Megapixel SPAD cameras for time-resolved applications

Kazuhiro MORIMOTO

EPFL

École
polytechnique
fédérale
de Lausanne

2020

Acknowledgements

This PhD thesis would not have been completed without supervision and support by many people. First of all, I would like to express my deepest gratitude to my supervisor, Professor Edoardo Charbon for his 3 years of intensive guidance and support for my life in Switzerland. I clearly remember when I first met Edoardo in Shizuoka University, Japan in 2015. After his invited talk, I got a chance to talk with him in person for 30 minutes, and I conveyed my wish to join AQUALab as an external PhD student. When I passed the interview in 2016, he said to me, "you are the third staffs joining our group from Japan. The previous staffs made significant contributions, and I expect the same from you. DON'T DISAPPOINT ME". During my PhD study, I always felt a pressure from those words, which also gave me a lot of energy for research activities.

I would like to thank all the AQUALab members in EPFL and TU Delft. I am especially grateful to Dr. Claudio Bruschini for fruitful technical discussion and his help in IP contract management, and Ms. Brigitte Khan for her kind support and advice to the administrative procedures and my family's daily lives. I also want to give a special thanks to my colleagues, Andrei Ardelean, Ming-Lo Wu, Arin Can Ulku, Preethi Padmanabhan, Ivan Michel Antolovic, and Myung-Jae Lee, who have collaborated for the megapixel SPAD project and co-authored journals. I have learnt a lot from the colleagues who are highly experienced in device, IC, PCB design, firmware and software development, etc. I would also like to thank other colleagues and former staffs, Samuel Burri, Scott Lindner, Yatao Peng, Emanuele Ripiccini, Esteban Venialgo, Jad Benserhir, Simone Frasca, Francesco Gramuglia, Bedirhan Ilik, Utku Karaca, Pouyan Keshavarzian, Ekin Kizilkan, Andrada Alexandra Muntean, Andrea Ruffino, Jiuxuan Zhao, Augusto Ximenes, Augusto Carimatto, and Chao Zhang for sharing their knowledge and experience.

My PhD activity in EPFL was financially supported by "Studying Abroad Program" in Canon Inc. In spite of more than 30 years of successful history in human resource development across the company, this program had been locally suspended for 15 years in our semiconductor division for various reasons; my first application for the program in 2015 was immediately rejected. However, the former and present heads of our division, Shunsuke Inoue and Takeshi Ichikawa understood the importance of my PhD study abroad, and tried their best to change the long-lasting policy. At last, my second application in 2016 was accepted thanks to their efforts, and I got a precious opportunity to live in Switzerland for two and a half years. I would like to express my special appreciation and respect to them for their decision and guidance.

Acknowledgements

Abstract

Progress in optics technology has enriched our daily lives in many ways. Various types of single-photon detector technologies have emerged in the last century. Among these, single-photon avalanche diode (SPAD) technology outperforms other detectors for its room temperature operation, picosecond timing resolution, manufacturability, and scalability. Since 2003 when the first complementary-metal-oxide-semiconductor (CMOS)-compatible SPAD array was demonstrated, the exponential growth of CMOS process technology has served as a driving force of scaling and miniaturization of SPAD arrays. In 2017, the largest SPAD array size of 512×512 was reported, while the smallest SPAD pixel pitch of 3 μm was demonstrated in the same year. In parallel to such a continuous scaling, researchers have adopted the SPAD sensors to diverse application fields. Particularly, SPAD image sensors are considered promising for single-photon sensitive time-resolved applications, such as time-of-flight (ToF) ranging, fluorescence lifetime imaging microscopy (FLIM), space applications, and Raman spectroscopy, as well as other scientific imaging applications. Common to these applications is the growing demand for larger SPAD arrays to enhance spatial resolution, dynamic range, data acquisition speed, and overall functionality.

Achieving a megapixel array has been one of the most important milestones for SPAD research for over 10 years, but also one that has proven as elusive. To accomplish this goal, we start with a theoretical analysis of scaling laws for SPAD performance characteristics to clarify the underlying tradeoff relations in the SPAD pixels below 10 μm in pitch. To address the severe degradation of fill factor in the miniaturization, we propose a novel guard-ring-sharing technique, potentially pushing the limit of pixel miniaturization in conventional SPAD structures. The proposed concept is verified by extensive characterization in 4×4 test SPAD arrays, where the world's smallest pixel pitch of 2.2 μm was achieved.

Next, we propose a new-generation SPAD structure: charge focusing SPAD. The proposed concept is promising for overcoming a fundamental tradeoff between photon detection efficiency (PDE) and dark count rate (DCR), while suppressing major correlated noise, hot pixel population, and SPAD-originated power consumption. A proof-of-concept 128×128 charge focusing SPAD image sensor is developed for feasibility study towards low-light imaging applications, and we demonstrate the world's lowest DCR density of 0.015 cps/μm^2 at room temperature.

To open the door to scalable time-resolved SPAD arrays, we propose novel system concepts of successive approximation ToF and coded time gating. The concepts are implemented in backside-illuminated (BSI) 3D-stacked SPAD sensors based on 45 nm/22 nm CMOS process

Abstract

with the resolution up to 0.5 megapixel.

As a highlight of this thesis, we demonstrate the world's first 1 megapixel SPAD image sensors based on time-gating approach. A proposed readout circuit sharing technique achieves significant fill factor improvement and orders of magnitude suppression of power consumption under strong illumination. The best-in-class noise characteristics and well-controlled non-uniformity in the timing performance ensure feasibility of the SPAD sensor towards a broad range of applications. We showcase the experimental results in 2-dimensional (2D) and 3-dimensional (3D) imaging applications based on the megapixel SPAD image sensor.

Finally, we apply the developed megapixel time-gated SPAD image sensor to the light-in-flight imaging. Unprecedentedly high spatial resolution and sub-100-picosecond time resolution enable the observation of an astronomical phenomenon called superluminal motion in the laboratory scale. We propose a new algorithm to reconstruct extra-dimensional information from 3D (x,y,t) spatio-temporal dataset, and successfully present 4-dimensional (x,y,z,t) light-in-flight imaging for the first time.

Key words: Time-of-flight (ToF) ranging, fluorescence lifetime imaging microscopy (FLIM), Raman spectroscopy, light-in-flight imaging, low-light imaging, single-photon avalanche diode (SPAD), megapixel, time gating, guard-ring sharing, charge focusing, backside illumination (BSI), 3D stacking, superluminal motion

Résumé

Les progrès de la technologie optique ont enrichi notre vie quotidienne à bien des égards. Divers types de technologies de détection à photon unique ont vu le jour au cours du siècle dernier. Parmi celles-ci, la technologie de diode à avalanche à photon unique (SPAD) surpasse les autres détecteurs pour son fonctionnement à température ambiante, sa résolution temporelle en picosecondes, sa fabricabilité et son évolutivité. Depuis 2003, lorsque la première matrice SPAD compatible CMOS (complémentaire métal-oxyde complémentaire) a été démontrée, la croissance exponentielle de la technologie de processus CMOS a servi de moteur à la mise à l'échelle et à la miniaturisation des matrices SPAD. En 2017, la plus grande taille de matrice SPAD de 512×512 a été signalée, tandis que le plus petit pas de pixel SPAD de 3 μm a été démontré la même année. Parallèlement à une telle mise à l'échelle continue, les chercheurs ont adopté les capteurs SPAD pour divers domaines d'application. En particulier, les capteurs d'images SPAD sont considérés comme prometteurs pour les applications à résolution temporelle sensibles à un seul photon, telles que la plage de temps de vol (ToF), la microscopie d'imagerie à vie par fluorescence (FLIM), les applications spatiales et la spectroscopie Raman, ainsi que d'autres scientifiques applications d'imagerie. La demande croissante de baies SPAD plus grandes pour améliorer la résolution spatiale, la plage dynamique, la vitesse d'acquisition des données et la fonctionnalité globale est commune à ces applications.

Atteindre un réseau de mégapixels a été l'un des jalons les plus importants de la recherche sur les SPAD depuis plus de 10 ans, mais aussi un qui s'est révélé insaisissable. Pour atteindre cet objectif, nous commençons par une analyse théorique des lois d'échelle pour les caractéristiques de performance SPAD afin de clarifier les relations de compromis sous-jacentes dans les pixels SPAD en dessous de 10 μm pitch. Pour répondre à la dégradation sévère du facteur de remplissage dans la miniaturisation, nous proposons une nouvelle technique de partage d'anneaux de garde, repoussant potentiellement la limite de la miniaturisation des pixels dans les structures SPAD conventionnelles. Le concept proposé est vérifié par une caractérisation approfondie dans des matrices SPAD de test 4×4, où le plus petit pas de pixel au monde de 2,2 μm a été atteint.

Ensuite, nous proposons une nouvelle génération de structures SPAD : SPAD à focalisation de charge. Le concept proposé est prometteur pour surmonter un compromis fondamental entre l'efficacité de détection de photons (PDE) et le taux de comptage sombre (DCR), tout en supprimant le bruit corrélé majeur, la population de pixels chauds et la consommation électrique d'origine SPAD. Un capteur d'image SPAD à focalisation de charge de 128×128 de preuve de concept est développé pour une étude de faisabilité vers des applications d'ima-

Abstract

gerie à faible luminosité, et nous démontrons la plus faible densité DCR du monde de 0,015 cps/μm^2 à température ambiante.

Pour ouvrir la porte à des matrices SPAD évolutives à résolution temporelle, nous proposons de nouveaux concepts de système d'approximation ToF successive et de synchronisation temporelle codée. Les concepts sont mis en œuvre dans des capteurs SPAD à rétroéclairage illuminé (BSI) basés sur un processus CMOS 45 nm/22 nm avec une résolution allant jusqu'à 0,5 mégapixels.

Comme point culminant de cette thèse, nous démontrons les premiers capteurs d'image SPAD 1 mégapixels au monde basés sur une approche de synchronisation temporelle. Une technique de partage de circuits de lecture proposée permet une amélioration significative du facteur de remplissage et une suppression de l'ordre de grandeur de la consommation d'énergie sous un fort éclairage. Les meilleures caractéristiques de bruit de sa catégorie et la non-uniformité bien contrôlée des performances de synchronisation garantissent la faisabilité du capteur SPAD dans une large gamme d'applications. Nous présentons les résultats expérimentaux dans des applications d'imagerie bidimensionnelle (2D) et tridimensionnelle (3D) basées sur le capteur d'image SPAD mégapixels.

Enfin, nous appliquons le capteur d'image SPAD à déclenchement temporel mégapixel développé à l'imagerie de la lumière en vol. Une résolution spatiale sans précédent et une résolution temporelle inférieure à 100 picosecondes permettent d'observer à l'échelle du laboratoire un phénomène astronomique appelé mouvement supraluminique. Nous proposons un nouvel algorithme pour reconstruire des informations extra-dimensionnelles à partir d'un ensemble de données spatio-temporelles 3D (x, y, t), et présentons avec succès l'imagerie en vol en 4 dimensions (x, y, z, t) pour la première fois .

Mots clefs : Gamme de temps de vol (ToF), microscopie d'imagerie à vie par fluorescence (FLIM), spectroscopie Raman, imagerie lumière en vol, imagerie à faible luminosité, diode à avalanche à photon unique (SPAD), mégapixels, déclenchement temporel, anneau de garde partage, mise au point de la charge, éclairage arrière (BSI), empilement 3D, mouvement supraluminique

Contents

Contents

List of Figures

List of Tables

List of Acronyms

AI artificial intelligence

APD avalanche photodiode

BNW buried n-well

BPW buried p-well

BSI backside-illuminated

CCD charge-coupled device

CMOS complementary metal-oxide-semiconductor

cps counts per second

d-ToF direct ToF

DCR dark count rate

DTI deep trench isolation

EMCCD electron multiplying charge-coupled device

FF fill factor

FLIM fluorescence lifetime imaging microscopy

FoV field-of-view

FPGA field-programmable gate array

FRET Förster Resonance Energy Transfer

FSI frontside-illuminated

FWHM full width at half maximum

GMM Gaussian mixture model

List of Tables

GR guard-ring

GS global shutter

i-ToF indirect ToF

IC integrated-circuit

ICCD intensified charge-coupled device

LIBS laser-induced breakdown spectroscopy

LiDAR light detection and ranging

MUX multiplexer

NIR near-infrared

NLOS non-line-of-sight

NW n-well

PDE photon detection efficiency

PDP photon detection probability

PMT photomultiplier tube

PVT process, voltage and temperature

PW p-well

QIS quanta image sensor

QKD quantum key distribution

RS rolling shutter

RTN random telegraph noise

SAR successive approximation

sCMOS scientific CMOS

SEM scanning electron microscope

SMLM super resolution localization microscopy

SPAD single-photon avalanche diode

SRH Schockley-Read-Hall

STI shallow trench isolation

TCSPC time-correlated single-photon counting

TDC time-to-digital converter

ToF time-of-flight

UV ultraviolet

VAPD vertical avalanche photodiode

VR/AR virtual and augmented reality

1 **Introduction**

Progress in optics technology has revolutionized our lives in many ways. Recently, techno-
logical advances have made it possible to produce, detect, and manipulate a single photon:
an elementary particle representing a quantum of light. A major driver for such a technology
is a rapid growth of quantum information science [1], where use of quantum superposition
and entanglement enables quantum computing and ensure ultra-secure communications.
In addition to the quantum information fields, single-photon detector technology has at-
tracted attention of a broad range of fields such as spectroscopy, optical communications,
medical and biomedical imaging, and 2-dimensional (2D) and 3-dimensional (3D) imaging.
Single-photon detection capability has been demonstrated in various systems [2]: photo-
multiplier tubes (PMTs) [3], avalanche photodiodes (APDs), superconductive photodetectors
[4], quantum dot-based devices [5], intrinsic defects in semiconductors [6], and complemen-
tary metal-oxide-semiconductor (CMOS)-based quanta image sensors (QISs) [7]. Among
these, single-photon avalanche diode (SPAD), or Geiger-mode APD, is considered a promising
candidate technology towards compact and cost-effective implementation of single-photon
detector systems. Major advantages of SPADs over other approaches are room temperature op-
eration, picosecond timing resolution, manufacturability, and scalability; since 2003 when the
first CMOS-compatible silicon SPAD array was demonstrated [8], research and development
on SPAD arrays have been accelerated along with exponential scaling of the CMOS process
technology. This chapter presents general perspectives for application fields of time-resolved
SPAD arrays, and feasibility of SPADs for photon counting-based intensity imaging, followed
by history of SPAD pixel scaling and its approaches for miniaturization.

1.1   Applications of time-resolved SPAD arrays

Growing demands in scientific and industrial fields for sensing ultrafast phenomena by optical
means have motivated researchers to develop photodetectors with high temporal resolution
up to nanoseconds or picoseconds. Such an optical sensing system often relies on active
approaches; the sensor detects optical signals synchronous to active light illumination applied
to the target. Some of those applications require full acquisition of the spatial distribution of

physical properties such as structure, depth, material composition, and optical or electrical parameters for the target. A straightforward solution based on a single photodetector is a point scanning. However, this could suffer from a long acquisition time to get a high spatial resolution. An alternative is widefield illumination combined with photodetector arrays. Parallel acquisition of optical signals from different positions of the target could reduce total acquisition time. In addition, elimination of mechanical scanning could simplify the sensing system, which is critical for some specific applications.

Due to defocused light illumination and consequently limited signal-to-noise ratio, the widefield approach necessitates precise detection of weak light signals, often at single-photon level, in each photodetector. A CMOS-based SPAD array is suitable for the widefield time-resolved optical sensing systems owing to its single-photon sensitivity, sub-100-picosecond time resolution and scalability. The digital nature of SPADs enables to combine with compact and high speed logic circuits, including digital counters and time-to-digital converters (TDCs), for pixel-parallel data acquisition and processing.

Fig. 1.1 shows possible application fields of time-resolved SPAD arrays. Time-of-flight (ToF) imaging and light detection and ranging (LiDAR) are among the most promising applications of the time-resolved SPAD arrays, where the time of arrival of laser pulses reflected by objects gives a direct estimation of object distances. In this field, typical SPAD timing jitter below a hundred picoseconds allows sub-centimeter depth resolution in widefield ranging. SPAD
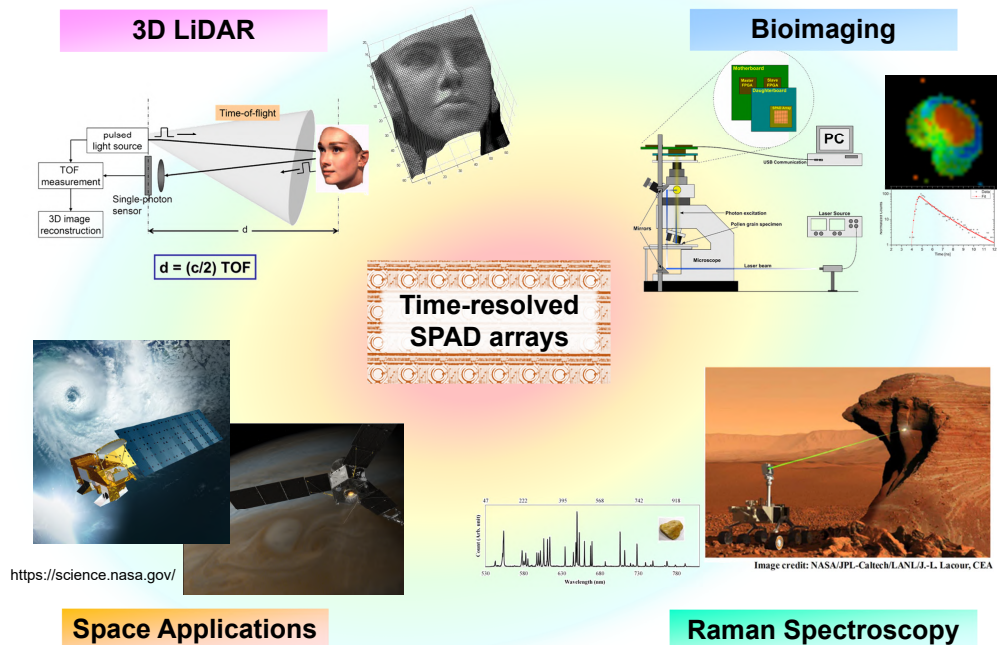
Figure 1.1 – Application fields of time-resolved SPAD arrays.

arrays can also be useful in biomedical imaging, particularly fluorescence lifetime imaging microscopy (FLIM), where nanosecond-scale decay of fluorescent photons needs to be resolved. For those two applications, time-correlated single-photon counting (TCSPC) is often a technique of choice to reconstruct photon time-stamp histogram synchronized with laser pulse illumination. Space applications include imaging and ranging for interplanetary exploration to global scale optical communications. Another attractive direction is the Raman spectroscopy, which is widely used for analyzing the chemical composition and crystal orientation of specimen. In this particular field, background fluorescent photons can be rejected by nanosecond global time-gating schemes. This section introduces previously reported SPAD arrays for each of aforementioned applications and potential motivations for further scaling of the array size.

1.1.1 3D time-of-flight ranging

3D depth sensing plays an essential role in the era of artificial intelligence (AI), autonomous vehicles, robots, virtual and augmented reality (VR/AR), and a variety of other emerging technologies. Depth sensing methods can be classified into passive and active approaches; stereoscopic vision is a commonly used passive approach, where multiple cameras for visible light capture scenes from different angles to reconstruct a depth map based on parallax. Active approaches, in contrast, involve a wave emitter and a detector. The wave can take any forms including acoustic wave and electromagnetic wave. Active depth sensors based on ultrasonic wave are often used for observation of atmosphere and underwater environments. Electromagnetic waves often used in applications of interest in this thesis include visible-to-infrared light, ultraviolet (UV) light, terahertz wave, microwave, X-rays and gamma rays. In addition to its limited interference to human vision, the near-infrared (NIR) light is useful for its abundant emitters, detectors, and optics components such as lenses, mirrors and filters. NIR-based depth sensing comprises structured light, ToF ranging, and some other special techniques.

The ToF ranging performs the depth measurement based on the principle of constant speed of light; object distances can be calculated using time duration from light emission to reception of reflected light at the object surface. The ToF method comprises two classes: indirect ToF (i-ToF) and direct ToF (d-ToF). The former estimates depth from relative phase shift of periodic light emission and detected light signals, while the latter estimates the depth from direct measurement of the time difference from light emission to detection. The i-ToF method can be implemented with simpler pixel circuits and it is scalable. The challenge of i-ToF lies in the impact of ambient light; the maximum measurable range and depth precision are limited by photon-shot noise of the ambient light. D-ToF, in contrast, requires complex pixel circuits, thus potentially limiting the scaling of the array. A major advantage of the d-ToF over the i-ToF is the immunity to ambient light to a greater extent; pulsed light emissions can concentrate the light energy into nanosecond time scale, and time-of-arrival histogramming enables efficient separation of time-correlated signals from time-uncorrelated noise caused by the ambient light.

