a separate array.

The block diagram of the IC is presented in Fig. 21. The SPADs are placed in 32 rows and 128 columns. Two adjacent rows (2×128 SPADs) can be connected to 256 TDCs in each measurement (i.e. a full scan of the whole SPAD array takes 16 measurement cycles). Row selection for connection to the TDCs is performed before every measurement based on control data received and stored in dedicated registers. Two TDC channels fit within the pitch of each detector column. The I/O pads were removed from the light-sensitive side of the IC (the bottom in Fig. 20), which makes it possible to bond two ICs in close proximity on a PCB and double the imaging resolution (Burri et al. 2014). The main part of the IC layout consists of 256 similar slices, each including 16 SPADs, their interfacing logic and 1 TDC.
interpolator. The red area in Fig. 20 shows the placement of such slices within the IC floorplan.

Fig. 20. Micrograph of the 32×128 SPAD/256+1 TDC array receiver IC (under CC BY from Paper V © 2020 Authors).

Fig. 21. Block diagram of the IC (under CC BY from Paper V © 2020 Authors).
The timing core, common to all the 257 TDCs (1 start and 256 stops), is composed of a counter and a delay-locked delay-line. Stabilization is achieved by means of a 200 MHz external reference oscillator. The delay-adjustable delay-line creates time samples for interpolation with ~78 ps resolution when the reference clock signal propagates through 64 identical delay elements. The 7-bit counter begins counting the clock cycles when the electrical start signal arrives at the IC. Each of the 257 channels consists of 64+7 registers (latches), which store the state of the sampled delay-line and the counter value. A synchronization logic between the delay-line and the counter registers synchronizes the counter sampling, so that the results are compatible (Jansson, Mäntyniemi, Kostamovaara 2009).

The TDC architecture provides a linear dynamic range of up to 640 ns under varying PVT conditions for all the time interval digitization between the start and the 256 stops. The power consumption of the TDC array, measured when operating with a 240 MHz reference frequency, is ~150 mW in its idle state (inactive SPAD array) owing to the constantly running DLL. During the active gate window, the outputs of the 64 delay elements simultaneously drive the 257 interpolation latches through long ~5 mm wires, leading to a transient power consumption of ~1.5 W. This translates into an average power consumption of 175 mW, for example, for measuring distances from targets located at up to 10 m at a pulsing rate of 250 kHz operating with a 240 MHz reference frequency.

The high power consumption of the TDC array leads to a temperature increase of nearly 10 degrees in the IC, which directly affects the breakdown voltages of the SPADs. As a result, adjustment of the bias voltage of the SPADs ($V_b$, in Fig. 11) according to the IC temperature (e.g. using off-chip voltage regulators) is needed to make sure that they are always operating with the highest possible $V_{ex}$ when active but are not conducting current when deactivated.

The measured RMS precision of the TDCs was 45 ps, but if the nonlinearities of the TDC are compensated for by post-processing of the results with an INL-LUT (integral nonlinearity look-up table) the precision can be improved to ~34 ps (Jansson et al. 2019), which matches well with the precision limitation due to quantization noise, LSB/$\sqrt{6}$. The INL-LUT was not used in the optical measurements, because the TDC precision is not the limiting factor in the system jitter.

A single measurement cycle of the IC is shown in Fig. 22. At the beginning of each measurement cycle the measurement settings are read into the IC: two adjacent rows are selected for measurement and the position and width of the gate window is defined with a resolution equal to the clock period of the TDC (~5 ns)
with respect to the start signal (the laser pulse). Gate windows with various widths (from 5 ns up to 640 ns in 5 ns steps) and various delays with respect to the laser pulse (from 5 ns to 635 ns in 5 ns steps) can be used depending on the measurement conditions.

Fig. 22. Measurement cycle for scanning the array and readout (under CC BY from Paper V © 2020 Authors).

At the falling edge of the Init pulse the measurement setting registers are updated and the SPADs are quenched. After the arrival of the electrical start signal, one TDC measures its arrival time. The chosen 256 SPADs are biased above the breakdown voltage for the predefined gate window (synchronized to the Start and TDC Clk). Upon breakdown in an active SPAD, one or more SPADs avalanche and their respective TDC(s) measure the detection time at the rising edge of the anode voltage (e.g. SPAD 1 in Fig. 22). If no detection occurs, the SPADs are actively quenched at the end of the gate window (e.g. SPAD 2 in Fig. 22).

After the measurement window is over, the measured data are transferred out of the chip once the DataReady signal is raised by the IC. The measured data include 258 16-bit words (i.e. the address of the active part of the array, the arrival time of the start signal and the breakdown time of the 256 active SPADs), which are read out of the IC through a 16-bit bidirectional data path. DataDir determines
the direction of data transfer, and DataClk is the data transfer clock in Fig. 21 and Fig. 22. The IC supports a maximum DataClk frequency of 100 MHz.