Fig. 1.2 introduces some of the previously reported SPAD-based ToF sensors. Fig. 1.2(a) is the
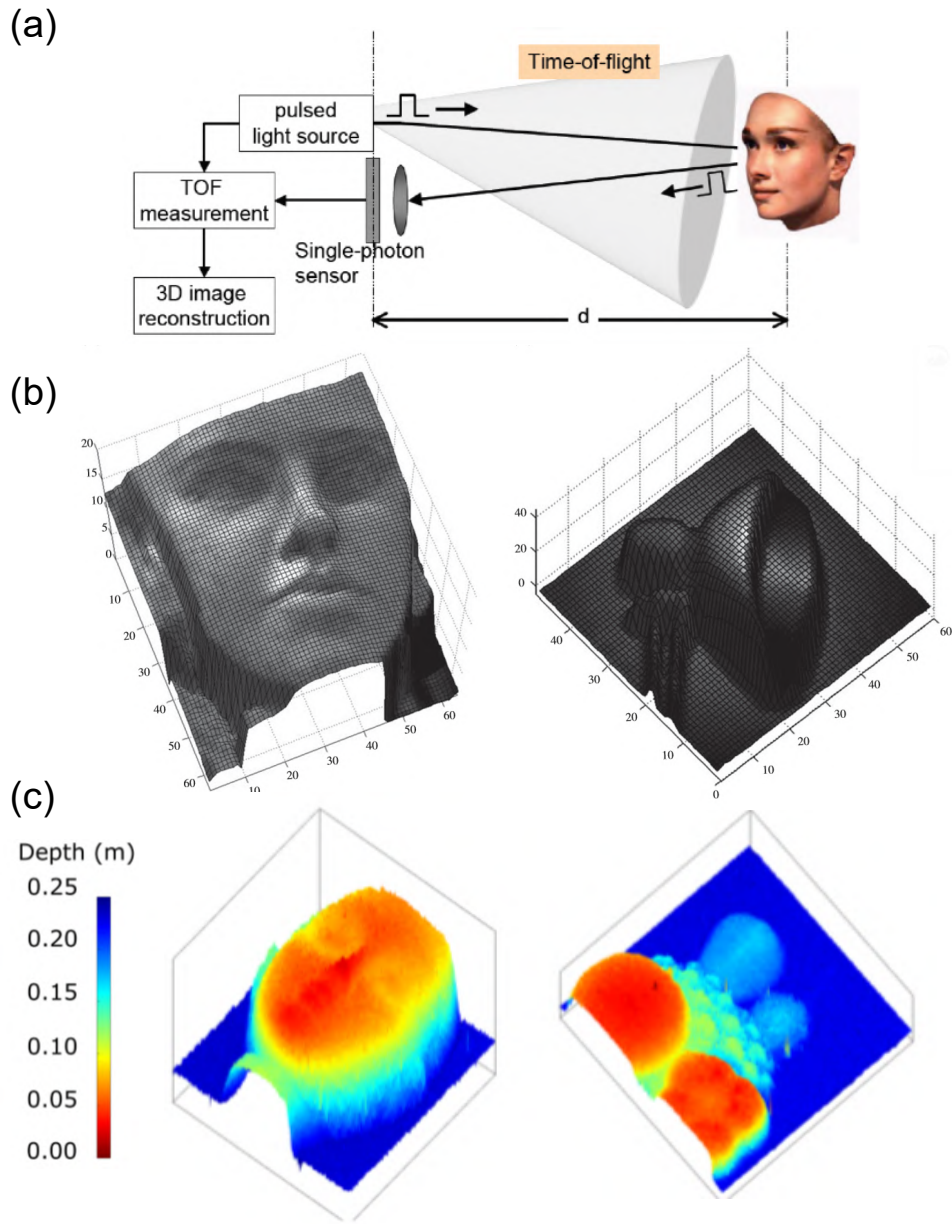
Figure 1.2 – SPAD-based ToF ranging. (a) Conceptual view of ToF ranging system. (b) 3D images captured by TCSPC-based SPAD-ToF imager [9]. (c) 3D images captured by time-gated SPAD-ToF imager [10].

conceptual view of SPAD-ToF camera; pulsed laser is emitted with a wide angle towards an object, and reflected light is collected to SPAD array using optical lens. The laser and the SPAD sensor are synchronized for the d-ToF measurement. Fig. 1.2(b) shows 3D images acquired by a TDC-based 64×64 SPAD image sensor [9]. Here, the 3D images are synthesized with 2D intensity images captured by the same SPAD camera for better visibility. On-chip TDC circuits enabled precise time stamping of single-photon detection events to generate temporal histograms. Fig. 1.2(c) shows 3D color-coded images captured by 256×256 time-gated SPAD sensor [10]. 10 ns time-gating window is scanned in time domain to acquire temporal histograms. Simplified pixel circuit enabled larger pixel arrays, whereas longer acquisition time of the depth map limits the frame rate.

There are several motivations to increase the size of the SPAD array in ToF ranging applications. First, large-scale SPAD arrays enable 3D depth mapping with high spatial resolution, potentially improving detectability of small objects and fine structures in the further distance. Second, scaling of the SPAD arrays can enhance the field-of-view (FoV) of the ToF camera. Third, large number of pixels is desirable to perform on-chip ambient light rejection technique called coincidence detection [11]. In light of these aspects, the large-scale SPAD arrays for ToF ranging are expected to provide a promising detector solution for next-generation flash LiDAR systems.

1.1.2 Biomedical imaging

Advances in optical microscopy have brought a powerful tool for observation of micrometer-to nanometer-scale structures and dynamics in biotechnology fields, as well as microelectronics, mineralogy, nanophysics, etc. In biomedical fields, fluorescence microscopy has enabled researchers to visualize structures and functions of biological tissues by labeling target components with fluorescent probes [12]. Fluorescent molecules, or fluorophores, emit photons typically within sub-nanosecond to microseconds after absorption of photons with shorter wavelength. The wavelength difference of the excitation photons and emitted photons enables selective filtering of background excitation photons from the fluorescent photon signals. In practical situations, some factors limit the illumination power and hence limiting the amount of photon emission; photobleaching is an irreversible process caused by excessive incident light, changing the properties of fluorophores and degrading the efficiency of fluorescence emission permanently. Typical allowable fluorescence cycles in organic fluorophores are 10,000 to 40,000. Phototoxicity is also an irreversible process caused by excessive light interacting with target organisms and cells that are permanently damaged. Those factors impose a requirement of single-photon detection capability on detectors for precise observation of fluorescent signal.

Fluorescence intensity imaging is one of the common approaches to visualize the spatial distribution of fluorophores based on intensity contrast. This approach sometimes has difficulty in resolving finer patterns with low intensity contrast. Alternative approach is FLIM; the distribution of fluorophores is recognized based on flurorescence lifetime, enabling highly resolved imaging for more complex situations, e.g. having multiple different fluorophores in the FoV.
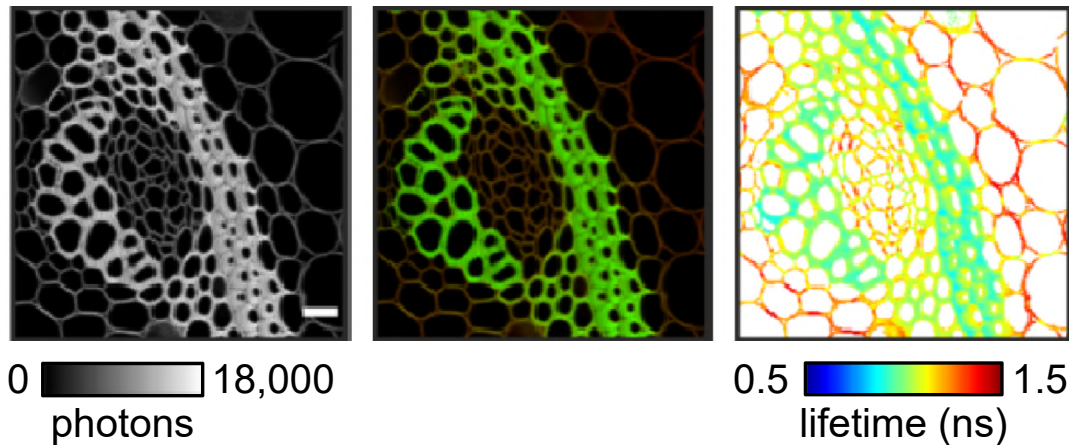
Figure 1.3 – SPAD-based FLIM, comprising of intensity image (left), color image (middle), and lifetime image (right). Images are reprinted from [13].

FLIM requires single-photon detection in picosecond to nanosecond time resolution, and hence the SPAD provides a promising solution. FLIM nowadays has a diverse applications in biomedical imaging thanks to a variety of fluorscence probe molecules developed for different purposes. Förster Resonance Energy Transfer (FRET), where emitted fluorescent photons are absorbed by nearby fluorophore with strong distance dependency of energy transfer, is one of the major applications of FLIM to probe molecular conformation, separation and association in 1 to 10 nm scale [14].

Fig. 1.3 shows previously reported experiments of fluorescence imaging based on SPAD sensor [13]. Left image is captured based on fluorescence intensity imaging, where number of detected photons are plotted in gray scale, whereas image in the center shows color imaging result. Compared to those images, right figure based on FLIM shows richer information on the finer structures with higher visual contrast. Here, a time-gated 64×4 linear SPAD array is implemented for scanning approach. The authors claim that the SPAD-based scanning FLIM demonstrates comparable performance to the conventional PMT-based TCSPC system (not shown in this figure). Later, SPAD arrays are applied to widefield FLIM; the array size has been increased from 128×128 [15] to 160×120 [16], 256×256 [17], and 512×512 [18]. Further scaling of the SPAD array over 1 megapixel for widefield approach enhances spatial resolution and reduces total acquisition time, thus pushing the limit of the existing FLIM systems.

In biomedical imaging field, researchers have investigated effective methods to overcome Abbe's diffraction limit to achieve nanometer-scale spatial resolution. Single molecule super resolution localization microscopy (SMLM) utilizes switching of the on and off states of fluorescence molecules to estimate the position of fluorophores beyond the optical diffraction limit. Fig. 1.4 demonstrates the experimental results of super resolution microscopy based on a 512×128 SPAD array [19]. Fig. 1.4(a) shows a comparison with scientific CMOS (sCMOS) camera, and Fig. 1.4(b) shows a comparison with electron multiplying charge-coupled device
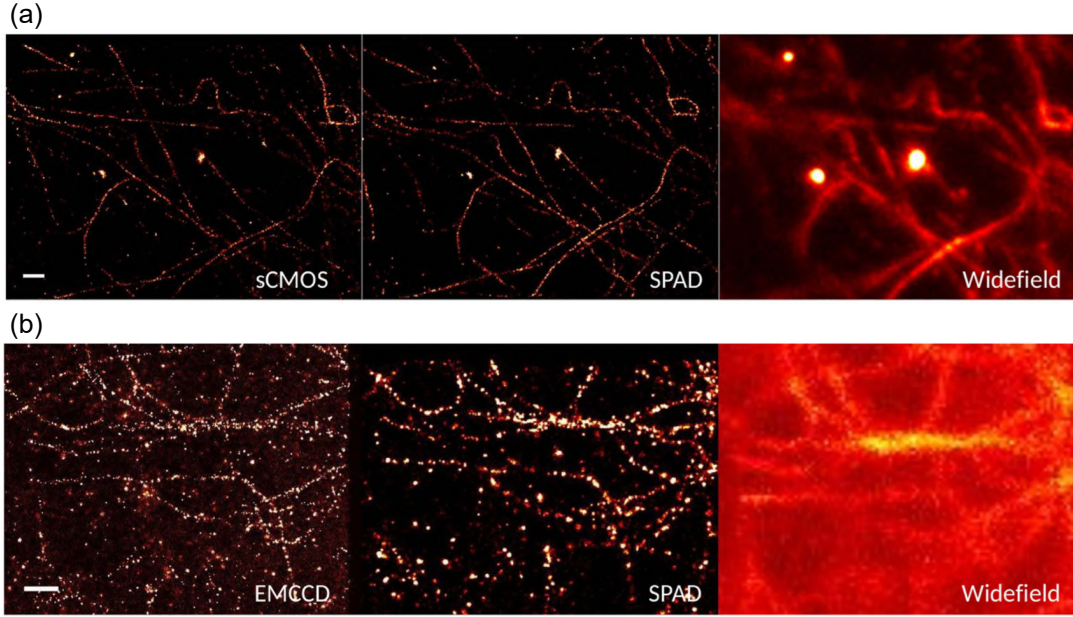
Figure 1.4 – SPAD-based super resolution microscopy [19]. (a) Comparison with sCMOS camera. (b) Comparison with EMCCD camera.

(EMCCD) camera. The authors claim that the SPAD array achieved super resolution performance close to sCMOS and EMCCD camera. The authors also mention that SPAD arrays with reduced pixel size, improved photon detection efficiency (PDE) and higher frame rate are desirable for future implementation of SPAD-based SMLM system competing with the existing technologies.

1.1.3 Space applications

Space exploration and its relevant applications have been strategically investigated worldwide for understanding fundamental physics and astronomy, as well as enriching our lives based on earth observation, satellite-based communication, etc. Space imaging refers to a variety of approaches including imaging astrological phenomena and celestial bodies from the earth, imaging from satellite, and imaging from planetary probes. Those imaging techniques may rely on electromagnetic waves such as UV-to-NIR light, X-ray, and gamma-ray, as well as various types of cosmic rays. In the last decade, SPAD detectors have drawn attention of the space imaging community. Fig. 1.5(a) shows conceptual images of SPAD-based gamma camera for space imaging [20]. The SPAD detector is coupled to low cost plastic scintillating fibers for compact implementation to fit in the limited space in satellites. In this application, the authors claim that large-area SPAD detectors are desirable for efficient collection of scintillating photons. Other requirements are high timing resolution and radiation hardness. For the latter, the authors demonstrate the proton irradiation dose dependence of dark count rate (DCR). The measured DCR increased by a factor of 6 after proton irradiation, which is

(a)

(b)

Figure 1.5 – SPAD-based space applications. (a) Conceptual schematics of SPAD-based gamma ray detectors [20] (right image from https://www.bo.imm.cnr.it/unit/projects/fiber-spad-read-out-scintillating-fibers-single-photon-avalanche-detectors). (b) Conceptual view of global scale quantum communication system [21].

acceptable for this application.

Long distance optical communications between satellites or ground-to-space have been studied intensively for recent years. In particular, global scale telecommunication system based on quantum key distribution (QKD) network is believed to realize a next generation of ultra-secure communication [21]. Fig. 1.5(b) shows a conceptual view of global scale QKD network. Here, efficient generation and detection of entangled polarization states of single photons are major challenges. The authors show the target specification of PDE > 40%, DCR < 1000 counts per second (cps), timing jitter < 100 ps, afterpulsing < 3%, and maximum count rate > 100 kHz. Wavelengths of 800 nm and 1550 nm can potentially be used for this application. Again, radiation hardness of the SPAD detectors is critical for installing on the satellites to ensure the stable operation for a long time. A large-scale SPAD array coupled to an optical lens can be a potential solution to boost the communication speed based on

pixel-parallel signal transmission and detection.

1.1.4 Raman spectroscopy

Spectroscopy is a widely used technique in chemistry, physics and astronomy for investigation of physical structure, composition, and electronic properties of matter from the atomic to macroscopic scale. Among various approaches, Raman spectroscopy is a major tool to nondestructively identify molecules and minerals by quantifying inelastic scattering of photons at the specimen. Weak photon signals in Raman spectroscopy are often overwhelmed by background fluorescence, limiting the signal-to-noise ratio. To address the issue, time-resolved Raman spectroscopy has been studied for temporal filtering of the fluorescent photons [22]. The concept was verified using a streak camera with picosecond time resolution [22] and a single SPAD [23], and later extended to a 128×128 SPAD array with 32 ns time gating [22]. The fluorescence rejection capability in the time-gated SPAD sensor is further refined in 1024×8 linear SPAD arrays achieving 0.7 ns time gating [25]. The authors employed balanced binary trees for spatially uniform gate and recharge clock distribution, and achieved minimum gate length of 700 ps with its skew under 100 ps over the entire array. As shown in Fig. 1.6(a), comparable Willemite Raman signature to the streak camera approach has been demonstrated with the linear SPAD array. The authors claim that the proposed SPAD sensor demonstrated the feasibility not only for Raman spectroscopy but also for laser-induced breakdown spectroscopy (LIBS). Fig. 1.6(b) is an artist's rendering of the LIBS instrument.
An attractive technology direction is a Raman imaging [26], where the system acquires 2D spatial map of the composition of targets. A large-scale 2D SPAD array, combined with 1D or 2D pulse laser illumination, is useful for evaluating 1D spatial distribution of spectral profile in a single measurement. Mechanical scanning of the sensing region enables to generate an image exhibiting the 2D distribution of the matters. In such a system, the SPAD array sizes in horizontal and vertical directions determine the resolution of Raman shift and vertical spatial resolution, whereas scanning pitch determines horizontal spatial resolution. The technical challenge lies in the sensor design to increase the array size without degrading the time-gating length and uniformity over the array. The large-scale time-gated SPAD array with nanosecond time gating is a promising candidate for high speed and high precision sensor towards compact and cost-effective Raman imaging camera.

1.1.5 Other scientific applications

Time-resolved SPAD arrays can also be employed for some of unique imaging modalities. Fig. 1.7(a) shows a demonstration of light-in-flight imaging based on a SPAD array [27]. TDC-based 32×32 SPAD image sensor is used to acquire time stamps of photons scattered from a laser beam. Pixel-wise histogramming combined with dedicated signal processing realizes a visualization of laser beam propagation in the air and reflection at the mirrors. Parallel photon detection significantly reduces measurement time with respect to previous light-in-flight works with a streak camera, intensified charge-coupled devices (ICCDs), etc. The spatial

(a)

(b)
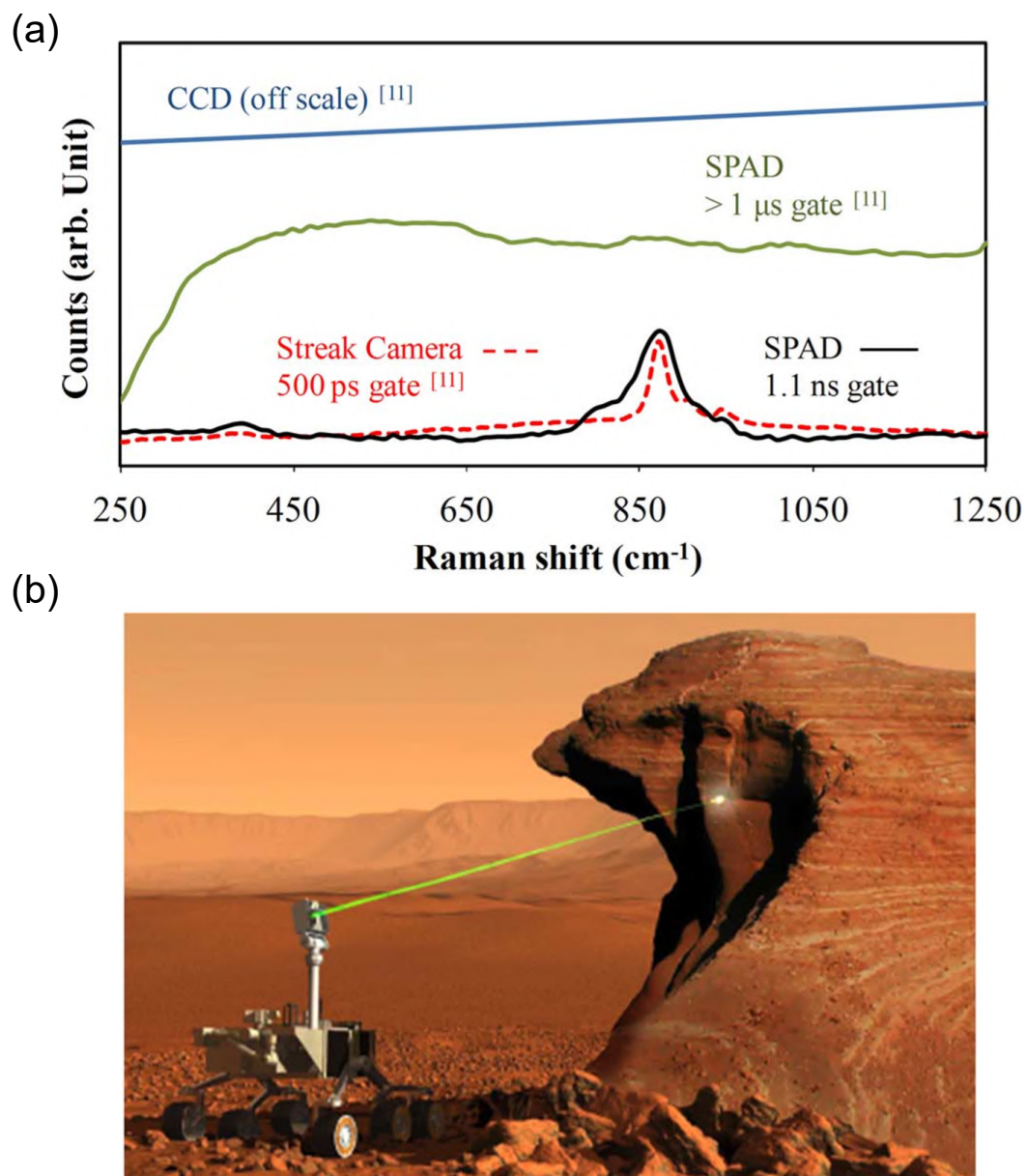
Figure 1.6 – SPAD-based Raman spectroscopy applications. (a) Performance comparison of Raman shift measurement [25]. (b) Artist's rendering of the LIBS instrument [25].