The measurement rate can vary depending on the DataClk frequency, the gate window width (covered range) and whether the settings need to be changed after a measurement cycle or not. With a distance range of ~10 m, if the data are transferred out of the IC with a DataClk frequency of 100 MHz and the measurement settings are changed after each measurement, loading the new measurement settings (20 ns), the initialization (20 ns), Start signal generation (20 ns), gate window width and DataReady signal generation (70 + 20 ns), and data readout (258 × 10 ns) would make a measurement rate of ~365 kHz possible. Therefore, a frame update rate of ~23 kHz (365/16) can be achieved with a laser pulse frequency of 365 kHz. Of course, the range and frame rate also depend in practice on the average power that the illuminator can provide.

The cumulative DCR of the SPADs is shown in Fig. 23(a) for each receiver IC separately (measured with $V_b = 22.4$ V). Three trends can be seen in the distribution. Nearly 60% of the SPADs have a DCR less than 37 kHz, but the distribution grows exponentially above 60% and 1% of the SPADs have DCRs higher than 1 MHz. There are a few hot pixels in each chip (with random placement), which are for practical purposes too noisy for use even in short-range measurements. The distribution of the DCR with a mean value of 139 kHz and a median of 22.7 kHz is shown in Fig. 23(b), while Fig. 23(c) shows the uniformity of the dark count rate in the two ICs.
Fig. 23. a) Cumulative percentage of DCR, b) DCR distribution and c) uniformity of DCR in the SPAD arrays of two ICs (under CC BY from Paper V © 2020 Authors).

The value of $V_b$ below which no triggering due to dark noise can be detected is shown for each SPAD on both chips in Fig. 24. A variation range of 550 mV can be seen among the SPADs of both arrays. This non-uniformity may be caused by a combination of factors: variation in $V_{ex}$ due to a drop in the high voltage distribution net, VDD/GND nets (dependent on the switching activity), $V_{br}$ non-uniformity due to imperfections arising in the manufacturing process, or temperature variation as a result of high current consumption in a section on-chip or other components on the PCB ($V_{br}$ grows linearly with temperature, ~40 mV per degree (Bronzi et al. 2012)). Because of this non-uniformity, most SPADs do not benefit from the highest possible excess voltage.
Fig. 24. $V_e$ distribution below which no triggering due to dark noise can be detected (under CC BY from Paper V © 2020 Authors).

The distribution of the measured arrival time differences recorded by each SPAD (by means of its respective TDC) when the whole detector array is exposed to the same laser pulse illumination is shown in Fig. 25. This distribution depicts the static delay difference between pixels caused on-chip from photon detection up until the respective TDC’s digital result is recorded. To obtain an accurate distribution, the laser beam was directed to the SPAD array by means of a multi-mode fibre and the results of numerous measurements were averaged for each pixel. The illumination power incident on the array was adjusted using neutral density filters to make sure the SPADs operated in single photon mode, thus avoiding multi-photon-induced walk error.

Three trends may affect the delay distribution pattern: in the horizontal direction, the placement of the TDCs with respect to the DLL (as can be seen in Fig. 20) has the dominant effect, i.e. the farther the TDC is from the DLL, the more delayed are the versions of the DLL transition edges that it receives. The TDC measuring the arrival time of the start signal is located after the farthest TDC from the DLL, and the photon travelling time for each SPAD detection is $t_{\text{SPAD}} - t_{\text{start}}$. Because of this floorplan and the TDC architecture, the delay difference pattern in the horizontal direction is mainly affected by the parasitics of the DLL signal distribution. The delay difference pattern in the vertical direction is mainly affected by the routing parasitics from the SPAD anodes to the interfacing electronics and from there to the respective TDC inputs. This leads to higher delay values for pixels of rows that are located farther from the TDC array. On the other hand, all the SPADs share the same deep n-well and hence the same cathode bias, but, as can be deduced from Fig. 24, $V_e$ is not equal for all the SPADs. This can cause weaker
avalanches, and as a result longer delay times for SPADs with lower $V_{ce}$. The delay map of Fig. 25 may be used as a calibration table in TOF measurements. The maximum measured delay differences among the arrays for the top and bottom ICs were 590 ps and 600 ps, respectively, which is in agreement with post-layout simulation results.

![Fig. 25. Colour map of the 2D on-chip delay distribution (under CC BY from Paper V © 2020 Authors).]

### 5.2 Solid-state 3D imager

A dTOF focal plane 3D imager (Fig. 26) was built by pairing the above-mentioned receiver IC with a quantum-well (QW) LD-based transmitter. Two ICs form a combined array of 8k SPADs (128×64) and are placed in the focal plane of a lens system with an aperture diameter of 5.6 mm and effective focal length of 6.67 mm, which is focused to infinity. A narrowband optical filter (FWHM ~20 nm) is combined with the lens to block unwanted ambient light. A zoomed photo of two bonded ICs with a separation of ~120 µm, which is about 3 times the pitch of one SPAD, is shown in Fig. 26(c).