(a)

Integrated image

1.2 ns 1.9 ns 2.6 ns

3.3 ns 4 ns 4.7 ns

Time frames

(b)

Visible wall

Hidden object

Occluder

Scanning galvanometer

Beam splitter

SPAD sensor

Lens

Picosecond laser
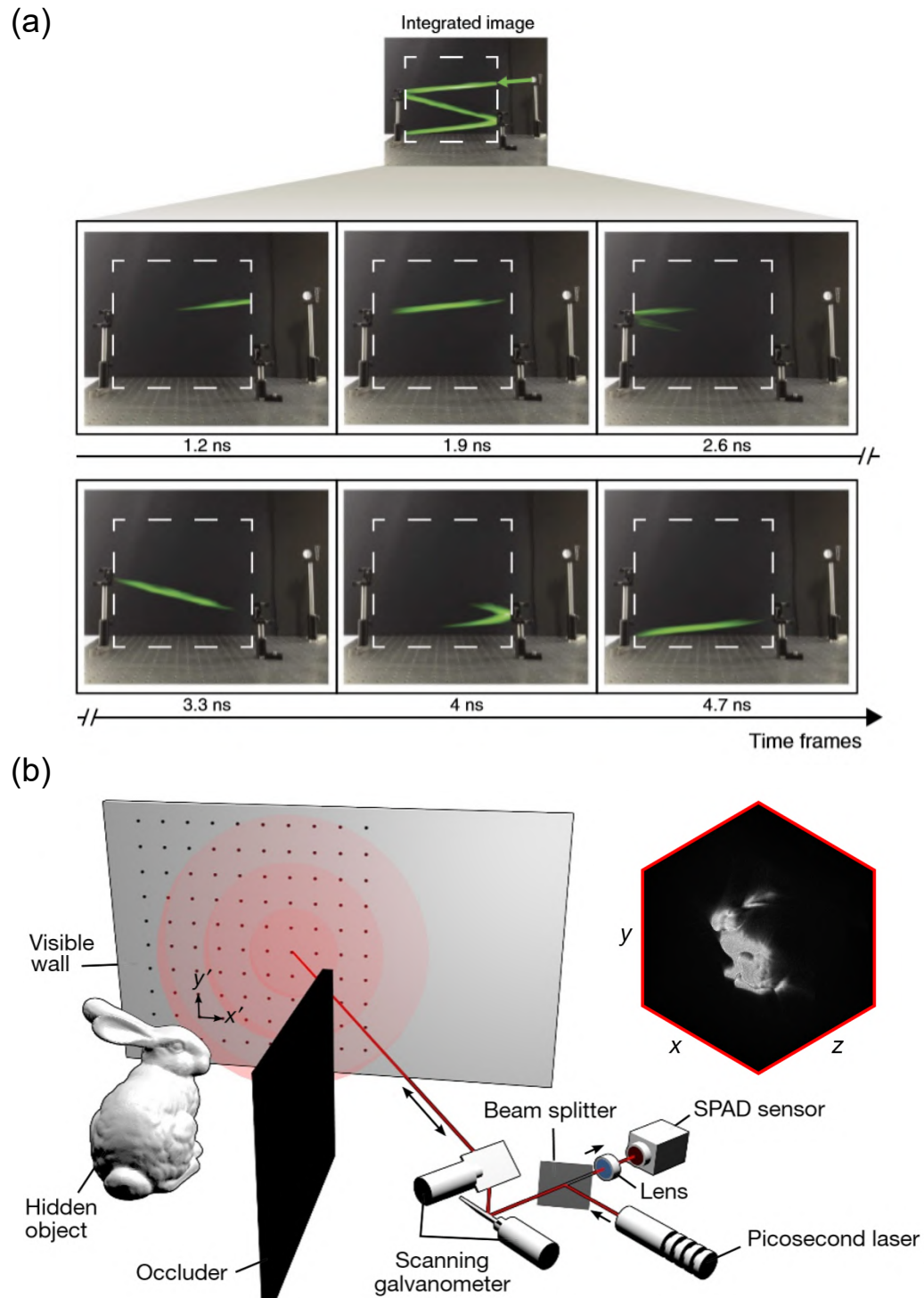
Figure 1.7 – SPAD-based scientific imaging applications. (a) SPAD-based light-in-flight imaging [27]. (b) SPAD-based non-line-of-sight imaging [28].

resolution of the captured images, however, is still limited, and spatial interpolation makes the movement of the laser beam rather clumsy in the captured video. Hence, larger SPAD arrays are promising to potentially improve the accuracy of the light-in-flight observation without interpolation.

Fig. 1.7(b) shows a schematic view of experimental setup for non-line-of-sight (NLOS) and the reconstructed image [28]. The focused laser is emitted towards a wall with diffusive surface. A part of the diffused light propagates towards the hidden target object, and reflected back to the wall. The light is further reflected by the wall and comes back to the imaging systems, eventually reaching the SPAD sensor. Essentially, the imaging system can computationally visualize the shape of the target object placed outside the FoV by utilizing multiple reflection of the pulsed laser. While the paper demonstrates the confocal approach for NLOS imaging, the authors also provide a future perspective for employment of 2-dimensional SPAD arrays with diffused laser source to ultimately achieve high spatial resolution single-shot NLOS imaging. Obviously, more advanced SPAD arrays are desirable for more precise acquisition of photon signals, which is critical for the image reconstruction. Larger arrays could make this reconstruction faster thanks to accelerated acquisition over the whole FoV.

1.2   Quanta image sensors

Apart from time-resolved applications, SPAD arrays enable a photon counting-based intensity imaging for visible-to-NIR light, thus potentially applicable to specialized imaging fields such as surveillance, defense, space imaging, endoscopy, and scientific imaging. In 2005, a new concept of image sensor, nowadays known as QIS, was proposed to overcome the limited imaging performance under aggressive pixel miniaturization [29, 30]. The QIS features spatio-temporal oversampling where each pixel receives only one to few photons in a single frame. The pixel could be designed smaller than a diffraction limit of visible wavelength. Integration of a single- or few-bit oversampled signals over pixels or frames enables to reproduce a nonlinear photoresponse characteristics similar to that of silver halide-based film, thereby achieving extended dynamic range. Applications of the QISs include scientific low-light imaging, professional and consumer photography, multi-aperture imaging, cryptography, detection of low-energy charged particles, and others [7].

Among various technologies, CMOS-based approach provides a scalable solution for QIS. Single-photon resolution in CMOS-based QIS can be achieved by advanced pixel and circuit design for exceptionally high conversion gain and suppressed readout noise. The pixel size can be scaled to 1 μm or less by employing well-established CMOS image sensor (CIS) process technologies. Recently, 1 megapixel QISs with 1.1 μm-pitch pixels are demonstrated in [31, 32]. The developed sensors presented mean temporal noise of 0.21 e-$_{rms}$, binary frame rate up to 1040 fps, and quantum efficiency above 70%.

In theory, SPAD-based QISs, or more generally, photon counting SPAD image sensors have several unique features with respect to the CMOS-based approach. First, in-pixel signal gain with avalanche multiplication process achieves effectively readout-noise-free operation. Sec-
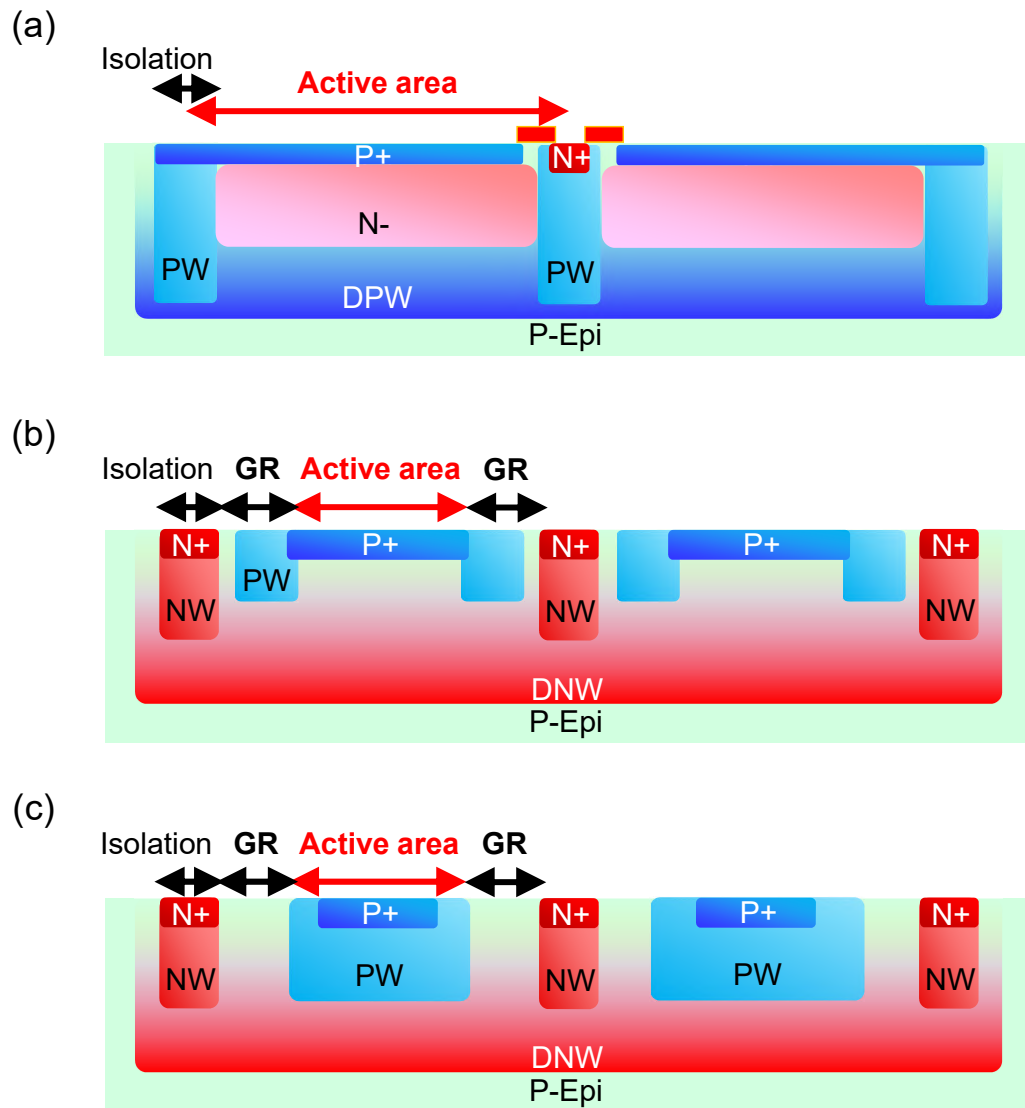
Figure 1.8 – Comparison of device structure for CIS and SPAD. (a) Schematic cross-section of CIS pixels. (b) Schematic cross-section of p+/deep-n-well SPAD. (c) Schematic cross-section of p-well/deep-n-well SPAD.

ond, natively digital operation allows pixel-wise signal processing as well as data compression. Third, global shutter (GS) function can be readily implemented by adding extra memories to each pixel, enabling parasitic-light-insensitive GS operation. SPAD-based QISs based on both analog and digital readout have recently been reported [18, 33].

In spite of the aforementioned advantages, SPAD image sensors are yet to be seen in the 2D camera markets. One of the major issues of SPADs for intensity imaging applications is a low signal-to-noise ratio, fundamentally limited by the high voltage operation. Fig. 1.8(a) shows a schematic cross-section of CIS pixels. The pixel contains a pinned photodiode with a surface P+ region and buried N- photodiode region for charge accumulation, connected to a transistor to transfer the photocharges to a floating diffusion (N+). Each pixel is separated by isolation p-well (PW) to suppress electrical crosstalk and blooming. Owing to the electrostatic potential distribution, the pixel is photosensitive up to the center of the isolation well; photocharges generated in the middle of the isolation region can travel to a nearby pixel by drift and diffusion, thereby enabling near-100% fill factor (indicated as active area in the figure). Fig. 1.8(b) shows a schematic cross-section of p+/deep-n-well SPAD pixels [34]. In the SPAD pixel, the photosensitive area, or active area, is defined by the central p-n junction between P+ and deep n-well (NW) to build uniform high electric field. The active area is surrounded by lightly doped PW, known as a guard-ring (GR), to suppress premature edge breakdown. The pixel is electrically separated from adjacent pixels by NW. Unlike the CIS, photocharges generated around GR and NW isolation are drained by the electrical contacts without reaching the central p-n junction. Hence, the fill factor in the SPAD pixel is significantly limited by the GR and isolations. This is more critical when the pixel size is reduced; widths of GR and isolation are determined by operation voltage and process limitations, and do not necessarily scale with the pixel size. Insufficient GR width induces a premature edge breakdown, degrading key SPAD performance and its uniformity. A similar trend can be expected in a p-well/deep-n-well SPAD shown in Fig. 1.8(c) [35], where the p-n junction between PW and deep NW forms a high electric field region to induce avalanche multiplication. In this device, premature edge breakdown is suppressed by virtual GR with no explicit doping.

In addition to the limited fill factor, SPAD pixels suffer from various noise factors such as DCR, afterpulsing and avalanche-induced crosstalk. Those factors could degrade overall image quality, preventing SPAD cameras from replacing CIS or charge-coupled device (CCD) cameras in the commercial markets. To address those issues in the existing SPADs, innovative concepts of SPAD device need to be introduced to overcome the fundamental limitations and tradeoffs.

1.3 History of SPAD pixel scaling

Analogous to a great success in the scaling of CISs, SPAD pixel-pitch scaling towards megapixel array plays a key role for further expansion of its potentials toward the aforementioned industrial and scientific applications. Fig. 1.9 explains the history of scaling in CMOS-compatible SPAD arrays. Starting from 32 pixels in 2003, the size of pixel arrays in SPAD sensors has continuously boosted by downsizing SPAD and pixel circuit with advanced CMOS technologies. The
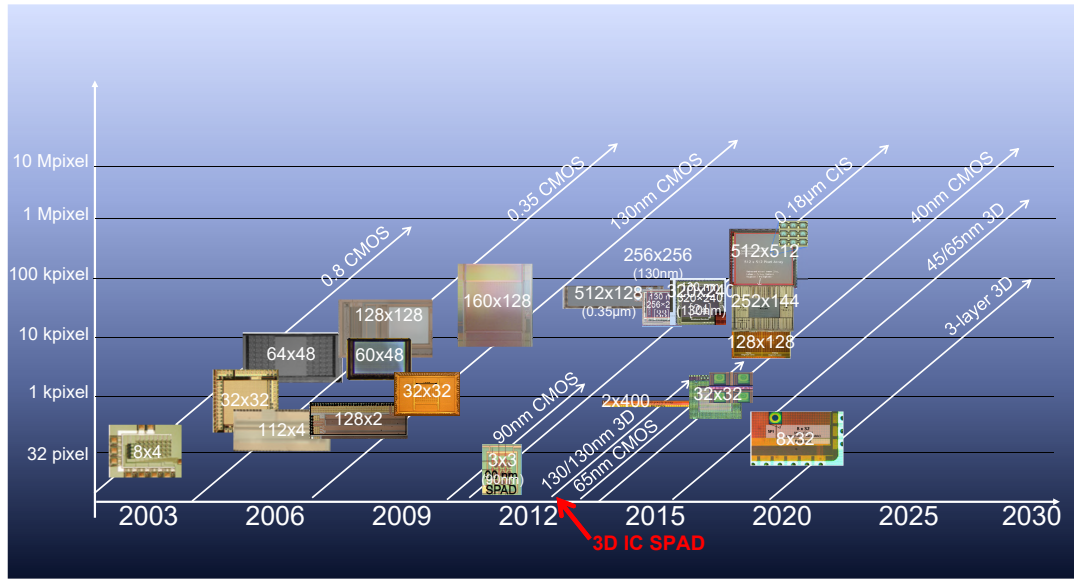
Figure 1.9 – History of CMOS-compatible SPAD development.

record of spatial resolution has been updated in 2017 by time-gated 512×512 SPAD sensor [36] (as of 2019). Aiming at high resolution widefield FLIM, the authors developed 16.38 μm-pitch frontside-illuminated (FSI) SPAD in 180 nm CMOS process. To achieved high photon detection probability (PDP) and low DCR, p-i-n diode-based SPAD is employed [37]. The pixel circuit consists of nMOS-only 11 transistors used for compact and low-power time-gating circuit and 1-bit dynamic memory.

In the context of pixel-pitch scaling, possibility of miniaturizing SPAD array has been explored in [38], where 4×4 of test pixels comprise 3 μm-pitch SPADs with all the readout circuit located external to the array. This work has focused on minimizing distance between neighboring SPADs in FSI, while performance degradation in DCR and PDP implies that aggressively miniaturized SPADs suffer from premature edge breakdown.

1.4 Approaches to miniaturize SPAD arrays

Various approaches have been proposed to address the issues. Fig. 1.10(a) shows a conceptual schematic of on-chip microlenses [39, 40, 41]. In monolithic configuration, SPAD and pixel circuit components share limited area. Due to the high voltage operation, the well region of the SPAD needs to be located distant from wells for the pixel circuit to avoid unwanted high electric field concentration. This requirement further reduces the fill factor. On-chip microlenses help collecting incident photons to the pixel area into the SPAD active area, improving the effective fill factor. With optimized design, the concentration factors of microlenses above 10 have been reported [39, 40], whereas the gain is limited for smaller f-numbers of an objective lens. A secondary effect of on-chip microlens is the reduction of DCR; the DCR often follows
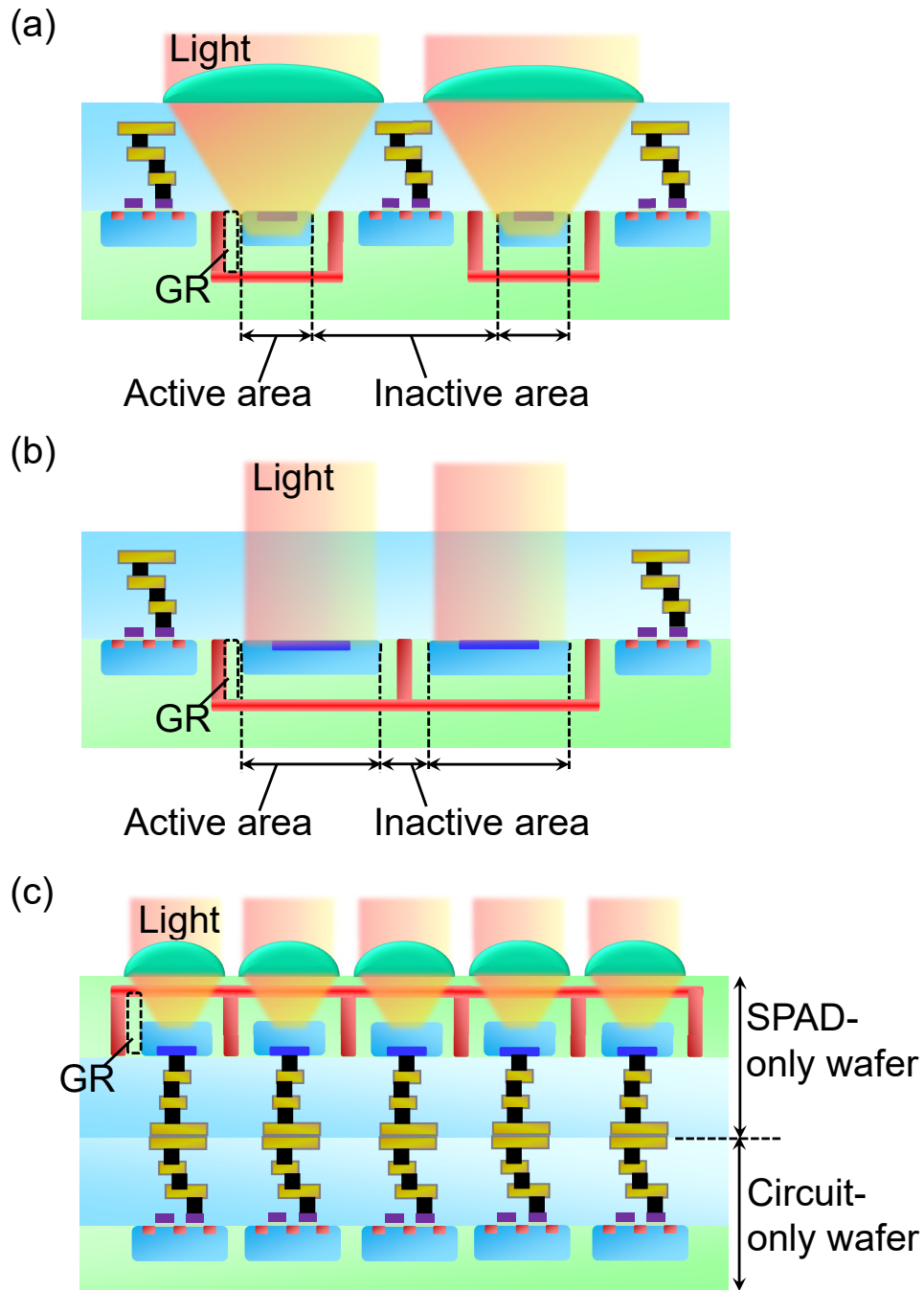
Figure 1.10 – Main approaches to miniaturize SPAD arrays. (a) On-chip microlens implementation [39, 40, 41]. (b) Well-shared SPAD arrays [42]. (c) 3D-stacked SPAD arrays combined with other techniques [44, 45, 46, 47].

proportionally to the active area, and the on-chip microlenses enable reduction of the active area while avoiding loss of the effective fill factor and PDE.