The customized QW LD (3 mm cavity length and 90 µm active stripe width) produces high-energy optical pulses with a central wavelength of 810 nm and a spectral width of 4 nm FWHM. The LD and its driver operation principle have already been discussed in Section 2.2.1. An LCR transient-based pulse shape control produces current pulses of width ~3 ns and a peak of ~10 A, leading to a laser pulse width and energy of 150 ps (FWHM) and ~3.8 nJ at a pulsing frequency of 250 kHz. A commercially engineered optical diffuser is used to shape the beam into a rectangle illuminating a target FOV of ~42° × 21°.
The 3D imager is connected via a USB 3 FPGA-based interface (OpalKelly XEM-7310) to a Qt-based Windows UI, which is used for real-time measurement control, data collection, data processing and 2D (intensity)/3D (depth) video reconstruction and streaming. The whole system can be powered by the USB 3 connection.

Fig. 26. a) The dTOF focal plane 3D imager, b) the transmitter-receiver PCB and c) a zoomed view of the receiver (under CC BY from Paper V © 2020 Authors).
The detection probability of SPADs measuring distances from a flat white target located 1.5 m away is shown in Fig. 27(a). The measurement was performed with low ambient light and a low average backscattered signal, to make sure that the probability of a SPAD being triggered by noise before the arrival of the signal or a multi-photon-triggered avalanche by the signal is negligible. As a result, this figure also represents the distribution of laser illumination power among the pixels of the detector array. Considering the laser radar equation given by Eq. 1 (Wang & Kostamovaara 1994) for $N_{\text{meas}} = 1$, we can calculate the average probability of photon detection per transmitted laser pulse ($N_{\text{signal}}$) for a SPAD located near the optical axis. The following parameter values for Eq. 1 reflect the measurement shown in Fig. 27:

$$E_{\text{tr}}/N_{\text{SPAD}}$$ is the transmitted pulse energy seen by a SPAD located near the optical axis (~0.516 pJ). This has been verified using a photodetector (with an active area diameter of 1 cm) at 1.5 m on the optical axis measuring 125 nW. The other parameters are $\epsilon_{\text{opt}} = 0.8$, $\rho_{\text{target}} = 1$, $A_{\text{rec}} = 0.24$ cm$^2$, $R = 1.5$ m, $hc/\lambda_{\text{tr}} = 2.3 \times 10^{-19}$ J, PDP = ~4% at 810 nm and FF = 35%. This leads to a ~0.08 probability of
laser photon detection for central SPADs, which is in agreement with the measured results of Fig. 27(a). The illumination power per pixel active area naturally declines outwards from the centre of the image as a result of the non-ideal diffuser and light fall-off (proportional to $\cos^4(\theta)$, where ‘$\theta$’ is the off-axis field angle).

Considering the measurements of Fig. 27(a), the distribution of detections for three SPADs located at the centre and opposite edges of the array are shown in Fig. 27(b). The histograms have a FWHM of ~250 ps, which is a combined result of the laser pulse width (~150 ps FWHM), SPAD jitter (~100 ps FWHM), TDC RMS single-shot precision (~100 ps FWHM) and effects of the fluctuations in supply voltage due to switching activity. This FWHM confirms that the single shot precision of the system in distance measurement is better than +/-2 cm. There are different sources of static timing differences between pixels, either optical or electrical, such as lens distortion, field curvature, off-axis angle of the target and static timing offset of the on-chip and off-chip electrical signals.

Fig. 27(c) shows the laser pulse energy distribution in time. Considering the fact that the probability of a signal detection per single emitted laser pulse is <<1 in dTOF 3D imaging, having such a short pulse envelope of 150 ps with high peak power (compared with ns pulses with the same pulse energy) improves the SNR and the single shot precision of the measurement.

An example 3D (point cloud) image taken of a lobby, with the furthest point of the scene nearly 6.5 m away from the 3D imager, is shown in Fig. 28(a). A high-resolution photo of the site concerned is also presented for reference in Fig. 28(b). Cross-correlation (sliding dot product) with a filter was used here to translate the histograms to distance. The filter and histogram before and after filtering for a pixel within Fig. 28 (Row 13/Col 33: armrest of the smaller sofa) are shown in Fig. 29. The filter is a normalized version of a histogram of hits collected by one SPAD in single photon mode measuring the location of a target placed at a specific distance from the camera. As can be seen in Fig 29(b), the signal might not be distinguishable from noise in cases of low SNR when only considering the peak, but the SNR can be improved to ~7 by the method described above and the signal can then be recovered. After filtering, a weighted average of a few points around the maximum of the filtered histogram was considered to represent the measured distance.
After resolving the measured distances for each pixel, calibrations were performed to remove the effect of barrel distortion introduced by the receiver lens and static
on-chip (Fig. 25) and off-chip (e.g. start signal delay) electrical offsets. The image was taken using a laser pulsing frequency of 250 kHz with a 1-second acquisition time in normal office lighting. It should be noted that no post-processing other than simple noise removal was performed on the point cloud data. A few points whose measured distances compared with those of all the neighbouring pixels differed more than a given threshold (e.g. 50 cm), were considered to represent noise and removed.