Fig. 1.10(b) shows a schematic cross-section of a well-shared SPAD array [45]. Adjacent SPADs are sharing an isolation well (shown in red) to reduce the distance between the active area. In contrast to Fig. 1.10(a), the pixel circuit can share the well region as well for area efficient layout. Note that this approach could lose the regularity of pixel array, thus distorting uniformity of spatial sampling of the photon signal.

Recently, 3D stacking has been adopted for commercial CISs and now the technology enables pixel-level wafer-to-wafer bonding with pitch below 5 μm [43]. The 3D-integration technology has been introduced for SPADs [44, 45, 46, 47], and has proven to overcome fundamental limitations in the monolithically integrated SPAD arrays. Fig. 1.10(c) shows a schematic view of a 3D-stacked SPAD array combined with other miniaturization techniques. The sensor consists of two wafers: a SPAD-dedicated wafer and a circuit-dedicated wafer. SPADs can be densely packed without special care on the electrical breakdown with circuit wells, thus significantly enhancing the fill factor. Each SPAD is bonded to the pixel circuit in the circuit-dedicated wafer. Physical separation of SPADs and circuits enable to employ different process technologies for the two wafers. The SPAD wafer can be optimized for high PDE and low noise performance, while advanced CMOS processes (e.g. 65 nm, 45 nm, or below) for the circuit wafer achieve compact, high speed and highly functional pixels.

1.5 Aims of the thesis

This thesis is motivated by the potential demands in diverse scientific and industrial application fields for compact and cost-effective megapixel time-resolved SPAD arrays. As described in the previous sections, time-resolved SPAD arrays have been adopted for various applications, some of which have shown outstanding performance and unique functionality, whereas others have shown performance not fully competitive to the established technologies such as PMT, CIS, EMCCD, sCMOS camera, streak camera, etc. Major limitations of the existing SPAD image sensors reside in the array size, pixel size, fill factor, PDP, DCR, and timing jitter, as well as correlated noise such as afterpulsing and crosstalk.

The first objective of this thesis is to formulate the scaling laws of SPAD pixels and exemplify fundamental tradeoffs in the miniaturized SPAD arrays below 10 μm pitch. As in any other semiconductor devices, miniaturization of the SPAD pixel is an everlasting challenge for researchers and engineers to achieve compact and inexpensive implementation of single-photon sensing systems. Some of the aforementioned characteristics improve as the pixel size is reduced, whereas the others are severely degraded due to the miniaturization. Systematic analysis of the tradeoff relations is of great importance to understand the physical limit of the miniaturization, and to investigate the possibility to push the limit of performance in the existing architectures.

The second objective of this thesis is to propose and implement novel concepts to overcome the tradeoffs in miniaturized SPADs based on insights from semiconductor device physics, fab-

rication process technology, and integrated-circuit (IC) design methodology. The concepts can be verified by theoretical analysis, simulations, and experiments. In addition to the sensitivity and noise characteristics, controlling stochastic variation and pixel-to-pixel non-uniformity of the performance is critical for robust manufacturing of single-photon detector arrays. The state-of-the-art comparison provides a useful measure to systematically clarify the advantages and limitations of the proposed concepts. Those concepts are ultimately targeting the demonstration of the world's first megapixel SPAD image sensors with excellent performance.

The third objective is to employ the large-scale SPAD arrays towards various time-resolved applications. System design and algorithm development are key for demonstrating the feasibility of the SPAD image sensors in the target applications. In addition to the continuous performance improvement, the developed large-scale SPAD image sensors are expected to achieve discontinuous innovations which have never been approached by the existing detectors. Such innovations can pave the way for completely new application fields requiring even higher level of SPAD array design, potentially inspiring the further research and development towards the multi-megapixel time-resolved SPAD image sensors.

1.6   Thesis organization

Chapter 2 describes a theoretical analysis of scaling laws for SPAD pixels to clarify the fundamental tradeoffs between various performance. In Chapter 3, a novel guard-ring-shared SPAD device technology is proposed to push the limit of the pixel miniaturization. The proposed concept is verified with 4×4 SPAD arrays achieving 2.2 μm pitch, the smallest ever reported. Chapter 4 explains a newly proposed charge focusing SPAD device technology to overcome a critical tradeoff between PDE and DCR, enabling miniaturized SPAD arrays with near-100% fill factor and ultralow DCR. A proof-of-concept 128×128 charge focusing SPAD array is developed to demonstrate the feasibility of the charge focusing SPAD for low-light imaging applications. In Chapter 5, new approaches for highly functional time-gated SPAD arrays are proposed. The algorithms are implemented in backside-illuminated (BSI) 3D-stacked SPAD arrays with the resolution up to 0.5 megapixel. Chapter 6 demonstrate the first megapixel SPAD array, fabricated in 180 nm CMOS. The megapixel time-gated SPAD image sensor based on new pixel architecture is applied to 2D and 3D imaging. The developed megapixel SPAD camera is further utilized for light-in-flight imaging in Chapter 7. Owing to high spatio-temporal resolution and a newly developed reconstruction algorithm, 4-dimensional light-in-flight observation is demonstrated for the first time. The conclusion and future perspectives are summarized in Chapter 8.

References

[1] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited Review Article: Single-photon sources and detectors," Rev. Sci. Instrum. 82(071101) (2011).

[2] R. H. Hadfield, "Single-photon detectors for optical quantum information applications," Nat. Photonics 3, 696-705 (2009).

[3] G. A. Morton, " Photomultipliers for scintillation counting," RCA Rev. 10, 525–553 (1949).

[4] A. Divochiy, F. Marsili, D. Bitauld, A. Gaggero, R. Leoni, F. Mattioli, A. Korneev, V. Seleznev, N. Kaurova, O. Minaeva, G. Gol'tsman, K. G. Lagoudakis, M. Benkhaoul, F. Lévy, and A. Fiore, " Superconducting nanowire photon-number-resolving detector at telecommunication wavelengths," Nat. Photonics 2, 302-306 (2008).

[5] A. J. Shields, and M. P. O'Sullivan, "Detection of single photons using a field-effect transistor gated by a layer of quantum dots," Appl. Phys. Lett. 76, 3673-3675 (2000).

[6] H. Kosaka, D. S. Rao, H. D. Robinson, P. Bandaru, T. Sakamoto, and E. Yablonovitch, "Photoconductance quantization in a single-photon detector," Phys. Rev. B 65(201307) (2002).

[7] E. R. Fossum, J. Ma, S. Masoodian, L. Anzagira, and R. Zizza, "The quanta image sensor: every photon counts," Sensors 16, 1260 (2016).

[8] A. Rochas, M. Gösch, A. Serov, P. A. Besse, R. S. Popovic, T. Lasser, and R. Rigler, "First Fully Integrated 2-D Array of Single-Photon Detectors in Standard CMOS Technology," IEEE Photonics Technology Lett. 15(7), 963-965 (2003).

[9] E. Charbon, "Single-photon imaging in complementary metal oxide semiconductor processes," Phil. Trans. R. Soc. A 372 (2014).

[10] X. Ren, P. W. R. Connolly, A. Halimi, Y. Altmann, S. McLaughlin, I. Gyongy, R. K. Henderson, and G. S. Buller, "High-resolution depth profiling using a range-gated CMOS SPAD quanta image sensor," Opt. Express 26(5), 5541-5557 (2018).

[11] C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, "A 100-m Range 10-Frame/s 340×96-Pixel Time-of-Flight Depth Sensor in 0.18μm CMOS," IEEE J. Solid-State Circuits 48(2), 559-572 (2013).

References

[12] J. W. Lichtman, and J.-A. Conchello, "Fluorescence microscopy," Nat. Methods 2, 910-919 (2005).

[13] M. Popleteeva, K. T. Haas, D. Stoppa, L. Pancheri, L. Gasparini, C. F. Kaminski, L. D. Cassidy, A. R. Venkitaraman, and A. Esposito, "Fast and simple spectral FLIM for biochemical and medical imaging," Opt. Express 23(18), 23511-23525 (2015).

[14] E. A. Jares-Erijman, and T. M. Jovin, "FRET imaging," Nat. Biotechnology 21, 1387-1395 (2003).

[15] Y. Maruyama, and E. Charbon, "An all-digital, time-gated 128X128 spad array for on-chip, filter-less fluorescence detection," IEEE Int. Solid-State Sensors, Actuators and Microsystems Conference (2011).

[16] M. Perenzoni, N. Massari, D. Perenzoni, L. Gasparini, and D. Stoppa, "A 160×120-Pixel Analog-Counting Single-Photon Imager with Time-Gating and Self-Referenced Column-Parallel A/D Conversion for Fluorescence Lifetime Imaging," IEEE J. Solid-State Circuits 51(1), 155-167 (2016).

[17] I. Gyongy, N. Calder, A. Davies, N. A. W. Dutton, R. R. Duncan, C. Rickman, P. Dalgamo, and R. K. Henderson, "A 256×256, 100-kfps, 61% Fill-Factor SPAD Image Sensor for Time-Resolved Microscopy Applications," IEEE Trans. Electron Devices 65(2), 547-554 (2018).

[18] A. C. Ulku, C. Bruschini, I. M. Antolovic, Y. Kuo, R. Ankri, S. Weiss, X. Michalet, and E. Charbon, "A 512×512 SPAD Image Sensor with Integrated Gating for Widefield FLIM," IEEE J. Selected Topics in Quantum Electronics 25(1), 1-12 (2018).

[19] I. M. Antolovic, S. Burri, C. Bruschini, R. A. Hoebe, and E. Charbon, "SPAD imagers for super resolution localization microscopy enable analysis of fast fluorophore blinking," Sci. Rep. 7(44108) (2017).

[20] M. Marisaldi, P. Maccagnani, F. Moscatelli, C. Labanti, F. Fuschino, M. Prest, A. Berra, D. Bolognini, M. Ghioni, I. Rech, A. Gulinatti, A. Giudice, G. Simmerle, D. Rubini, A. Candelori, and S. Mattiazzo, "Single Photon Avalanche Diodes for space applications," IEEE Nuclear Science Symposium Conference (2011).

[21] E. Kerstel, A. Gardelein, M. Barthelemy, The CSUG Team, M. Fink, S. K. Joshi, and R. Ursin, " Nanobob: a CubeSat mission concept for quantum communication experiments in an uplink configuration," EPJ Quantum Technology 5(6) (2018).

[22] J. Blacksberg, G. R. Rossman, and A. Gleckler, "Time-resolved Raman spectroscopy for in situ planetary mineralogy," Appl. Optics 49(26), 4951-4962 (2010).

[23] I. Nissinen, J. Nissinen, A.-K. Länsman, L. Hallman, A. Kilpelä, and J. Kostamovaara, "A Sub-ns Time-gated CMOS Single Photon Avalanche Diode Detector for Raman Spectroscopy," Proc. the European Solid-State Device Research Conference (2011).

[24] J. Blacksberg, Y. Maruyama, E. Charbon, and G. R. Rossman, "Fast single-photon avalanche diode arrays for laser Raman spectroscopy," Opt. Lett. 36(18), 3672-3674 (2011).

[25] Y. Maruyama, J. Blacksberg, and E. Charbon, "A 1024×8, 700-ps Time-Gated SPAD Line Sensor for Planetary Surface Exploration With Laser Raman Spectroscopy and LIBS," IEEE J. Solid-State Circuits 49(1), 179–189 (2014).

[26] S. Sil, R. Gautam, and S. Umapathy, "Applications of Raman and Infrared Microscopy to Materials and Biology," Molecular and Laser Spectroscopy, Advances and Applications, 117-146 (2018).

[27] G. Gariepy, N. Krstajic, R. Henderson, C. Li, R. R. Thomson, G. S. Buller, B. Heshmat, R. Raskar, J. Leach, and D. Faccio, "Single-photon sensitive light-in-flight imaging," Nat. Commun. 6(6021) (2015).

[28] M. O'Toole, D.B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on light-cone transform," Nature 555(7696), 338 (2018).

[29] E. R. Fossum, "Gigapixel digital film sensor (DFS) proposal," in Nanospace Manipulation of Photons and Electrons for Nanovision Systems, The 7th Takayanagi Kenjiro Memorial Symposium and the 2nd International Symposium on Nanovision Science, Hamamatsu, Japan (2005).

[30] E. R. Fossum, "Some Thoughts on Future Digital Still Cameras," in Image Sensors and Signal Processing for Digital Still Cameras, CRC, 305–314 (2006).

[31] J. Ma, S. Masoodian, D. A. Starkey, and E. R. Fossum, "Photon-number-resolving megapixel image sensor at room temperature without avalanche gain," Optica, 4(12), 1474-1481 (2017).

[32] A. Gnanasambandam, O. Elgendy, J. Ma, and S. H. Chan, "Megapixel photon-counting color imaging using quanta image sensor," Opt. Express 27(12), 17298 (2019).

[33] N. A. W. Dutton, I. Gyongy, L. Parmesan, S. Gnecchi, N. Calder, B. R. Rae, S. Pellegrini, L. A. Grant, and R. K. Henderson, "A SPAD-Based QVGA Image Sensor for Single-Photon Counting and Quanta Imaging," IEEE Trans. Electron Devices 63(1), 189-196 (2016).

[34] C. Niclass, A. Rochas, P.-A. Besse, and E. Charbon, "Design and Characterization of a CMOS 3-D Image Sensor Based on Single Photon Avalanche Diodes," J. Solid-State Circuits 40(9), 1847-1854 (2005).

[35] J. A. Richardson, E. A. G. Webster, L. A. Grant, and R. K. Henderson, "Scaleable Single-Photon Avalanche Diode Structures in Nanometer CMOS Technology," IEEE Trans. Electron Devices 58(7), 2028-2035 (2011).

[36] A. C. Ulku, C. Bruschini, X. Michalet, S. Weiss, and E. Charbon, "A 512×512 SPAD Image Sensor with Built-In Gating for Phasor Based Real-Time siFLIM," Int. Image Sensor Workshop (2017).

References

[37] C. Veerappan, and E. Charbon, "A Low Dark Count p-i-n Diode Based SPAD in CMOS Technology," IEEE Trans. Electron Devices 63(1), 65-71 (2015).

[38] Z. You, L. Parmesan, S. Pellegrini, and R. Henderson, "3μm Pitch, 1μm Active Diameter SPAD Arrays in 130nm CMOS Imaging Technology," Int. Image Sensor Workshop (2017).

[39] J. M. Pavia, M. Wolf, and E. Charbon, "Measurement and modeling of microlenses fabricated on single-photon avalanche diode arrays for fill factor recovery," Opt. Express 22(4), 4202-4213 (2014).

[40] I. M. Antolovic, A.C. Ulku, E. Kizilkan, S. Lindner, F. Zanella, R. Ferrini, M. Schnieper, E. Charbon, and C. Bruschini, "Optical-stack optimization for improved SPAD photon detection efficiency," in Proc. SPIE, Quantum Sensing and Nano Electronics and Photonics XVI, San Francisco, CA, USA (2019).

[41] I. Gyongy, A. Davies, B. Gallinet, N. A. W. Dutton, R. Duncan, C. Rickman, R. K. Henderson, and P. A. Dalgarno, "Cylindrical microlensing for enhanced collection efficiency of small pixel SPAD arrays in single-molecule localisation microscopy," Opt. Express 26(3), 2280-2291 (2018).

[42] T. Al Abbas, N. Dutton, O. Almer, F. M. D. Rocca, S. Pellegrini, B. R. Rae, D Golanski, and R. K. Henderson, "8.25μm Pitch 66% Fill Factor Global Shared Well SPAD Image Sensor in 40nm CMOS FSI Technology," Int. Image Sensor Workshop (2017).

[43] Y. Kagawa, N. Fujii, K. Aoyagi, Y. Kobayashi, S. Nishi, N. Todaka, S. Takeshita, J. Taura, H. Takahashi, Y. Nishimura, K. Tatani, M. Kawamura, H. Nakayama, T. Nagano, K. Ohno, H. Iwamoto, S. Kadomura, and T. Hirayama, "Novel Stacked CMOS Image Sensor with Advanced Cu2Cu Hybrid Bonding," IEEE Int. Electron Devices Meeting (2016).

[44] J. M. Pavia, M. Scandini, S. Lindner, M. Wolf, and E. Charbon, "A 1 × 400 Backside-Illuminated SPAD Sensor With 49.7 ps Resolution, 30 pJ/Sample TDCs Fabricated in 3D CMOS Technology for Near-Infrared Optical Tomography," IEEE J. Solid-State Circuits 50(10), 2406-2418 (2015).

[45] T. Al Abbas, N. A. W. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. K. Henderson, "Backside illuminated SPAD image sensor with 7.83 μm pitch in 3D-stacked CMOS technology," IEEE Int. Electron Devices Meeting, 811-814 (2016).

[46] A. R. Ximenes, P. Padmanabhan, M.-J. Lee, Y. Yamashita, D.-N. Yaung, and E. Charbon, "A Modular, Direct Time-of-Flight Depth Sensor in 45/65-nm 3-D-Stacked CMOS Technology," IEEE J. Solid-State Circuits 54(11), 3203-3214 (2019).

[47] R. K. Henderson, N. Johnston, S. W. Hutchings, I. Gyongy, T. Al Abbas, N. Dutton, M. Tyler, S. Chan, and J. Leach, "A 256×256 40nm/90nm CMOS 3D-Stacked 120dB Dynamic-Range Reconfigurable Time-Resolved SPAD Imager," IEEE Int. Solid-State Circuits Conference (2019).

2 Scaling Laws in SPAD Pixels

SPAD pixel scaling leads to the shrinkage of its active area and affects the performance measures, such as fill factor, PDP, PDE, DCR, correlated noises, and power consumption. To understand the impact of pixel miniaturization, it is useful to introduce simplified models for each parameter which take the dimension of the pixel into consideration. In this chapter, a theoretical study on the scaling laws of the pixel performance is investigated.

2.1 Analysis criteria

To proceed with the theoretical analysis of the scaling laws, some assumptions must be made. First, SPAD pixel array configuration is assumed to be a square grid, while it is not difficult to generalize the discussion to other configurations, e.g. honeycomb structure [1]. Second, circular-shaped SPADs are assumed to simplify the discussion on the curvature change with scaling. In some prior works, rounded-corner rectangle or square SPADs are also adopted to improve fill factor [2, 3, 4]. However, these designs are not always suitable for scaling with geometric similarity preserved, where the electric field concentration at the corners can induce premature edge breakdown and also change the breakdown voltage. Third, a 3D-stacked configuration with SPAD-only array in a single plane is assumed. In non-3D-stacked FSI or BSI configuration, SPAD and pixel circuit coexist in the same plane. In a given pixel pitch, SPAD and circuit have to share the limited area, and the circuit complexity can affect the size of SPAD active area and its performance. The main focus of this analysis is to formulate the scaling laws of SPAD performance, and hence the SPAD array without circuit components is desired for more systematic and quantitative analysis. Fourth, active-to-active distance is assumed to be fixed at a certain dimension irrespective of the scaling parameter. This is justified by the following discussion.

For analysis of the scaling laws in the SPAD pixel, it is natural to assume that the doping profile along z-axis for each implantation layer is unchanged, and the breakdown voltage of the p-n junction in the SPAD stays in the same range. This implies that, unlike the scaling in MOS transistors where the lower supply voltage is adopted for the smaller devices, the power supply voltage for the SPAD does not scale as a function of its dimension. Another premise

on the SPAD pixel design is that the guard-ring width has to be sufficiently large to avoid the premature edge breakdown. Given the fact that the lateral diffusion length of doped ions cannot readily be controlled, the electrostatic potential distribution around the guard ring is not much dependent on the active diameter. The optimum guard-ring width ensuring no edge breakdown in the operating condition is defined with the following equation [5]:

$$V_B^{gr}(W_{gr}) = V_B^{p-n} + V_{ex}^{max}, \tag{2.1}$$

where $V_B^{gr}(W_{gr})$ is the breakdown voltage at the guard ring with given guard-ring width W_{gr}, V_B^{p-n} the breakdown voltage at the vertical p-n junction, and V_{ex}^{max} the maximum excess bias used in the system. Based on the discussion above, all the terms in the above equation are not dependent on the pixel size, and the optimum W_{gr} can be defined regardless of scaling. These considerations impose a constraint in the pixel scaling that the guard-ring width has to be unscaled and fixed at a certain value over all the SPAD pixel dimensions to guarantee the stable Geiger-mode operation without unwanted edge breakdown. The optimum W_{gr} should be in the similar range as the depletion width of the main SPAD p-n junction, and is typically 1 to 2 μm [6]. In addition, the optimum width of an isolation layer, typically formed with deep-well implantation, is determined by a process design rule for minimum drawing width, and should not be scaled with the pixel dimension. The pixel pitch L_p, which will be employed as a scaling parameter in the following discussion, can be expressed as:

$$L_p = D_a + L_{a-a} = D_a + 2 \cdot W_{gr} + W_{iso}, \tag{2.2}$$

where the well-sharing configuration is assumed, and D_a is the active diameter, L_{a-a} the active-to-active distance, and W_{iso} the isolation width. In the following discussion, W_{gr} and W_{iso} are both assumed to be 1 μm unless otherwise noted, and L_p is assumed solely dependent on D_a.

Fig. 2.1 shows the conceptual views depicting the SPAD pixel scaling. Fig. 2.1(a) is the example of top-view layout for a 2×2 pixel array. As discussed above, the active-to-active distance L_{a-a} is fixed when shrinking the pixel pitch L_p. As a result, the active diameter D_a is reduced proportionally to L_p. This assumption can be applied to any types of existing SPAD device structures [7, 8]. For example, Fig. 2.1(b) shows the cross-sectional view of p+/NW SPAD. D_a is defined as the diameter of the inner circle of guard-ring p-well, whereas L_{a-a} corresponds to a sum of the NW separation width and twice the width of p-well guard ring. For PW/deep-NW SPAD or p-i-n SPAD, D_a equals to the diameter of p-well, and L_{a-a} is a sum of NW separation width and twice the width of virtual p-epi guard ring. This indicates that the scaling law analysis can be performed with only three key dimensional parameters, L_p, D_a and L_{a-a}, without losing generality.

In summary, the main assumptions for the analysis of scaling laws are listed below:
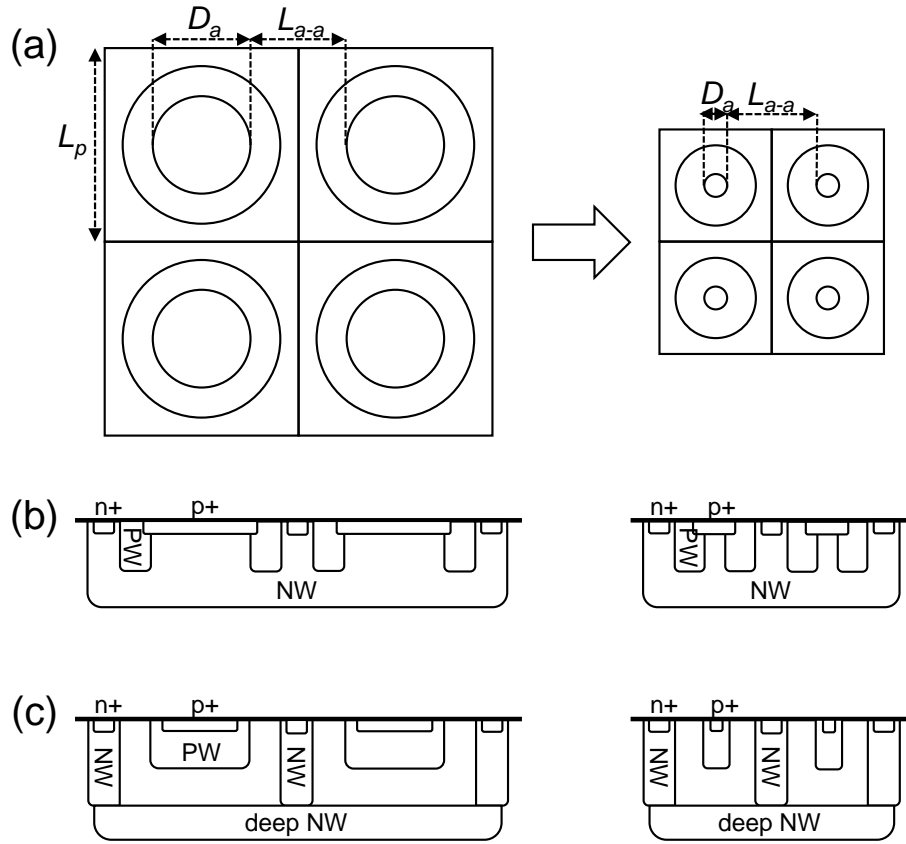
Figure 2.1 – Conceptual views of SPAD pixel scaling; (a) top-view layout examples of pixel miniaturization, (b) cross-section example of p+/NW SPAD, (c) cross-section example of PW/deep-NW SPAD or p-i-n SPAD.

- uniform square grid,

- circular shape for active area and inner/outer borders of guard ring,

- 3D-stacked configuration with full separation of SPAD and pixel circuit into different wafers,

- active-to-active distance unscaled with the SPAD pixel dimension,

- pixel pitch L_p employed as a scaling parameter.

2.2 Fill factor

Fill factor (FF) of the SPAD pixel, defined as the ratio between the drawn active area and the pixel area, is one of the fundamental parameters determining single-photon sensitivity. FF is a

purely geometric parameter, and is straightforward to be formulated as a function of the pixel pitch L_p:

$$FF = \frac{\pi(L_p - L_{a-a})^2}{4L_p^2}.$$

(2.3)

It is obvious from the above equation that FF goes down to zero when $L_p = L_{a-a}$, and cannot be defined for $L_p < L_{a-a}$. For sufficiently large L_p, FF converges to $\pi/4 \times 100 = 78.5\,\%$.

Fig. 2.2 shows the calculated FF as a function of the pixel pitch for several different active-to-active distances. FF curves show monotonic increase with the pixel pitch L_p. Relatively steep increase of FF is observed at smaller L_p, whereas saturating behavior of FF is shown at larger L_p. Slower saturation for larger L_{a-a} indicates that if active-to-active distance is large, larger pixel pitch is required to get higher FF, e.g. above 50%. Note that in the actual sensor design, the effective FF can be enhanced by employing on-chip microlenses [9, 10], although designers should bear in mind that microlenses are less effective for smaller f-numbers of the objective lens.
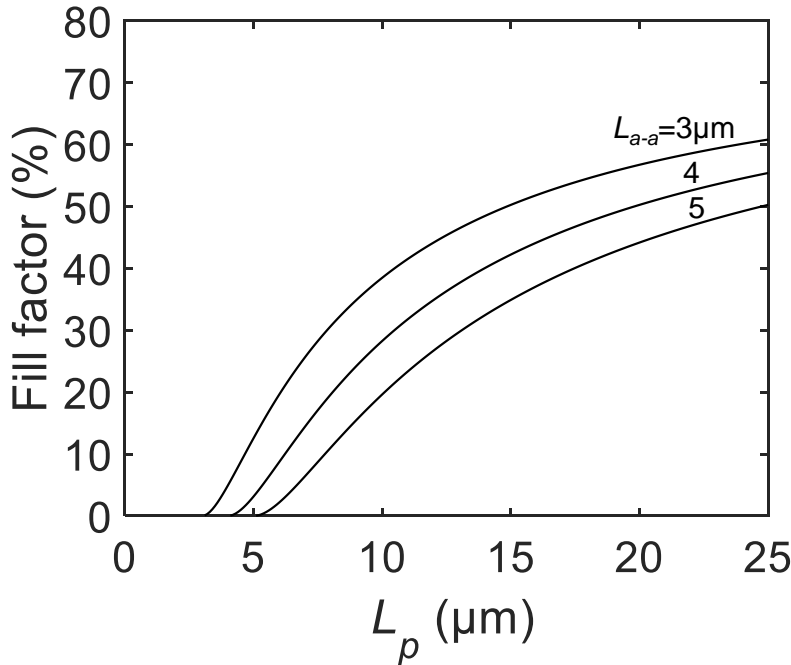
Figure 2.2 – Calculated FF as a function of SPAD pixel pitch L_p for active-to-active distances $L_{a-a} = 3, 4, 5\ \mu$m.

2.3 PDP and PDE

PDP in the SPAD pixels is defined in the following equation [11]:

$$PDP = QE \times P_{ava},$$ (2.4)

where QE is the quantum efficiency and P_{ava} the avalanche triggering probability. In ideal SPAD device, PDP represents the single-photon sensitivity normalized by the active area, and it is not scaled with the active diameter and the pixel pitch. In practice, discrepancy between "drawn" active area and "effective" active area leads to considerable dependencies of PDP with the scaling parameter L_p [12].

The discrepancy between designed and actual active size stems from two possible reasons: nonideality in the process fabrication, and nonideality in the device design. One example of the process nonideality is the lateral diffusion of doped ions [13]. The lateral diffusion length is determined by type of dopant ion, implantation energy, and thermal annealing conditions, and is typically in the order of 0.1 to 1 μm for deep well implantation. This lateral diffusion induces the decrease of doping concentration at the edge of the active area. The electric field at the edge of active area can be locally reduced with respect to the electric field at the center of the active area, thus lowering the sensitivity at the border of the active area. On the other hand, the device design nonideality is caused by lateral electric field near the guard ring. Photocharges generated in the neutral region of the SPAD randomly move around due to the thermal diffusion until they reach the nearby depletion region and are drifted to an electrode. If the photocharges reach the main p-n junction with high electric field, they induce the avalanche multiplication, thereby generating the photon detection signal. However, photocharges close to the border of active area can reach the depletion region towards the guard ring before reaching the main junction. In such a case, the carriers don't cause avalanche multiplication, and no photon detection signal is observed. This so-called "border effect" [14, 15] causes the photon detection loss at the edge of active area, which becomes more significant in the smaller pixels.

For both process- and device-originated nonidealities, PDP correction can be performed by introducing inactive radius r_{in}, representing the effective width of photon-insensitive region at the edge of active region. The corrected equation for the scaling law of PDP is given by:

$$PDP = PDP_{max} \times (\frac{L_p - L_{a-a} - 2r_{in}}{L_p - L_{a-a}})^2,$$ (2.5)

where PDP_{max} is the virtual maximum PDP with sufficiently large active size.

Fig. 2.3 shows the calculated PDP as a function of pixel pitch for different r_{in}. The curve with $r_{in} = 0$ μm corresponding to the ideal case with no border effect shows no dependency with L_p. For finite r_{in}, PDP starts from zero at $L_p = L_{a-a} + 2r_{in}$, grows and saturates to PDP_{max} with increasing L_p. Similar to the scaling law for FF, slower increase is observed for the larger
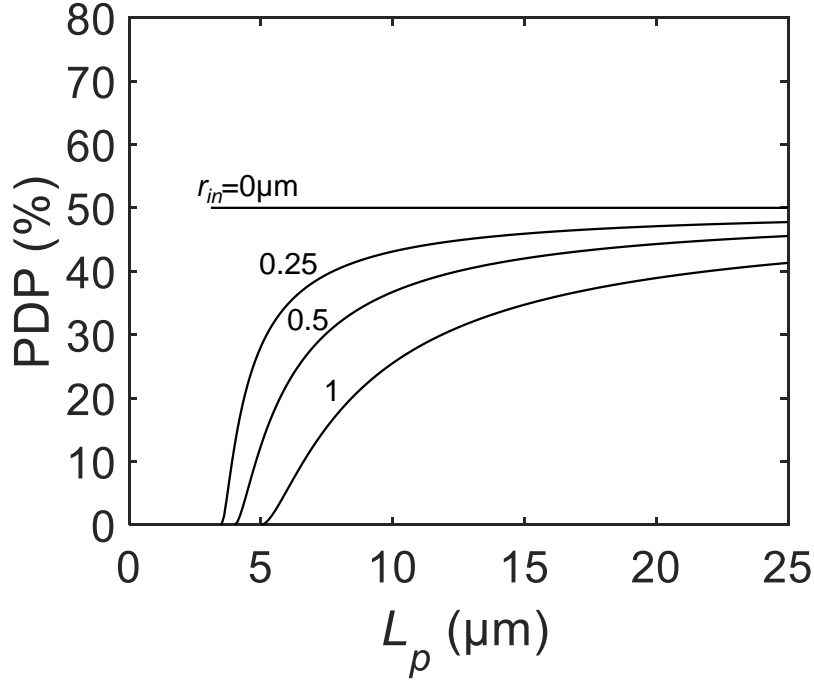
Figure 2.3 – Calculated PDP as a function of SPAD pixel pitch L_p for $PDP_{max} = 50\%$, active-to-active distance $L_{a-a} = 3$ μm and inactive radius $r_{in} = 0, 0.25, 0.5, 1$ μm.

r_{in}.

PDE is another indicator of single-photon sensitivity. Unlike PDP where the sensitivity is normalized by the active area, PDE is defined as the single-photon sensitivity normalized by the pixel area. The following equation holds [11]:

$$PDE = PDP \times FF. \tag{2.6}$$

Based on the previous equations, PDE can be explicitly formulated as:

$$PDE = PDP_{max} \times \frac{\pi(L_p - L_{a-a} - 2r_{in})^2}{4L_p^2}. \tag{2.7}$$

Fig. 2.4 is the calculated PDE as a function of L_p for different r_{in}. Similar to FF and PDP, the curves start from zero at smaller L_p, and saturate at larger L_p. The maximum PDE is given by $PDP_{max} \times 78.5\% = 39.3\%$, assuming $PDP_{max} = 50\%$. Again, introducing on-chip microlenses will potentially increase overall PDE.
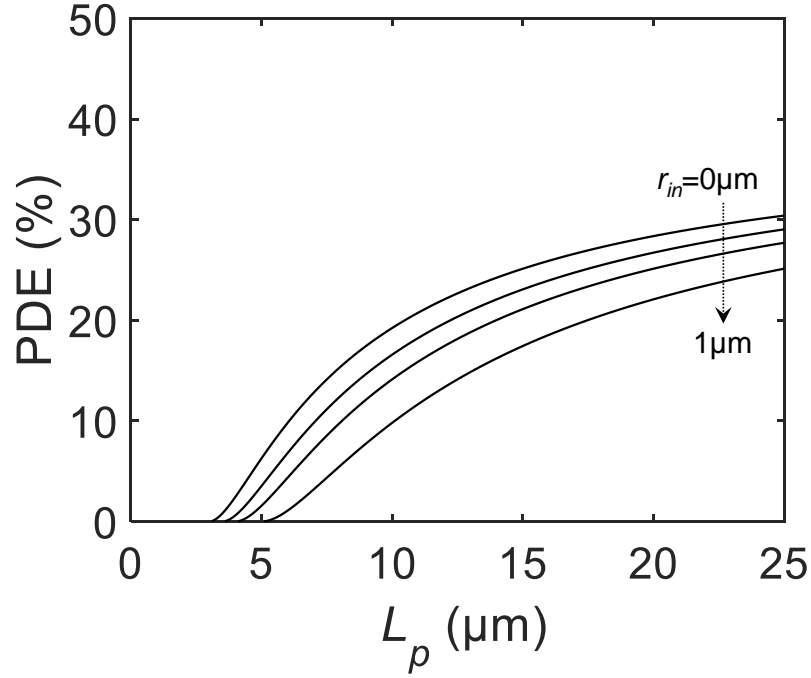
Figure 2.4 – Calculated PDE as a function of SPAD pixel pitch L_p for PDP_{max} = 50%, active-to-active distance L_{a-a} = 3 μm and inactive radius r_{in} = 0, 0.25, 0.5, 1 μm.

2.4 DCR

DCR has several different causes, such as band-to-band tunneling, trap-assisted tunneling, trap-assisted thermal generation and diffusion current [16, 17]. Experimentally, the source of the DCR can be classified based on an Arrhenius plot [18, 19, 20]. In silicon SPADs, the activation energies E_a for band-to-band tunneling, trap-assisted tunneling, trap-assisted thermal generation and diffusion current are known to be approx. 0 eV, 0-0.55 eV, 0.55 eV, and 1.1 eV, respectively. In practice, the measured E_a can be intermediate values, e.g. 0.8 eV, indicating the mixture of multiple DCR components.

Based on the assumption that the premature edge breakdown is suppressed, the tunneling components at the edge of active region can be neglected. Contributions of the thermal generation and diffusion current are also negligible in the depletion region to the guard ring due to insufficient electric field for avalanche triggering by the generated carriers. Therefore, the contribution from the main p-n junction of the SPAD dominates over that from the edge of the active region. Interestingly, all the aforementioned DCR components are proportional to the "effective" active area. Tunneling current, regardless of band-to-band or trap-assisted, is proportional to the total volume of region with highly concentrated electric field, which is obviously proportional to the active area. Thermal generation and diffusion carriers are detected only when those carriers are generated in the vicinity of the active region. Assuming that the thermal generation and diffusion current are spatially uniform around the active region, those components are also naturally assumed to be proportional to the active area.
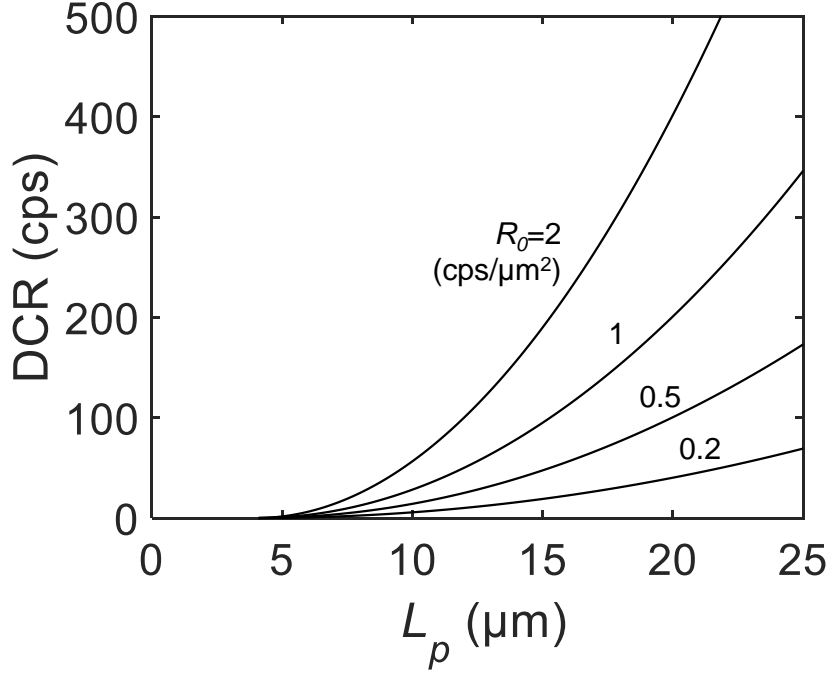
Figure 2.5 – Calculated DCR as a function of SPAD pixel pitch L_p for R_0 = 0.2, 0.5, 1 and 2 cps/μm^2, active-to-active distance L_{a-a} = 3 μm and inactive radius r_{in} = 0.5 μm.

The scaling law for DCR can be formulated as follows:

$$DCR = R_0 \times \frac{\pi(L_p - L_{a-a} - 2r_{in})^2}{4},$$

(2.8)

where R_0 is a DCR per unit active area.

Fig. 2.5 is the calculated DCR as a function of L_p for different DCR per unit area R_0. Starting from 0 cps at $L_p = L_{a-a}+2r_{in}$, DCR shows parabolic increase with L_p. DCR is highly dependent on R_0, which is a function of excess bias, temperature, and process quality such as trap and impurity densities. Opposite to FF, PDP and PDE, smaller pixel pitch is desirable to improve the DCR performance. The designer should consider the best tradeoff between PDE and DCR to find the optimum L_p to provide reasonable S/N ratio.