3D point cloud images of a person standing in front of a wall with a ball in the bottom right corner taken at frame rates of 2, 5, 10, 20 and 40 fps are shown in Fig. 30. The wall is located at a distance of ~3.3 m, and the measurements were performed in normal office lighting. In these measurements the LD transmitter provides a flood illumination with a total average optical power of ~1 mW, i.e. the effective average optical power illuminating the 1/16 of the FOV (corresponding to the active block in the detector array under measurement) is only ~65 µW. These results demonstrate the dependence of the quality of the recorded 3D range image on the effective frame rate. Uncalibrated histograms of detections before and after filtering for a specific SPAD (Row: 13/Col: 34) for two frame rates of 2 and 40 fps (corresponding to the highest and lowest frame rates of the 3D images in Fig. 30) are shown in Fig. 31.

Fig. 30. 3D images of a scene taken at different frame rates (under CC BY from Paper V © 2020 Authors).
Fig. 31. Distribution of photon detections for a SPAD (Row: 13/Col: 34) with different frame rates based on the measurements in Fig. 30.

The measurements presented above were carried out in office lighting (corresponding to ~200 lux ambient light). Fig. 32 shows the results of measurements of a scene in the presence of background noise. The ambient light is created by lamps, and the distance of the wall from the camera is ~3.5 m (a high-resolution photo is presented in Fig 32(a) for reference). Fig. 32(b) shows the measured average noise detection probability of the SPADs per TDC bin, when measuring a 20 ns gate window which is set to start after the arrival of the laser pulse echo (to make sure that all the detections are generated by noise). As can be seen, the background lighting condition is not uniform within the entire scene. The maximum background level was measured to be ~0.06 Wm⁻²nm⁻¹) at 810 nm corresponding to ~6.5 klux.

Fig. 32(c) shows the detection distribution results of a SPAD measuring the left side of the ball in case of a gate window starting ~15 ns before the target for a total of 3125 laser shots (corresponding to a frame rate of 5 fps). The background noise level of 6.5 klux, according to Eq. 2 for the 3D imager specification presented above, leads to a mean time of ~90 ns between background photons hitting the SPAD. With a target situated at a distance of 3 m (~20 ns), the SPADs would be blocked from detecting the target, as the result of background noise, in ~15% of the laser shots. Fig. 32(d) shows a point cloud presentation of the measured scene at a frame rate of 5 fps.
Fig. 32. a) High-resolution photo of the scene, b) average noise detection probability per TDC bin measured with a gate window of ~20 ns starting after the arrival of laser pulse echo, c) detection distribution for a SPAD measuring the surface of the ball with a gate window starting ~15 ns before the target at a frame rate of 5 fps, and d) 3D point cloud with a gate window starting ~15 ns before the target at a frame rate of 5 fps.
6 Discussion

The main aim of this research was to propose and verify a customizable compact solid-state approach 3D range imaging based on dTOF which combines a short, intensive pulsed laser transmitter with a SPAD array/TDC array single-chip receiver. Two receiver ICs were developed and evaluated in two 3D imaging systems. The final result was a solid-state 3D imager with a range of a few metres and a spatial resolution of 8k pixels (42° × 21° FOV) and a frame rate of up to 20 fps. The 3D imager was a combination of two single-chip SPAD/TDC (128×32/256+1) array receiver ICs and a customized LD-based transmitter (150 ps/3.8 nJ/810 nm/250 kHz) in a volume of 5×7×4 cm³. The receiver ICs were designed with separate TDC and SPAD arrays to allow a higher FF of ~40%. This strategy differs from the mainstream smart-pixel implementations in SPAD/TDC-based TOF receiver chips, in which a single pixel includes interfacing and time interval measurement electronics, inevitably resulting in low FF. Another feature of the receiver IC was that each TDC was shared between multiple SPADs, instead of the regular flash implementation (one photon arrival-time measurement unit per SPAD).

On the transmitter side, an average optical illumination power of ~1 mW was produced at a relatively moderate pulsing rate of 250 kHz. Moreover, the transmitter was implemented in a small size, consisting of a single custom-designed LD and its on-board driver. This transmitter approach, i.e. high-energy short laser pulses, allows the system complexity to be markedly reduced, since the high pulse energy can already achieve a range of a few metres to non-cooperative targets with a reasonable frame rate without the need for a high pulsing rate (receiver complexity). At the same time, the inherent precision is defined by the short pulse width of ~150 ps.