DCR density R is defined by DCR normalized by drawn active area, and is often used for comparison of the SPAD process quality between devices fabricated in different processes [21]. Just as PDP, nonideality such as the border effect leads to the dependence of R on L_p as follows:

$$R = R_0 \times \frac{\pi(L_p - L_{a-a} - 2r_{in})^2}{4(L_p - L_{a-a})^2}.$$

(2.9)
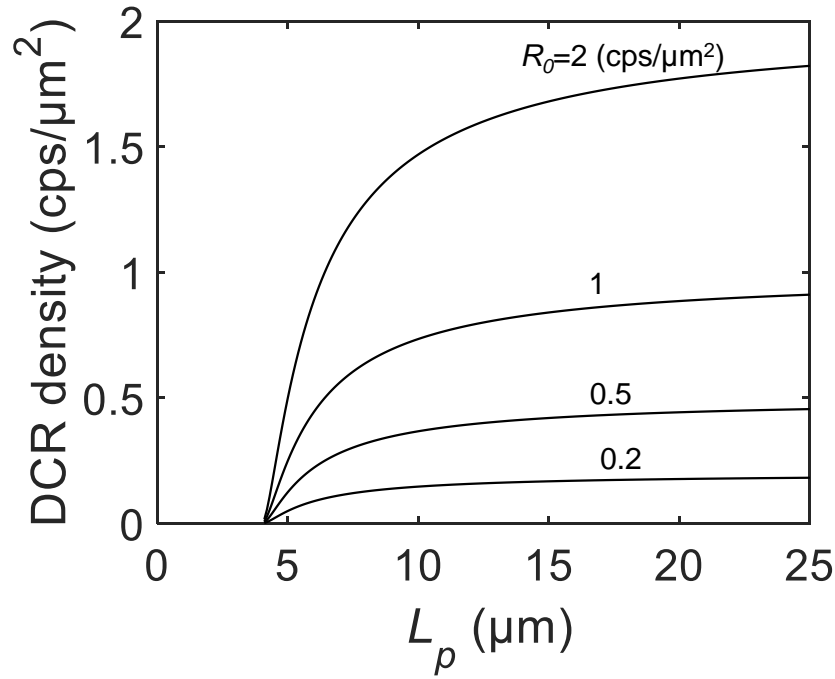
Figure 2.6 – Calculated DCR density as a function of SPAD pixel pitch L_p for $R_0 = 0.2, 0.5, 1, 2$ cps/μm^2, active-to-active distance $L_{a-a} = 3$ μm and inactive radius $r_{in} = 0.5$ μm.

At larger L_p, the DCR density saturates to R_0. Fig. 2.6 shows the L_p dependence of the DCR density for various R_0. As can be seen from the similarity to the equation for PDP, the DCR density starts from zero at $L_p = L_{a-a} + 2r_{in}$, rapidly increases and saturates for larger L_p. This implies that in the actual measurement, the DCR density can be underestimated at the smaller pixel pitch due to the existence of the photon-insensitive region at the edge of the active region.

Note that the above discussion is based on the assumption that the guard-ring width is optimized to avoid edge breakdown for entire range of the pixel pitch. In the actual device design, sometimes abrupt increase of DCR and DCR density is observed at smaller pixel pitch, even with fixed active-to-active distance. To the author's knowledge, no systematic analysis has been conducted for this phenomenon. One possible reason is the enhanced curvature at the edge of active region inducing the electric field concentration near guard ring. Analogous to antennas, electric field tends to be concentrated in regions of high curvature, which may enhance the corner electric field and induce premature edge breakdown when scaling down the pixel. Another possible explanation is the nonideality in the photoresist formation process. In most of the SPAD devices, the diffusion regions for p-n junction, guard ring or isolation are formed by well doping, where high energy doping is employed. In such a process, thicker photoresist is desired to avoid penetration of accelerated ions through the resist. The opening size of the photoresist for such a thick resist (typically 3 to 10 μm) requires careful calibration to match the actual shape and size to the designed layout. The layout for well doping is usually supported only for 0 or 90 degree lines, whereas SPAD layout often involves circular or ring

shape having arbitrary angles. This could give the deviation of the actual resist opening size from the design especially in the smaller pixel dimension, leading to the unwanted edge breakdown.

2.5 Afterpulsing

Correlated noise, such as afterpulsing and crosstalk, is critical for some applications, where temporal and spatial correlations of photon detection signals play key roles [22, 23, 24]. Afterpulsing is caused by an avalanche-generated carrier captured at deep trap state near the multiplication region, which is released by thermal activation or tunneling after nanosecond to microsecond trapping time and induces another avalanche multiplication event. This mechanism implies that the afterpulsing probability P_a is dependent on the trap density D_{trap} and the total number of avalanche-generated carriers N_{ava}. Higher trap density and more avalanche carriers result in higher P_a. If P_a is not too large, e.g. smaller than 10%, a linear relation between P_a and $D_{trap} \times N_{ava}$ can be assumed in a first-order approximation [25].

Assuming the spatially uniform distribution of the deep trap states, D_{trap} is independent of the scaling parameter. N_{ava}, on the other hand, can be dependent on the scaling parameter. N_{ava} is calculated based on the following:

$$eN_{ava} = C_{par} \times V_{ex} = (C_{p-n} + C_0) \times V_{ex}, \tag{2.10}$$

where e is the elementary charge, V_{ex} the excess bias, C_{par} the total parasitic capacitance at the SPAD output node, either cathode or anode which is connected to the quenching resistor, C_{p-n} the p-n junction capacitance at active region, C_0 the sum of the other parasitic capacitance contributions from connected metal wires, diffusion regions, gates, etc. C_{p-n} is proportional to the active area, whereas C_0 does not scale with the pixel size or the active size. In summary, the scaling law of P_a is given by:

$$P_a = A \times \left[\frac{\pi \epsilon (L_p - L_{a-a})^2}{4W_{eff}} + C_0 \right], \tag{2.11}$$

where A is the temperature-, bias- and process-dependent coefficient, ϵ the permittivity, and W_{eff} the effective depletion region width determined by the p-n junction doping profile.

Fig. 2.7 shows the L_p dependence of afterpulsing probability for various W_{eff} and C_0 (dashed lines for $C_0 = 5$ fF, and solid lines for $C_0 = 30$ fF). For all the parameter combinations, the parabolic increase of P_a is shown with the offset corresponding to $A \times C_0$. Larger W_{eff} shows weaker dependence of P_a on L_p, indicating less contribution of p-n junction capacitance to the total parasitic capacitance C_{par}. In any case, scaling down of the pixel has a positive impact on the afterpulsing probability thanks to the reduced parasitic capacitance.

Figure 2.7 – Calculated afterpulsing probability as a function of SPAD pixel pitch L_p for $C_0 = 5$ fF (dashed lines) and 30 fF (solid lines), active-to-active distance $L_{a-a} = 3\ \mu$m, $A = 1 \times 10^{11}$ F^{-1}, and $W_{eff} = 0.5, 1, 2\ \mu$m.

Note that the dead time is assumed constant for all L_p in this analysis. In a real device design, fixed quenching resistance results in the L_p dependence of the dead time. This secondary effect makes the P_a less sensitive to L_p compared to the case where constant dead time is assumed. If the dependence of P_a on dead time is strong enough to compensate the trend as shown in Fig. 2.7, then, it will be possible to flatten or even reverse P_a for larger L_p.

2.6 Crosstalk

Crosstalk is another type of correlated noise in SPAD pixels. Unlike afterpulsing, where only a single pixel is involved, crosstalk involves two pixels in the noise generation process. When avalanche multiplication is triggered in a pixel, thousands to millions of electrons and holes are generated. When those carriers are recombined with counterpart charges, either photons or phonons can be emitted to preserve the energy conservation law. Silicon is a material with indirect bandgap, and hence the probability to emit photons is very low. For the photon energy higher than the silicon bandgap, only several to tens of photons are emitted out of one million of avalanche-generated carriers [26, 27]. Still, those photons can move towards a neighboring pixel and be detected.

Similar to afterpulsing, crosstalk probability P_c is dependent on the number of avalanche-generated carriers N_{ava}. A larger number of carriers leads to the higher P_c. Again, in a first-

order approximation, P_c is considered to be proportional to N_{ava}. In addition, the distance between pixels is another important factor for scaling. Given that the emitted secondary photons decay exponentially with travel length, shorter pixel-to-pixel distance results in higher crosstalk. The emitter-to-receiver distance dependence of crosstalk can be approximated by [28]:

$$P_c = B\frac{e^{-\alpha r}}{r^2},\tag{2.12}$$

where B is a coefficient that will be explained later, r the distance from one SPAD of interest to the other, and α the effective decay length of the emitted light. Regarding the crosstalk between two nearest-neighbor SPAD pixels, r in the above equation corresponds to the pixel pitch L_p. Note that this equation implicitly assumes that the light emission occurs at the center of the active region for the emitting SPAD, and the average photon intensity reaching the active region of receiver is approximated by the photon intensity at the center of the active region of the receiving SPAD. In reality, the finite size of the active region for both emitter and receiver may cause slight deviation of measured crosstalk from the above model. For simplicity, the following analysis will be based on the above model where effect of finite active size is neglected.

The coefficient B is dependent on both emitter and receiver characteristics. Considering the emitter, B should depend on the total number of emitted photons which is proportional to N_{ava}. On the other hand, B should also be correlated with the sensitivity of the receiver. The probability of detecting an emitted photon is proportional to PDP and the effective active area, which coincide with PDE by definition. Thus, the crosstalk probability between two nearest-neighbor SPAD pixels can be expressed as:

$$P_c = B' \times \left[\frac{\pi\epsilon(L_p - L_{a-a})^2}{4W_{eff}} + C_0 \right] \times \frac{e^{-\alpha L_p}}{L_p^2} \times \frac{(L_p - L_{a-a} - 2r_{in})^2}{L_p^2},\tag{2.13}$$

where B' is an excess-bias dependent coefficient.

Fig. 2.8 shows the L_p dependence of the calculated crosstalk probability for various α and C_0. All curves show the peak close to $L_p = L_{a-a} + 2r_{in}$, followed by decreasing trend for larger L_p. The decreasing rate is highly dependent on the combination of parameters: the curve with $\alpha = 0.2\ \mu\mathrm{m}^{-1}$ and $C_0 = 30$ fF shows sharp reduction towards zero, whereas the curve with $\alpha = 0.05$ $\mu\mathrm{m}^{-1}$ and $C_0 = 5$ fF shows mild decrease. Note that $\alpha = 0.2\ \mu\mathrm{m}^{-1}$ and $0.05\ \mu\mathrm{m}^{-1}$ correspond to the cases with effective light emission wavelengths of 700 nm and 850 nm, respectively. Except for L_p very close to $L_{a-a} + 2r_{in}$ where PDE drops to zero, the scaling of the SPAD pixel has an adverse effect on the crosstalk probability. This is in contrast with afterpulsing, where smaller L_p is preferable for suppressing noise.

To deal with the degraded crosstalk at smaller pixel pitch, several countermeasures can be considered. First, lowering V_{ex} helps suppressing the crosstalk probability at the expense of
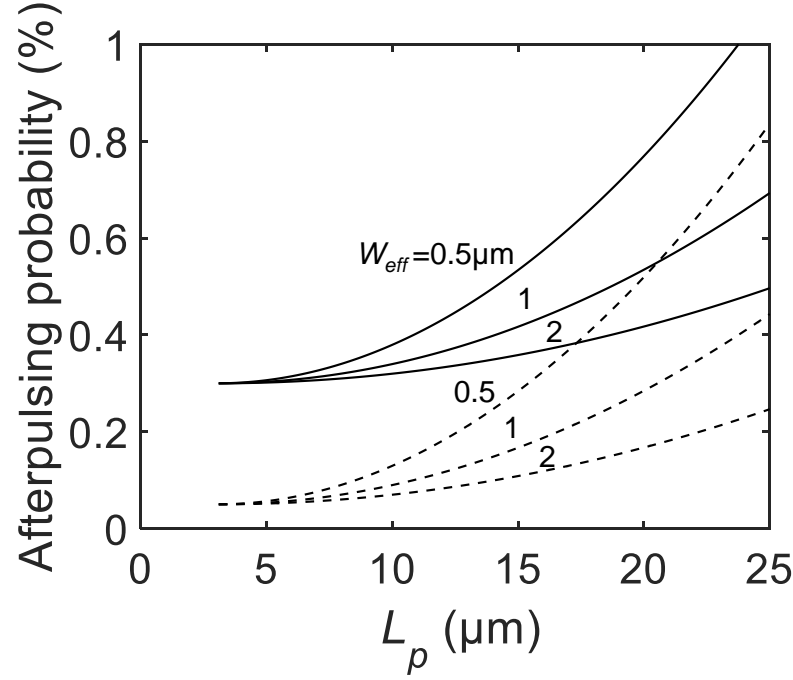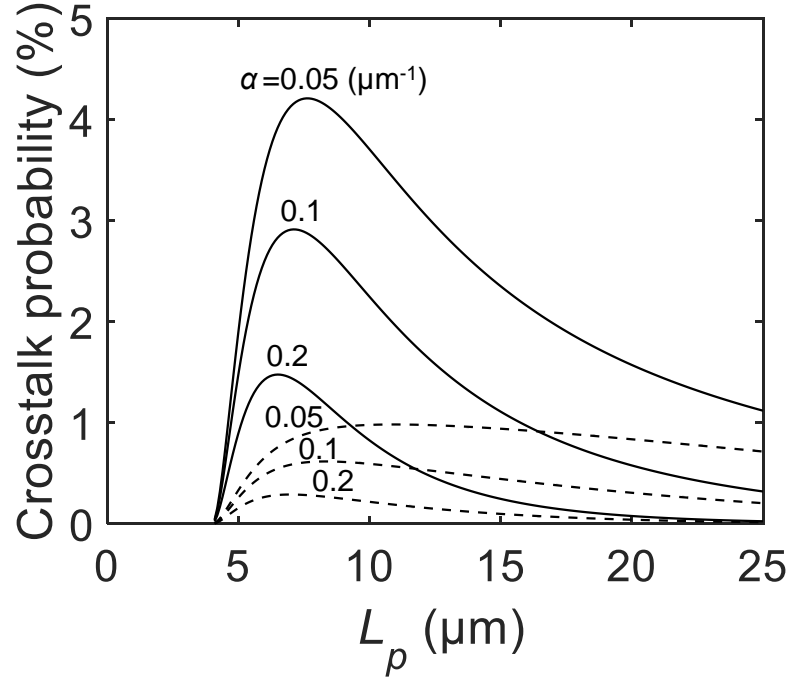
Figure 2.8 – Calculated crosstalk probability as a function of SPAD pixel pitch L_p for $C_0 = 5$ fF (dashed lines) and 30 fF (solid lines), active-to-active distance $L_{a-a} = 3$ μm, $B' = 5 \times 10^{14}$ m^2· F^{-1}, $W_{eff} = 1$ μm, $r_{in} = 0.5$ μm, and $\alpha = 0.05, 0.1, 0.2$ μm^{-1}.

PDP and PDE. V_{ex} affects both N_{ava} in the emitter and the sensitivity of the receiver, and hence the crosstalk probability follows the square law with respect to of V_{ex}. Second, formation of narrow deep trench isolation (DTI) could suppress the crosstalk. The trench material with lower refractive index can reflect the emitted photons and eventually confine the photons in the emitter. This could lead to an order of magnitude improvement of the crosstalk probability.

2.7 Power consumption

Avalanche-originated power consumption in large-scale SPAD arrays is a key parameter as it grows proportionally to the number of pixels. The total power consumption in the SPAD array depends on the incident photon flux. For a systematic comparison, the following discussion focuses on the energy consumption per single avalanche event, E_{ava}, in a single SPAD pixel. The power consumption at readout circuits is not taken into account here. E_{ava} is a product of eN_{ava} and $(V_{ex} + V_B)$, expressed as the following:

$$E_{ava} = C_{par} \times V_{ex} \times (V_{ex} + V_B) = D \times \left[\frac{\pi\epsilon(L_p - L_{a-a})^2}{4W_{eff}} + C_0 \right], \tag{2.14}$$

35
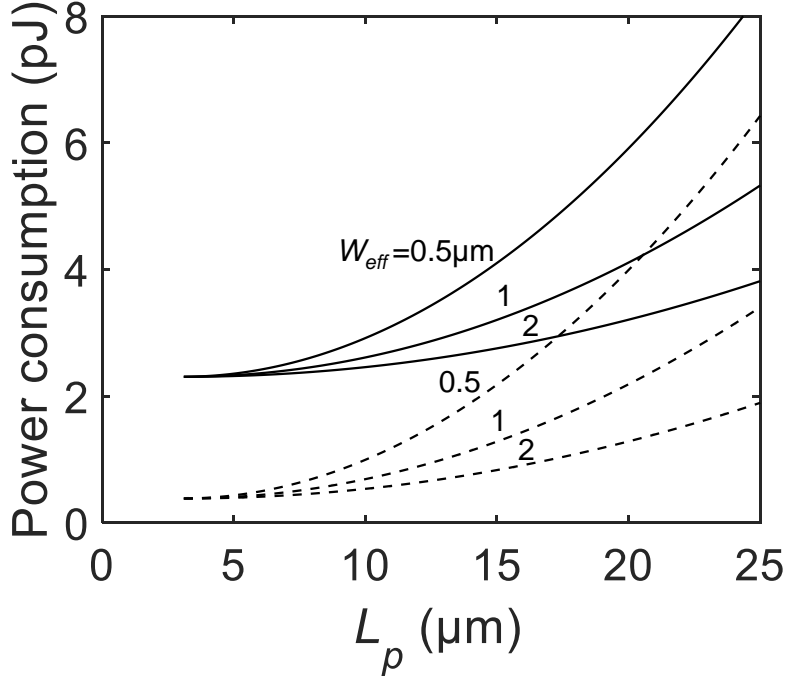
Figure 2.9 – Calculated power consumption as a function of SPAD pixel pitch L_p for $C_0 = 5$ fF (dashed lines) and 30 fF (solid lines), active-to-active distance $L_{a-a} = 3$ μm, $V_B = 20$ V, $V_{ex} = 3.3$ V, and $W_{eff} = 0.5, 1, 2$ μm.

where $D = V_{ex} \times (V_{ex} + V_B)$ is the bias-dependent coefficient, and V_B the breakdown voltage of the SPAD. Apart from the detail of the coefficient, the equation has the same structure as that of the afterpulsing probability. Naturally, the calculated trend of single-event power consumption E_{ava} as a function of L_p in Fig. 2.9 shows the similarity to Fig. 2.7.

2.8 Timing jitter

Timing jitter in the SPAD is determined by multiple factors such as device configuration, doping profile, detection threshold, excess bias, and temperature, and it is not straightforward to formulate the scaling law for it. Qualitatively, larger pixel pitch gives higher timing jitter for several reasons: first, spatial expansion of avalanche multiplication process takes more time in the larger L_p due to the finite lateral avalanche propagation velocity [29]. Second, larger L_p requires slower rising of the output voltage due to larger parasitic capacitance, leading to enhanced statistical variability. Further systematic analysis should be conducted for deeper understanding of scaling on timing jitter.

2.9 Conclusion

In the above sections, the scaling laws of the key SPAD characteristics with pixel dimensions have been investigated. Miniaturization of the SPAD pixel improves DCR, afterpulsing, power consumption and timing jitter, whereas it has an adverse effect on fill factor, PDP, PDE and crosstalk. The equations for the scaling laws are summarized in the Table 2.1. Especially, the degradation of single-photon sensitivity is inevitable in the conventional SPAD pixel when its pitch becomes smaller than 10 μm. Further technological breakthrough is a prerequisite for SPAD pixel miniaturization towards multi-megapixel arrays. Some of the novel technologies to push the limit of miniaturization will be introduced in the following chapters.

*NOTE: The theory for the scaling laws is supported by existing data from our group and from others in the literature. A systematic comparison of the theoretical and experimental plots has yet to be performed due to the COVID-19 disruption of our plans, and will be continued after the submission of this thesis.

Table 2.1 – Summary of scaling laws in the SPAD pixels with the pixel pitch L_p as a scaling parameter. The coefficient is omitted in the equations.