In the prototype presented in Chapter 5, the receiver IC was paired with flood-pulsed illumination, but a receiver architecture was used in which the SPAD array was divided into N predefined blocks (N_{SPAD}/N_{TDC}=N>1), and the TDC resources were allocated to only one of the blocks in each measurement. In other words, for the sake of simplicity, the complete light-sensitive area of the receiver array was illuminated, despite the fact that only one-sixteenth of the SPADs could be activated at any given time. To address this inefficiency, a major improvement at the system level can be achieved using block-based illumination. In a block-based transmitter architecture all the illumination power (typically limited by the
available high-speed LD drive current) is concentrated spatially only in the region of interest (in the selected active part of the detector array in each measurement).

This is a more effective method than the mainstream approach of distributing the illumination power over the whole array and having a TDC for each detector (a per-pixel TDC). When N illumination blocks are used (Fig. 33, \(N=16\)), the number of on-chip TDCs is reduced by a factor of \(N\), but the measurement time remains the same (as with flood illumination and per-pixel TDCs). The reason is that the number of laser pulses needed per pixel element is \(N\) times less due to the \(N\) times higher probability of detecting a back-reflected/scattered photon per measurement (assuming the same average illumination powers and bearing in mind that in practical measurement situations the signal detection probability per pixel per single emitted pulse is \(<1\)). Because of this reduced number of pulses per valid detection, the number of background hits is also reduced by \(N\), leading to an improvement in the SNR by a factor of \(N^{0.5}\). This in turn translates into a frame rate improvement of \(N\), while maintaining the same range (and SNR), or improving the measurement range by \(N^{1/4}\), assuming the frame rate is kept constant. An additional advantage of the block-based illumination approach is that it enables the illumination to be focused on a specific area of interest within the FOV of the system at a given time.

In practice, block-based illumination can be realized in solid-state systems by means of an array of LD bars or a VCSEL array with addressable element areas, which can be separately chosen and activated to be driven by a current driver for a selected number of laser pulses. A simplified presentation of such an architecture is given in Fig. 33.

![Fig. 33. Block-based illumination scheme using an addressable LD bar as the emitter (under CC BY from Paper V © 2020 Authors).](image-url)
A comparison between the performances of two hypothetical examples of transmitter/receiver architectures based on the results acquired by measuring the prototype presented in Chapter 5 is shown in Table 1. It is assumed that spatial resolution (128×64 SPADs), LD illumination power and the optics are kept the same as in the above prototype, and the achievable performance with different transmitter architectures and receiver complexities (i.e. number of TDCs) are compared. The first row of the table presents the performance of the system for whose measured results were presented in this thesis. A frame rate of 5 fps was achieved when covering a 5 m range, which serves as a basis for calculating the performance of the other rows while maintaining the same frame point cloud quality (SNR, noise dominated by background illumination).

The second row depicts the case of a flash-type receiver (all the SPADs can be measured in one measurement, i.e. for each laser pulse) combined with flood illumination in the transmitter. Since there is always a trade-off between achieving a higher frame rate or a longer range, both extremes of the trade-off are shown in the row, i.e. a higher frame rate of 80 fps with a maximum range of 5 m, or a longer range of 10 m while keeping the frame rate at 5 fps. The third row presents the calculated performance of a system using the block-based receiver architecture when paired with the same block-based architecture in the transmitter.

As can be seen in Table 1 (and based on the explanations presented above), the block-based transmitter/receiver solution would reduce the complexity of the receiver IC drastically and would simultaneously offer better performance than a flood illumination/flash detector architecture, but at the cost of a more complex transmitter.

Table 1. Comparison of homogeneous and block-based illumination (under CC BY from Paper V © 2020 Authors)

<table>
<thead>
<tr>
<th>System approach</th>
<th>Illumination type</th>
<th>Number of TDCs (64x128 SPADs)</th>
<th>Achievable Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Achievable Performance</td>
<td>Optimized for highest range</td>
</tr>
<tr>
<td>This work</td>
<td>flood</td>
<td>4x128 (16 blocks)</td>
<td>5 m @ 5 fps (measured)</td>
</tr>
<tr>
<td>Flash</td>
<td>flood</td>
<td>64x128 (per-pixel TDCs)</td>
<td>10 m @ 5 fps 5 m @ 80 fps</td>
</tr>
<tr>
<td>Block-based</td>
<td>16 blocks</td>
<td>4x128 (16 blocks)</td>
<td>20 m @ 5 fps 10 m @ 80 fps</td>
</tr>
</tbody>
</table>

Although the block-based illumination was not implemented in the research reported here, the receiver IC was designed with later development of the
architecture in mind. On the other hand, comprehensive tests on the IC and the 3D imager prototype (with flood illumination) were carried out, verifying achievement of the performance needed in medium-range applications, e.g. within the fields of robotics, automation and security.

Although the major improvement to the system is the matching block-based illumination, there are a few other possibilities for achieving better performance. Diffusers can be grown on top of the LDs to make the system even more compact, and the laser pulse energy can be increased further to ~10 nJ by using a multi-junction structure, for example. It is also possible to use a standard commercially available DH pulsed LD, which would increase the maximum pulse energy to ~100 nJ while limiting the minimum pulse width to 1–2 ns (Heinrich et al. 2018; Liero et al. 2020).

Since the IC can support laser pulsing frequencies of up to 365 kHz (while covering a 10 m range), the average transmitted power could also be increased by using higher LD pulsing rates. As a result of having a TDC depth (640 ns) that is nearly 5 times larger than the practical range, a train of multiple laser pulses (e.g. 100 ns apart) can be used instead of a single laser pulse in one measurement under low ambient light conditions to improve the exposure time per frame. It should be noted that the 3D image results presented here are based on simple data processing. Further advanced computational imaging approaches can be used for both distance approximation and point cloud data processing.

As pointed out in the introduction, there have been many studies on the development of optical solid-state 3D imagers, e.g. (Henderson et al. 2019; Horaud et al. 2016; Mitev & Pollini 2015; Payne et al. 2014; Vornicu, Carmona-Galán & Rodríguez-Vázquez 2017; Walker, Richardson & Henderson 2011), to name just a few. A thorough comparison of state-of-the-art solid-state 3D imagers would be a difficult task, since they are based on different approaches and aim at different applications, i.e. they are designed with specific circuit and system-level performance requirements in mind (FOV, 2D resolution, frame rate, range, etc.). Table 2 is an attempt at comparing a few recently developed state-of-the-art pulsed TOF solid-state 3D imagers and the corresponding receiver ICs, the emphasis being on ones with system-level realization (i.e. both transmitter and receiver) and aimed at small size/short range applications (up to tens of metres). The 3D imager realization (Ruokamo, Hallman & Kostamovaara 2019), presented in the second column of Table 2 was developed in the same research group, thus benefitting from the same transmitter approach while implementing a different receiver architecture, one based on the sliding time-gate technique in a SPAD array, in which the time-
gate is swept over the desired range in small steps over successive measurements and the presence of a detection is recorded in each step for each SPAD within the array. The length of the time-gate can be set in order to reduce the effect of background noise.

Table 2. Recently developed pulsed TOF receiver ICs and solid-state 3D imager systems (under CC BY from Paper V © 2020 Authors)

<table>
<thead>
<tr>
<th>Ref</th>
<th>This work (Ruokamo et al. 2017; Bellisai et al. 2013; Lussana et al. 2015; (Zhang, Chao et al., 2018))</th>
<th>(Ruokamo et al., 2019)</th>
<th>(Bellisai et al. 2013; Guerrieri et al., 2010)</th>
<th>(Villa et al., 2014)</th>
<th>(Zhang, Chao et al., 2018)</th>
<th>(Henderson et al., 2019)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOS Tech. (µm)</td>
<td>0.35 HV</td>
<td>0.35 HV</td>
<td>0.35 HV</td>
<td>0.35</td>
<td>0.18</td>
<td>0.04/0.09&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Pixel size (µm)</td>
<td>40 µm</td>
<td>50x100</td>
<td>20 µm</td>
<td>30 µm</td>
<td>28.5 µm</td>
<td>38.4 µm</td>
</tr>
<tr>
<td>2D resolution (µm²)</td>
<td>128x32</td>
<td>80x25</td>
<td>32x32</td>
<td>32x32</td>
<td>32x32</td>
<td>256x256&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Array FF (%)</td>
<td>35</td>
<td>32</td>
<td>3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3</td>
<td>28</td>
<td>51</td>
</tr>
<tr>
<td>Chip size (mm²)</td>
<td>5.5x6.6</td>
<td>5x5.7</td>
<td>3.5x3.5</td>
<td>9x9</td>
<td>2x5</td>
<td>2.5x2.5</td>
</tr>
<tr>
<td>IC Power (mW)</td>
<td>180</td>
<td>66</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>77.6</td>
</tr>
<tr>
<td>Single-shot prec.</td>
<td>&lt;2 cm</td>
<td>2 cm</td>
<td>27 cm</td>
<td>1 cm</td>
<td>&lt;1 cm</td>
<td>&lt;2cm</td>
</tr>
<tr>
<td>FOV (degrees)</td>
<td>21x42</td>
<td>18x28</td>
<td>40x40</td>
<td>6.9x6.9</td>
<td>40x40</td>
<td>1.2x1.2</td>
</tr>
<tr>
<td>Range</td>
<td>3.5 m</td>
<td>3.5 m</td>
<td>14 m</td>
<td>8 m</td>
<td>0.7 m</td>
<td>50 m</td>
</tr>
<tr>
<td>@ Fr. rate</td>
<td>@5 fps</td>
<td>@14 fps</td>
<td>@25 fps</td>
<td>@6 fps</td>
<td>@&lt;1 fps</td>
<td>@30 fps &lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>λ (nm)</td>
<td>810</td>
<td>870</td>
<td>808</td>
<td>750</td>
<td>637</td>
<td>671</td>
</tr>
<tr>
<td>Illum. P. (mW)</td>
<td>0.7</td>
<td>6.2</td>
<td>~200</td>
<td>90</td>
<td>2</td>
<td>1.8</td>
</tr>
<tr>
<td>Pulse width</td>
<td>150 ps</td>
<td>110 ps</td>
<td>150 ns</td>
<td>100 ps</td>
<td>40 ps</td>
<td>100 ps</td>
</tr>
<tr>
<td>Pulsing rate</td>
<td>250 kHz</td>
<td>700 kHz</td>
<td>1.7 MHz</td>
<td>40 MHz</td>
<td>40 MHz</td>
<td>1.9 MHz</td>
</tr>
<tr>
<td>pTOF technique</td>
<td>direct</td>
<td>gate scan</td>
<td>indirect</td>
<td>direct</td>
<td>direct</td>
<td>direct</td>
</tr>
<tr>
<td>Background noise</td>
<td>10 klux</td>
<td>10 klux</td>
<td>NK</td>
<td>&lt;50 lux</td>
<td>dark</td>
<td>1 klux</td>
</tr>
<tr>
<td>System size (cm³)</td>
<td>5x7x4</td>
<td>NK&lt;sup&gt;‡&lt;/sup&gt;</td>
<td>6x5x8&lt;sup&gt;‡&lt;/sup&gt;</td>
<td>6x8x8&lt;sup&gt;‡&lt;/sup&gt;</td>
<td>NK</td>
<td>NK</td>
</tr>
</tbody>
</table>