Characteristics	Equation
Fill factor (%)	$\frac{(L_p - L_{a-a})^2}{L_p^2}$
PDP (%)	$(\frac{L_p - L_{a-a} - 2r_{in}}{L_p - L_{a-a}})^2$
PDE (%)	$\frac{(L_p - L_{a-a} - 2r_{in})^2}{L_p^2}$
DCR (cps)	$(L_p - L_{a-a} - 2r_{in})^2$
DCR density (cps/μm^2)	$\frac{(L_p - L_{a-a} - 2r_{in})^2}{(L_p - L_{a-a})^2}$
Afterpulsing probability (%)	$\left[\frac{\pi\epsilon(L_p - L_{a-a})^2}{4W_{eff}} + C_0 \right],$
Crosstalk probability (%)	$\left[\frac{\pi\epsilon(L_p - L_{a-a})^2}{4W_{eff}} + C_0 \right] \times \frac{e^{-\alpha L_p}}{L_p^2} \times \frac{(L_p - L_{a-a} - 2r_{in})^2}{L_p^2}$
Power consumption (pJ)	$\left[\frac{\pi\epsilon(L_p - L_{a-a})^2}{4W_{eff}} + C_0 \right]$

References

[1] Z. You, L. Parmesan, S. Pellegrini, and R. K. Henderson, "3μm Pitch, 1μm Active Diameter SPAD Arrays in 130nm CMOS Imaging Technology," Int. Image Sensor Workshop (2017).

[2] Y. Maruyama, J. Blacksberg, and E. Charbon, "A 1024×8, 700-ps Time-Gated SPAD Line Sensor for Planetary Surface Exploration With Laser Raman Spectroscopy and LIBS," IEEE Journal of Solid-State Circuits 49(1), 179–189 (2014).

[3] C. Niclass, H. Matsubara, M. Soga, M. Ohta, M. Ogawa, and T. Yamashita, "A NIR-Sensitivity-Enhanced Single-Photon Avalanche Diode in 0.18μm CMOS ," Int. Image Sensor Workshop (2015).

[4] I. Gyongy, N. Calder, A. Davies, N. A. W. Dutton, P. Dalgarno, R. Duncan, C. Rickman, and R. K. Henderson, "256×256, 100kfps, 61% Fill-factor Time-resolved SPAD Image Sensor for Microscopy Application," IEEE Int. Electron Devices Meeting (2016).

[5] A. Gulinatti, F. Ceccarelli, I. Rech, and M. Ghioni, "Silicon technologies for arrays of Single Photon Avalanche Diodes," Proceedings of International Society for Optical Engineering 9858, Advanced Photon Counting Techniques X, 98580A (2016).

[6] C. Veerappan, "Single-Photon Avalanche Diodes for Cancer Diagnosis," PhD thesis, TU Delft (2016).

[7] J. A. Richardson, E. A. G. Webster, L. A. Grant, and R. K. Henderson, "Scaleable Single-Photon Avalanche Diode Structures in Nanometer CMOS Technology," IEEE Trans. Electron Devices 58(7), 2028-2035 (2011).

[8] C. Veerappan, and E. Charbon, "A Low Dark Count p-i-n Diode Based SPAD in CMOS Technology," IEEE Trans. Electron Devices 63(1), 65-71 (2015).

[9] J. M. Pavia, M. Wolf, and E. Charbon, "Measurement and modeling of microlenses fabricated on single-photon avalanche diode arrays for fill factor recovery," Opt. Express 22(4), 4202-4213 (2014).

[10] I. Gyongy, A. Davies, B. Gallinet, N. A. W. Dutton, R. Duncan, C. Rickman, R. K. Henderson, and P. A. Dalgarno, "Cylindrical microlensing for enhanced collection efficiency of small

References

pixel SPAD arrays in single-molecule localisation microscopy," Opt. Express 26(3), 2280-2291 (2018).

[11] E. Charbon, "Single-photon imaging in complementary metal oxide semiconductor processes," Philosophical Transactions of the Royal Society A, Mathematical, Physical and Engineering Sciences (2014).

[12] I. Takai, H. Matsubara, M. Soga, M. Ohta, M. Ogawa, and T. Yamashita, "Single-Photon Avalanche Diode with Enhanced NIR-Sensitivity for Automotive LIDAR Systems," Sensors 16(4), 459 (2016).

[13] R. Tielert, "Two-Dimensional Numerical Simulation of Impurity Redistribution in VLSI Processes," IEEE Trans. Electron Devices 27(8) (1980).

[14] F. Acerbi, G. Paternoster, A. Gola, N. Zorzi, and C. Piemonte, "Silicon photomultipliers and single-photon avalanche diodes with enhanced NIR detection efficiency at FBK," Nuclear Inst. And Methods in Physics Research, A (2017).

[15] A. Gola, "NUV-HD and NIR-HD SiPMs and Applications," Int. SPAD Sensor Workshop (2018).

[16] N. Teranishi, "Dark current and white blemish in image sensors," IEEE Int. Symposium on VLSI Technology, Systems and Application (2013).

[17] E. Charbon, "Status and Future Prospects of High Time Resolution Photon Counting Sensor Arrays," Topical Workshop on Electronics for Particle Physics (2015).

[18] E. A. G. Webster, J. A. Richardson, L. A. Grant, D. Renshaw, and R. K. Henderson, "A Single-Photon Avalanche Diode in 90-nm CMOS Imaging Technology with 44% Photon Detection Efficiency at 690 nm," IEEE Electron Device Letters 33(5) (2012).

[19] M.-J. Lee, P. Sun, and E. Charbon, "A first single-photon avalanche diode fabricated in standard SOI CMOS technology with a full characterization of the device," Opt. Express 23(10) (2015).

[20] H. Xu, L. Pancheri, G.-F. D. Betta, and D. Stoppa, "Design and characterization of a p+/n-well SPAD array in 150nm CMOS process," Opt. Express 25(11) (2017).

[21] C. Veerappan, and E. Charbon, "A Substrate Isolated CMOS SPAD Enabling Wide Spectral Response and Low Electrical Crosstalk," IEEE Journal of Selected Topics in Quantum Electronics 20(6) (2014).

[22] M. Unternährer, B. Bessire, L. Gasparini, D. Stoppa, and A. Stefanov, "Coincidence detection of spatially correlated photon pairs with a monolithic time-resolving detector array," Opt. Express 24(25) (2016).

[23] L. Gasparini, B. Bessire, M. Unternährer, A. Stefanov, D. Boiko, M. Perenzoni, and D. Stoppa, "SUPERTWIN: towards 100kpixel CMOS quantum image sensors for quantum optics applications," Proceedings of SPIE, vol. 10111, Quantum Sensing and Nano Electronics and Photonics XIV (2017).

[24] G. Lubin, R. Tenne, I. M. Antolovic, E. Charbon, C. Bruschini, and D. Oron, "Quantum correlation measurement with single photon avalanche diode arrays," arXiv:1910.01376 (2019)

[25] M. Liu, C. Hu, J. C. Campbell, Z. Pan, and M. M. Tashima, "Reduce Afterpulsing of Single Photon Avalanche Diodes Using Passive Quenching With Active Reset," IEEE J. Quantum Electronics 44(5) (2008).

[26] W. Mönch, "On the Physics of Avalanche Breakdown in Semiconducors," Physica Status Solidi 36(9) (1969).

[27] L. W. Snyman, M. Du Plessis, and H. Aharoni, Injection-Avalanche-Based n+pn Silicon Complementary Metal–Oxide–Semiconductor Light-Emitting Device (450–750nm) with 2-Order-of-Magnitude Increase in Light Emission Intensity," Japanese Journal of Applied Physics 46(4B) (2007).

[28] I. Rech, A. Ingargiola, R. Spinelli, I. Labanca, S. Marangoni, M. Ghioni, and S. Cova, "Optical crosstalk in single photon avalanche diode arrays: a new complete model," Opt. Express 16(12), 8381 (2008).

[29] M. Assanelli, A. Ingargiola, I. Rech, A. Gulinatti, and M. Ghioni, "Photon-Timing Jitter Dependence on Injection Position in Single-Photon Avalanche Diodes," IEEE J. Quantum Electronics 47(2) (2011).

3 Miniaturized SPAD pixel arrays with guard-ring-sharing technique

Miniaturization of SPAD pixel is fundamentally limited by its high voltage operation; a guard-ring region is necessary for suppressing unwanted electric field concentration at the edge of the active area, while pixel isolation region is required for electrical separation of adjacent pixels. Theoretical analysis in Chapter 2 have shown that, in the SPAD pixels below 10 μm-pitch, those photon-insensitive regions occupy large portion of pixel area, thus degrading the fill factor. We address this issue on a basis of principles of semiconductor device physics, and propose a novel structure to significantly enhance the fill factor in the miniaturized SPADs. This chapter is based on results presented in, K. Morimoto et al. "High fill-factor miniaturized SPAD arrays with guard-ring-sharing technique," Optics Express 28(9), 13068-13080 (2020).

3.1 Status of SPAD pixel scaling

As in conventional CCDs and CISs, pixel array size contributes to signal-to-noise ratio and dynamic range, ultimately determining image quality. When the sensor format is fixed, larger array sizes can only be achieved by shrinking pixel pitch. For over 15 years, there has been a continuous effort to increase the size of SPAD arrays and thus reducing pixel pitch has been a great part of that effort. The smallest SPAD pixel pitch reported is 6 μm for large-scale imagers [1] and 3 μm for test structures [2]. Due to the existence of guard-ring structures, where no photons can be detected, shrinking SPAD pixel pitch below 10 μm could lead to the reduced fill factor. Several countermeasures to this issue have been proposed recently, such as well-sharing [3], on-chip microlens [4, 5, 6], and 3D-stacking [7, 8, 9, 10], which enable better fill factor in small SPAD pixels. Yet, the fundamental limit of SPAD pixel miniaturization lies in the photon-insensitive guard-ring and pixel isolation structures.

In this chapter, we demonstrate a novel guard-ring-sharing technique to push the limits of SPAD pixel miniaturization. The SPAD performance has been verified in guard-ring-shared 4×4 arrays of 2.2 μm-pitch SPAD, the smallest pixel pitch ever reported. In addition, significant improvement of fill factor, DCR, PDE, and timing jitter performance with respect to the state-of-the-art small SPAD pixels is observed in the guard-ring-shared 3 μm and 4 μm pixels. The 2.2 μm-pitch SPAD pixels showed excellent performance in afterpulsing and timing jitter,

while the DCR and PDP are deviated from larger pixels. Combined with on-chip microlens and 3D-stacking approaches, the proposed technique provides a promising solution to achieve a compact multi-megapixel SPAD array.

3.2 Physical limit of SPAD pixel miniaturization

To address the physical limit of SPAD pixel miniaturization, the need of pixel circuitry adds uncertainty to pixel pitch and fill factor, given that the circuit area is highly dependent on technology node, architecture, and layout. Thanks to recent progress in 3D-stacked technologies though, this uncertainty is largely removed and SPADs can be assumed to be densely packed, thus the influence of pixel circuit area is not taken into account in this chapter, unless otherwise noted. The discussion can be naturally extended to monolithic 1D [11] or small-scale 2D arrays [12, 13], where pixel circuits are located outside the array.

A schematic cross-section of the p-i-n SPAD [14], combined with well-sharing technique [3], is shown at the top of Fig. 3.1(a), where the p-well defines the active area of the photosensor. Key parameters in pixel miniaturization are active diameter D_a, active-to-active distance L_{a-a}, and pixel pitch L_p, where $L_p = D_a + L_{a-a}$. Assuming circular active area, the SPAD fill factor is given as:

$$FF = \frac{\pi D_a^2}{4L_p^2} = \frac{\pi D_a^2}{4(D_a + L_{a-a})^2}. \tag{3.1}$$

For a given pixel pitch L_p, smaller L_{a-a} is desirable to improve fill factor. In the well-shared SPADs, L_{a-a} consists of isolation well (NW) width and twice the guard-ring width. Typically, the guard-ring width is designed to be at least 1 μm to avoid premature edge breakdown, whereas the minimum isolation well width is 0.5 μm due to process requirements. Hence, L_{a-a} smaller than 2.5 μm is not practically useful to guarantee reasonable PDP and DCR, as well as high levels of manufacturability. Under the above assumptions, when targeting 20% fill factor, the theoretical limit of well-shared SPAD pitch is 5 μm, irrespective of fabrication technology. Pixel pitch smaller than this limit has been investigated previously [2], whereas the measured low PDP and high DCR suggest that the device suffered from premature edge breakdown.

3.3 Device structure, simulation and preliminary test

To overcome the limitations of SPAD miniaturization, we propose a novel guard-ring-sharing technique. The schematic cross-section of guard-ring-shared SPAD is shown at the bottom of Fig. 3.1(a). Compared to the conventional well-shared structure, the isolation well between

neighboring pixels is eliminated. The pixel is virtually isolated by the shared guard-ring region with a shallow trench isolation (STI). Assuming the shared guard-ring width of 1 μm, the theoretical limit of pixel pitch can be reduced to 2 μm for 20% fill factor. In Fig. 3.1(b), the electric field distribution is extracted by 2D-TCAD simulation. In contrast to well-shared devices (top), guard-ring-shared devices feature neighboring active regions in close proximity (bottom). The corresponding simulated electrostatic potential distribution is shown in Fig. 3.1(c). The potential profiles on the white dashed lines in Fig. 3.1(c) are plotted in Fig. 3.1(d). For well-sharing (top), potential barrier height between two active area is determined by cathode voltage V_C, typically in the order of 15 to 30 V.

In contrast, the minimum required potential barrier to operate the SPAD is determined by excess bias V_{ex}; when reducing the potential barrier height, an increase in the punch-through current between neighboring p-wells may induce electrical coupling of output signals, thus leading to the electrical crosstalk. The punch-through current is maximized when the potential difference between the p-wells is maximum, i.e. when one SPAD is quenched and the adjacent SPAD is not quenched. The maximum bias difference is V_{ex}, and hence a potential barrier higher than V_{ex} is sufficient to block punch-through. For guard-ring sharing (bottom), the isolation region between p-wells is fully depleted, and the potential barrier is reduced below 10 V, but it can be higher than V_{ex}, typically 1 to 6 V. Careful design of the potential barrier by optimizing doping conditions enables the significant reduction of active-to-active distance without degradation of the crosstalk and other SPAD properties.

To experimentally confirm the feasibility of the guard-ring-sharing technique, preliminary tests were performed with a 3-terminal two-SPAD device. Fig. 3.2 shows the measured results. As shown in the inset of Fig. 3.2(a), two guard-ring-shared SPADs are designed, where cathode voltage V_C and anode voltages V_{A1} and V_{A2} can be applied externally. In the experiment, V_{A2} is fixed at 0 V, and V_{A1} is swept from -8 V to 8 V for several different values of V_C (8 to 15 V with 1 V-step). The breakdown voltage of this SPAD is above 23 V, and no avalanche current is expected in the measured range. D_a and L_{a-a} for the measured device were 3 μm and 1 μm, respectively. As shown in Fig. 3.2(a), no significant current to anode 1 (I_{A1}) is observed for $V_{A1} < 3$ V. Exponential increase of I_{A1} is observed for positive V_{A1}, indicating punch-through current. The rising value of I_{A1} is dependent on V_C. The critical voltage for punch-through, V_{PT}, can be defined as V_{A1}, giving $I_{A1} = 100$ pA. Note that the absolute value of current to anode 2 (I_{A2}) was equivalent to I_{A1}, whereas cathode current (I_C) was below the measurement limit of 10 pA over the measured range. Fig. 3.2(b) shows the V_C dependence of V_{PT}, exhibiting linear increase of V_{PT} for V_C. The result is consistent with our simulations, where potential barrier height increases with increasing cathode voltage. Linear fitting of the data is shown as a dashed line, indicating that in the Geiger mode, with V_C above 20 V, the potential barrier is high enough to suppress the punch-through current with V_{ex} at up to 10 V. The results strongly suggest that in an optimized process and device conditions, adjacent p-wells can be electrically isolated without isolation well. Note that the proposed structure is distinct from STI-bounded SPADs [15], where the STI has a direct contact with the multiplication region and its surface traps degrade the noise performance. In our device, the STI is spatially isolated from the buried multiplication region to avoid influence on the noise performance. The STI
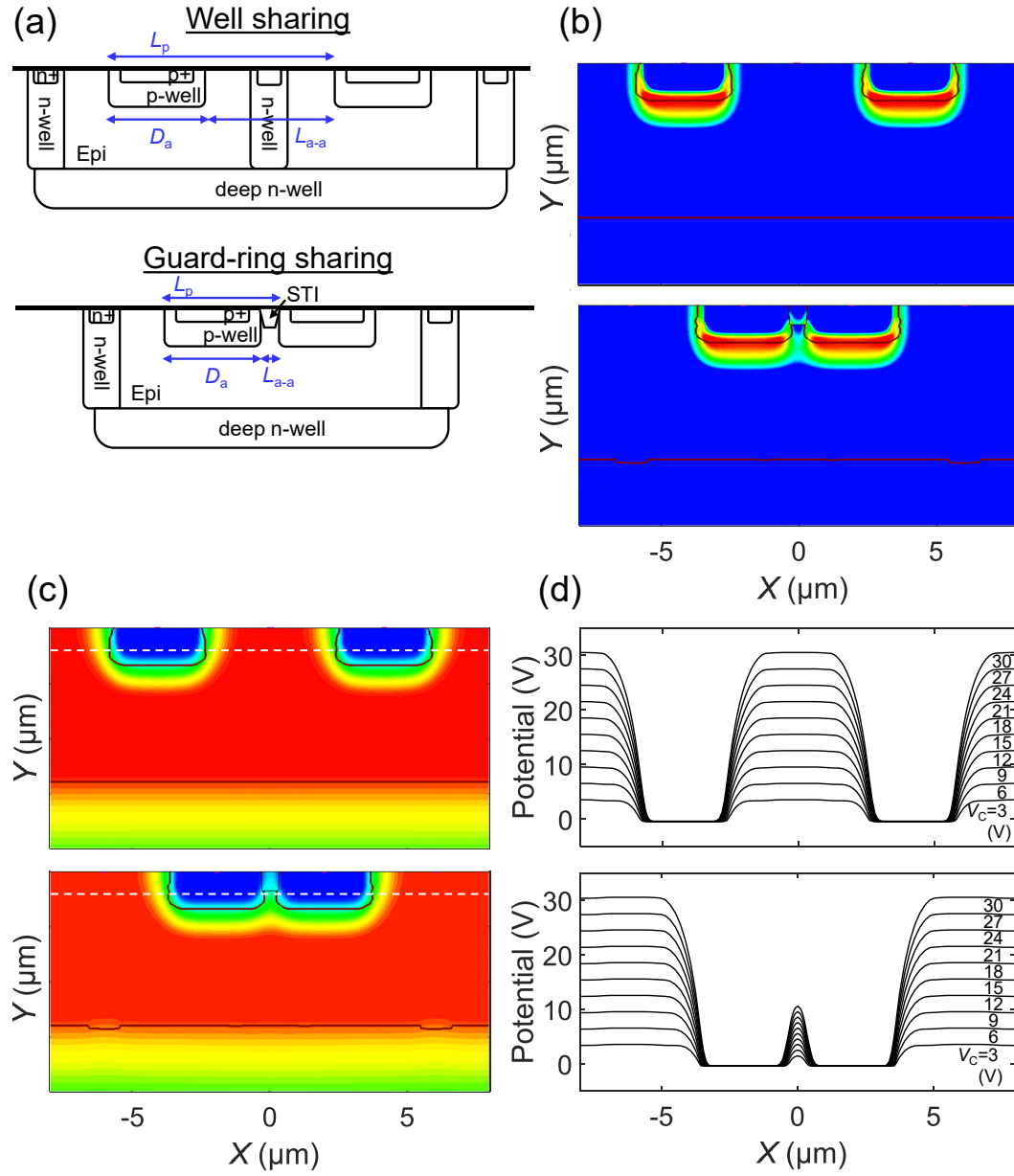
Figure 3.1 – Comparison of well-sharing (top) and guard-ring-sharing techniques (bottom). (a) Cross-sectional views of p-i-n SPADs. (b) Simulated electric field distributions. (c) Simulated electrostatic potential. (d) 1D potential profile on the white dashed line in (c) for cathode voltage V_C = 3 to 30 V with 3 V-step.
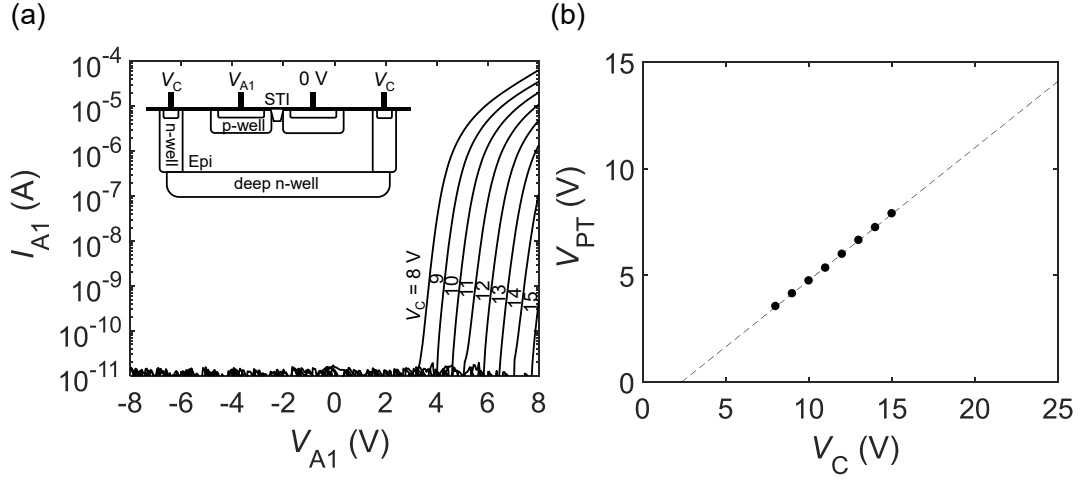
Figure 3.2 – Results of preliminary test with 3-terminal two-SPAD device. (a) I-V characteristics for anode 1 with V_{A2} fixed at 0 V and V_C at 8, 9, 10, 11, 12, 13, 14, 15 V. Inset shows a schematic cross-section of the test device. (b) Punch-through voltage V_{PT} as a function of V_C, where V_{PT} is defined as V_{A1} giving I_{A1} = 100 pA. Dashed line is the linear fit.

between p-wells is introduced to suppress punch-through and to enhance V_{PT}. Separate measurements were carried out with a guard-ring-shared device without the STI isolation (not shown), which showed a similar trend but with slightly lower V_{PT}.