<sup>a</sup> excluding the transmitter, <sup>b</sup> microlens array used to make up for low FF, <sup>c</sup> 3D stacked, <sup>d</sup> Not Known, <sup>‡</sup> receiver aperture not specified, <sup>§</sup> 64x64 in LiDAR measurements.

The 3D range imager prototype designed and characterized in this thesis is one of the rare complete realizations of solid-state pulsed TOF 3D range imagers, in the sense that both the illuminator and the receiver are realized in a single miniaturized entity. Most previous research has focused on the design of receiver electronics starting from the point of view of flood illumination, and demonstrating mainly circuit-level measurement results. A general conclusion to rise from the comparison (Table 2) is that even with flood illumination, the results achieved by the 3D range imager prototype developed in this thesis compare quite favourably with the prior
state-of-the-art. Thus, it can be concluded that the suggested high energy/high-speed semiconductor LD concept used in dTOF measurement with a CMOS SPAD/TDC array receiver could pave the way to miniaturization of future solid-state 3D imagers. Sharing time measurement units between detectors and concentrating the laser illumination in both time (short sub-ns pulses) and space (targeting only the active part of the detector array) would translate to a matched combination of solid-state scan of illumination and detector sections. As a result, all the optical power would still be used in each measurement, while the complexity of the receiver would be drastically reduced.
7 Summary

Improving environment awareness through the integration of 3D imaging technology into means of transport, industrial machines, hand-held gadgets, etc. has been a few of the important commercial applications of 3D imagers. Such applications require a 3D imaging architecture that can provide a short-range (up to ten metres), cm-level precision, a frame rate of tens of fps, small size (~1 cm³), low cost, a spatial resolution of ~ 10k pixels and eye safety measures. Above all, however, integration demands solid-state implementation and flexibility in tailoring the above performance parameters to the needs of each specific application.

Many solid-state 3D imager approaches exist, such as stereoscopic vision, structured light and TOF, each with its own architecture-dependent performance strengths and limitations. One promising solid-state approach for developing miniaturized 3D imagers is the laser dTOF technique. The transmitter in such a 3D imager is usually realized to provide pulsed illumination covering the area of interest (FOV of the system). A 2D SPAD array together with an arrival-time estimation technique is used to detect and measure the travel time of the laser pulse photons. The spatial location of each measured point in the 3D image is defined by the receiver optics and the placement of the SPAD within the 2D array, and the third dimension, i.e. depth, is defined by the travel time of the detected photons. Such dTOF 3D laser imagers have so far shown good performance, but at the cost of quite sophisticated transmitters or demanding speed and functionality as receiver requirements.

The work presented in this thesis involved the use of a customized LD-based transmitter, which through operation in enhanced gain-switching mode, can provide ~150 ps/~3.8 nJ pulses. Having such a short energetic pulse brings down the requirements on the measurement rate, such that measurements within the range of a few metres become accessible with pulsing rates of a few hundred kHz and frame rates of up to some tens of fps. The receiver is based on a combination of a SPAD array and a TDC array on a single CMOS chip. Separation of the realization of the detector array from that of the time measurement array is chosen in the receiver IC design to ensure high FF in the light-sensitive part of the chip. The 3D imaging architecture offers high spatial resolution (since arrays of up to some tens of thousands of SPADs can be implemented in CMOS technologies), cm-level single shot precision and a good SNR under bright conditions (using short pulses),
and has a potential for low-cost miniature implementation (as the result of simultaneous digitalization of both the transmitter and receiver).

In order to study and verify the proposed architecture, the first step was to design and develop a receiver IC in a 0.35 µm HVCMOS technology which included a 9×9 SPAD and a 9+1 TDC array. This IC was combined with a customized bulk LD transmitter (~1 nJ/100 ps/100 kHz) in a measurement system using a fibre and collimator together with an engineered diffuser on the transmitter side and a positive lens on the receiver side. A single shot precision of ~150 ps was achieved measuring 3D targets located at ~20 m. The measurement results suggested that the laser 3D imager concept was an adaptable low-cost solution for 3D imagers, despite the prototype’s non-idealities in terms of opto-mechanics and low spatial resolution (covering a 15 mrad divergence angle). To make sure that the strategy of using a separate dense SPAD array does not lead to unwanted correlated noise, crosstalk was studied comprehensively through coincidence measurements. The design of this receiver IC also paved the way for many further developments of 1D rangefinding and 2D line-profiling systems within the research group.

A second receiver IC (32×128 SPAD/256+1 TDC array) and 3D imaging system was designed to improve the prototype in terms of 2D resolution, illumination power, optical efficiency, compactness, etc. The solid-state 3D imager was characterized by means of measurements. The system used flood pulsed illumination from an LD-based transmitter which produced short, high-energy pulses (~150 ps/3.8 nJ). Two detector/TDC ICs formed an 8k pixel receiver targeting a FOV of ~42° × 21°.

Another receiver IC feature was sharing of the TDC between pre-determined SPAD array blocks with a 16/1 ratio, the aim being to support the system-level idea of block-based illumination, which leads to a simple, straightforward IC structure and a high FF of the sensor area of the chip, and at the same time has potential for improved system-level performance relative to the mainstream flood illumination and per-pixel TDC approach.
List of references


Ficorella, A., Pancheri, L., Dalla Betta, G. F., Brogi, P., Collazuol, G., Marrocchesi, P. S., ...
In 2016 46th European Solid-State Device Research Conference (ESSDERC) (pp. 101-104). IEEE.


Henderson, R. K., Johnston, N., Hutchings, S. W., Gyongy, I., Al Abbas, T., Dutton, N., ...


Ito, K., Niclass, C., Aoyagi, I., Matsubara, H., Soga, M., Kato, S., ...


Original publications

This thesis is based on the following publications, which are referred to throughout the text by their Roman numerals:


Reprinted with permission from IEEE (Papers I and II), The Optical Society (Papers III and IV) and under CC BY license from Authors (Paper V).

The original publications are not included in the electronic version of the dissertation.
735. Laukka, Aleksi (2020) The effects of microalloying on the scale formation of AISI 304 stainless steel in walking beam furnace conditions

736. Ali, Samad (2020) Learning-based predictive resource allocation for machine type communications


738. Wijewardhana, Uditha (2020) Sparse recovery algorithms for streaming and multidimensional signals

739. Li, Panpan (2020) Deep eutectic solvents in the production of wood-based nanomaterials

740. Nuortimo, Kalle (2020) Public acceptance in energy technology development and deployment: opinion mining case study


742. Määtyniemi, Pekka (2020) Chemical precipitation in the simultaneous removal of \(\text{NH}_4\text{N}^+\text{N}\) and \(\text{PO}_4\text{P}^2\text{P}\) from wastewaters using industrial waste materials


744. Frondelius, Tero (2020) Development of methods in engine design process

745. Särestöniemi, Mariella (2020) UWB in-body propagation and radio channel characteristics for wireless body area network applications

746. Ashraf, Faisal Bin (2020) River regimes and energy demand interactions in Nordic rivers

747. Kuula, Seppo (2020) Continuous co-creation of knowledge-intensive business services

748. Mustonen, Erno (2020) Vertical productisation over product lifecycle: co-marketing through a joint commercial product portfolio

749. Halttula, Heikki (2020) Enhancing data utilisation in the construction project lifecycle through early involvement and integration

750. Nguyen, Hoang (2020) Ettringite-based binder from ladle slag: From hydration to its fiber-reinforced composites

Book orders:
Virtual book store
http://verkkokauppa.juvenesprint.fi
Sahba S. Jahromi

SINGLE PHOTON AVALANCHE DETECTOR DEVICES AND CIRCUITS FOR MINIATURIZED 3D IMAGERS