3.4 Test chip design

4×4 SPAD arrays with selective readout circuits were designed with monolithic 180 nm CMOS process, which was customized for low noise SPAD, to verify the Geiger-mode operation of guard-ring-shared SPADs. Fig. 3.3(a) shows a circuit diagram; 8 different types of 4×4 SPAD arrays were designed in a single chip. The anode terminal of each SPAD in an array is connected to an array of pixel circuits in one-to-one correspondence. The pixel circuit comprises a quenching and a cascode transistor [16], followed by a level-shifting inverter and two parallel tristate inverters. Using demultiplexers and multiplexers, pulsed photon counting signals from two arbitrary SPADs in a single 4×4 array of interest can be streamed out through OUT1 and OUT2. The two parallel output configuration enables direct measurement of inter-avalanche time correlation for two SPADs, as well as the parallel measurement of DCR, PDP, afterpulsing probability and timing jitter for single SPAD.

Fig. 3.3(b) shows the chip micrograph. The size of a single chip is 1.8 mm×1.8 mm, and four dies with the same format and readout circuit, but with different SPAD structures and dimensions are designed. A single die contains two sets of devices shown in Fig. 3.3(a). In total, 64 different types of SPADs are tested. Three pads at the corners of the chip are for the preliminary test devices discussed in the previous section.
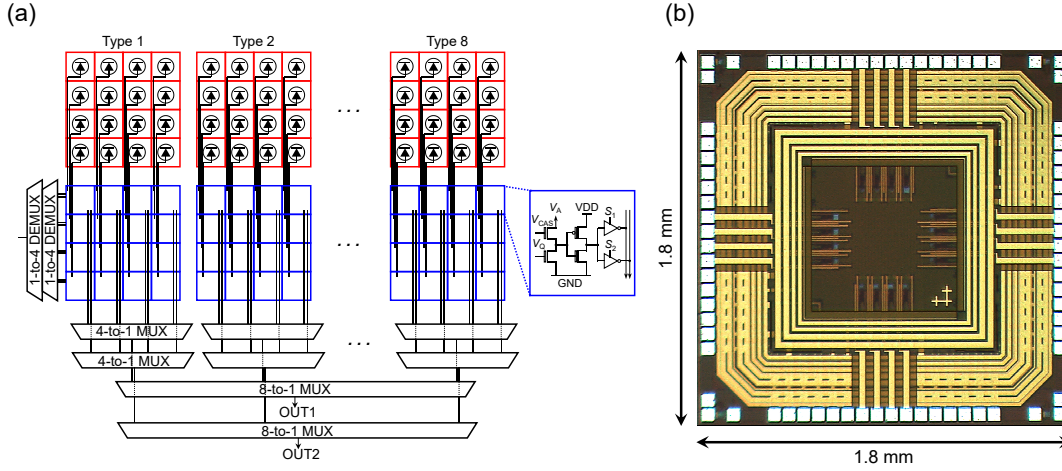
Figure 3.3 – Architecture of eight 4×4 SPAD arrays. (a) Circuit diagram of 4×4 SPAD arrays with two parallel selective readout circuits. (b) Chip micrograph with the size of 1.8 mm×1.8 mm.

3.5 Experimental results

3.5.1 Fill factor and PDP

Fig. 3.4(a) shows a geometrical fill factor of designed SPADs. The fill factor is defined as the drawn active area subtracted by the area of metal layers overlapping the active area. Active diameters D_a = 1.2, 2, and 3 μm are designed in combination with active-to-active distances L_{a-a} = 1, 1.6, 2.2, 2.8, and 3.4 μm. Owing to the guard-ring-sharing technique, fill factors of up to 19.5%, 32.3% and 42.4% were achieved for the pixel pitches L_p = 2.2, 3, and 4 μm, respectively. Note that, when adopting this technique to 3D-stacking process, the fill factor can be further enhanced due to the lack of shadowing effect by metal interconnects.

Fig. 3.4(b) shows the measured PDP as a function of wavelength. The plotted PDP is calculated as an average of PDPs from two pixels in an array. As in the conventional SPAD without guard-ring sharing, the PDP increases by increasing the excess bias voltage V_{ex}. For the SPAD with 3 μm active diameter and 1 μm active-to-active distance, maximum PDP of 33.5% is obtained with V_{ex} = 6 V at λ = 500 nm. The corresponding PDE, defined as a product of PDP and fill factor, is 14.2%, which is considerably higher than state-of-the-art miniaturized SPADs. Relatively broad spectrum is observed, and the PDP with V_{ex} = 6 V at λ = 900 nm is measured at 3.1%. The trend is consistent with prior art based on the p-i-n SPADs [14], indicating that the guard-ring sharing has no significant impact on PDP.

Fig. 3.4(c) shows the measured maximum PDP as a function of the active-to-active distance L_{a-a}. Slight increase of the maximum PDP is observed with reducing L_{a-a}. This could be caused by a deformed electric field distribution for smaller L_{a-a}, leading to the reduced loss of photocharge detection around the corner of the active area. Fig. 3.4(d) shows the excess bias dependence of the maximum PDP for different active diameters. The maximum PDPs for D_a = 1.2 μm and 2 μm are smaller compared to that of D_a = 3 μm, likely due to the border
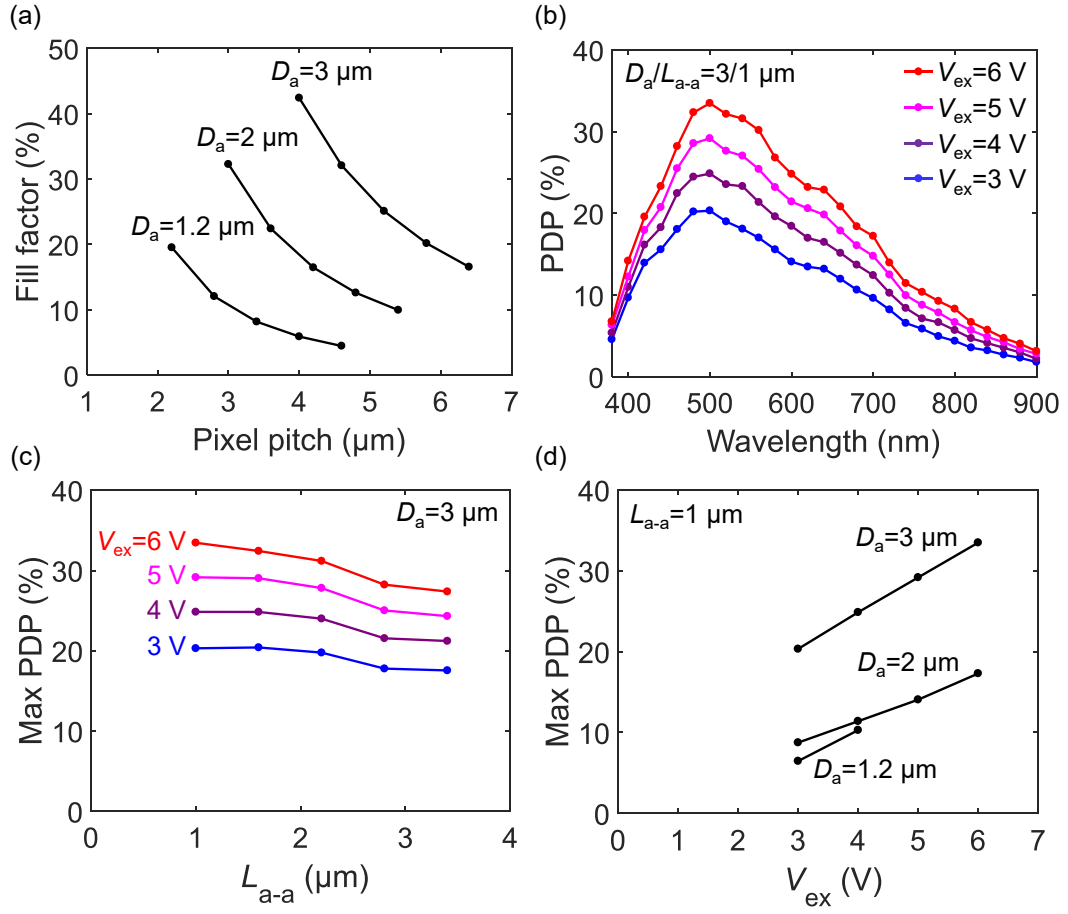
Figure 3.4 – Fill factor design and PDP measurement results for guard-ring-shared SPADs. (a) Fill factor of designed SPADs with different active diameters and pixel pitches. (b) Measured PDP for SPAD with $D_a = 3$ μm and $L_{a-a} = 1$ μm as a function of wavelength. (c) Maximum PDP as a function of L_{a-a}. (d) Maximum PDP as a function of excess bias and D_a.

effect for reduced active diameters [17].

3.5.2 DCR

Fig. 3.5 shows the measured room temperature DCR performance. The plotted DCR is median of 16 pixels in an array. Active-to-active distance dependence of DCR for $D_a = 3$ μm is plotted in Fig. 3.5(a). For $V_{ex} = 3$ V, no significant dependency of DCR is observed for smaller L_{a-a}. For $V_{ex} = 6$ V, the DCR is slightly increased at $L_{a-a} = 1$ μm, whereas the absolute values of the DCR are still much smaller compared to the prior arts thanks to small active area and low noise process.

Fig. 3.5(b) shows the excess bias dependence of room temperature DCR for different active diameters. Smaller DCR for $D_a = 2$ μm is observed with respect to $D_a = 3$ μm due to reduced active area. Limited increase of DCR is observed up to $V_{ex} = 6$ V, which implies that tunneling-
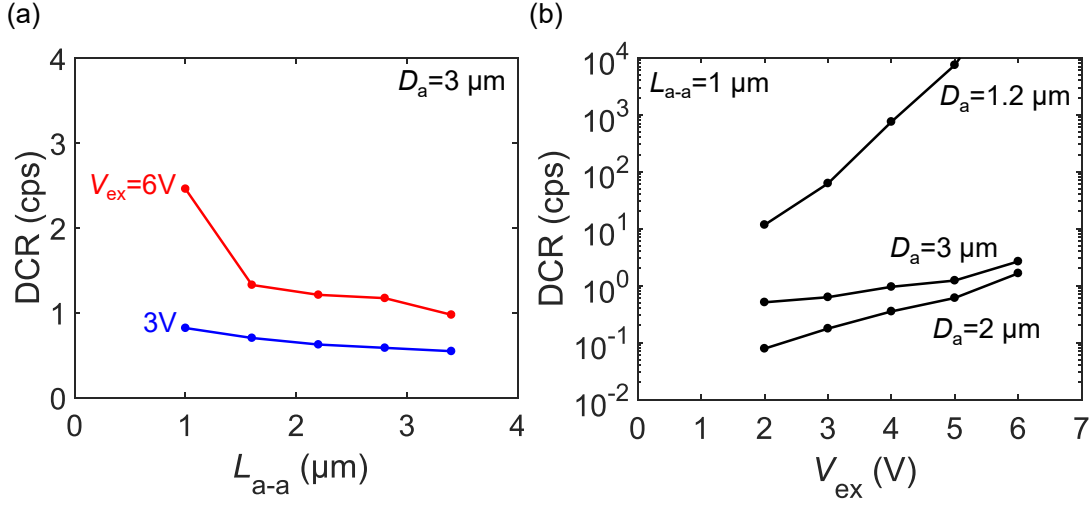
Figure 3.5 – DCR characteristics for guard-ring-shared SPADs. (a) Measured room temperature median DCR for SPAD with D_a = 3 μm as a function of L_{a-a}. (b) Median DCR as a function of excess bias and D_a with L_{a-a} = 1 μm.

induced DCR component is well-suppressed. The DCR for D_a = 1.2 μm is higher than that of larger active diameters, indicating that the aggressive design of the active area results in the higher risk of premature edge breakdown.

3.5.3 Crosstalk

As discussed previously, crosstalk is one of the most critical properties in miniaturized SPADs, which could potentially be influenced by guard-ring sharing. In general, crosstalk in SPADs can be classified in optical and electrical crosstalk [18, 19]. Optical crosstalk in SPAD is caused by avalanche-induced light emission, i.e. emitted photons travel towards nearby pixels and generate secondary electron-hole pairs, thus inducing other avalanche multiplication events. Electrical crosstalk is caused by avalanche-generated excess charges. One possible mechanism is an avalanche-induced hot carrier traveling across the pixel isolation to induce the additional avalanche multiplication event in the adjacent pixels. Another cause is a punch-through current between adjacent active regions; when a SPAD detects a photon, the voltage at the anode (or cathode in case of the reversed p-n configuration) swings with an amplitude of nearly V_{ex}. The punch-through current can pull the anode voltage of the adjacent pixel, which could go beyond the threshold of pixel circuit and generate an output pulse. The difference between the two electrical crosstalk mechanisms is whether the secondary output pulse is generated by avalanche effect or direct macroscopic leakage current. In typical SPAD arrays, both of the electrical crosstalk components are suppressed due to sufficiently high potential barrier at the pixel isolation.

Quantitative evaluation of optical and electrical crosstalk was experimentally performed by comparing well-shared and guard-ring-shared SPAD structures. Fig. 3.6(a) shows the cross-
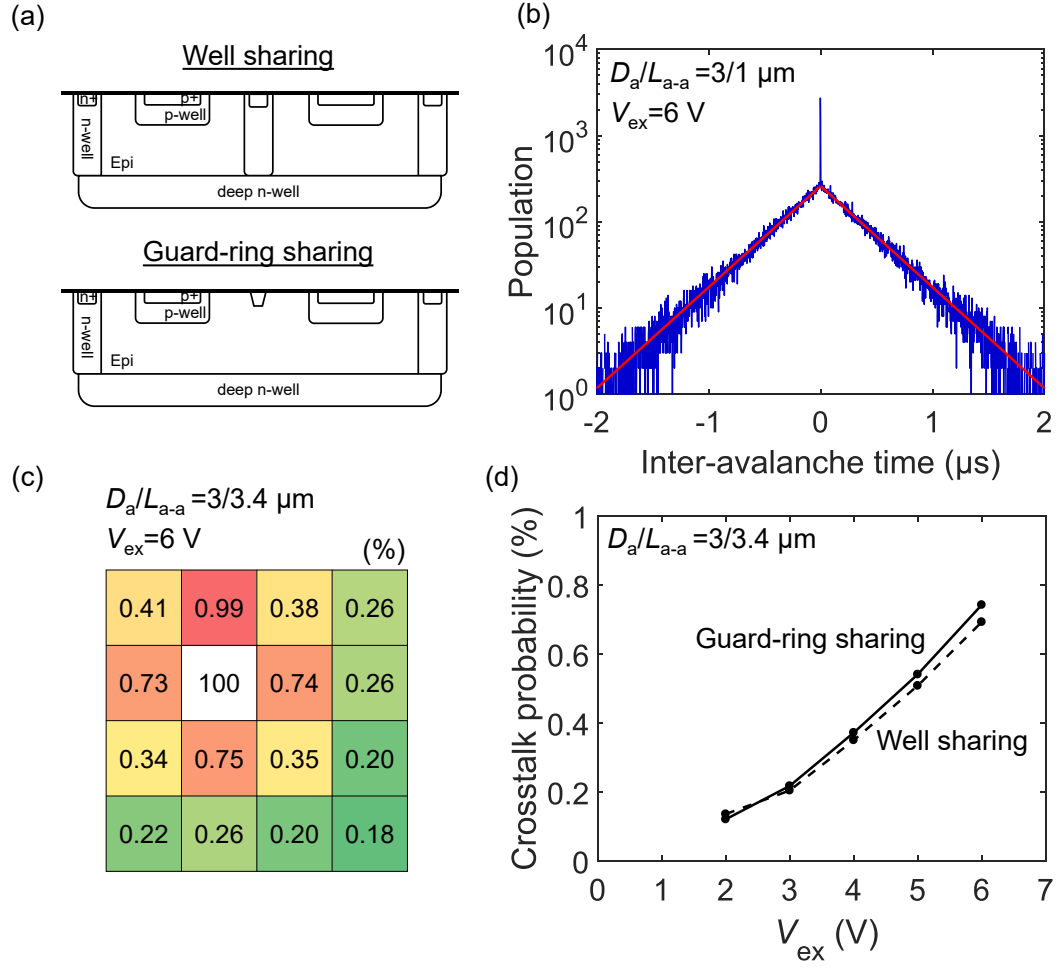
Figure 3.6 – Comparison of crosstalk between well sharing and guard-ring sharing. (a) Cross-sectional views of well-shared SPADs (top) and guard-ring-shared SPADs (bottom) with identical D_a and L_{a-a}. (b) Inter-avalanche time histogram for two adjacent pixels with $D_a = 3$ μm and $L_{a-a} = 1$ μm. Red curve is the exponential fit. (c) Crosstalk map in a 4×4 array with $D_a = 3$ μm and $L_{a-a} = 3.4$ μm. (d) Crosstalk probability as a function of excess bias with $D_a = 3$ μm and $L_{a-a} = 3.4$ μm.
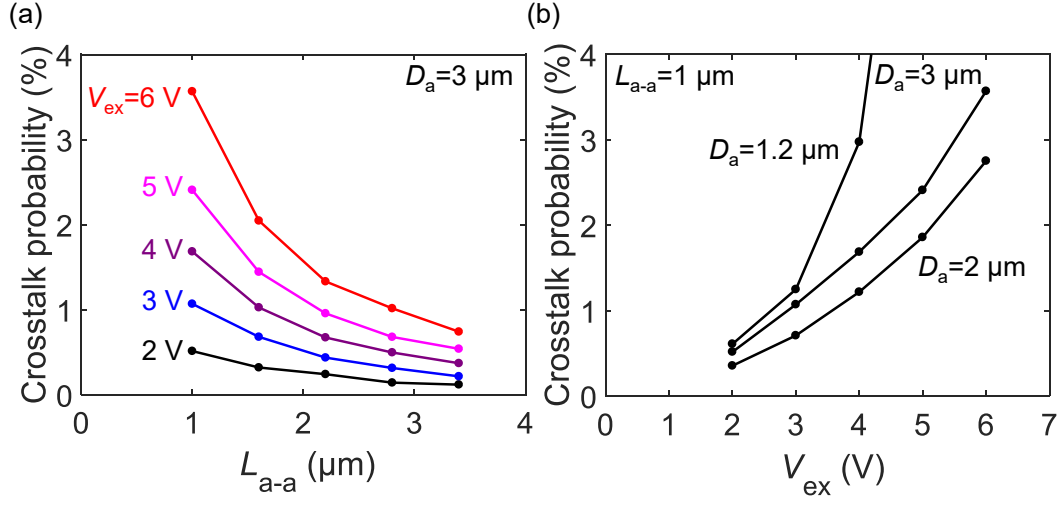
Figure 3.7 – Measured crosstalk for guard-ring-shared SPADs. (a) Crosstalk probability for SPAD with D_a = 3 μm as a function of L_{a-a}. (b) Crosstalk probability as a function of excess bias and D_a with L_{a-a} = 1 μm.

sections of well-shared and guard-ring-shared SPADs with identical pixel dimensions; when the active diameter and the active-to-active distance are identical for both structures, the effect of optical crosstalk should be identical because the amount of light emission per single avalanche event stays the same, and propagation of emitted photons is not affected by the doping profile of pixel isolation. Thus, differences in crosstalk for the two structures can be caused solely by electrical crosstalk.

In 4×4 SPAD arrays, crosstalk is measured by inter-avalanche time histogramming [2]. Assuming ideal SPAD array with no correlation of photon detection events between two pixels, an inter-avalanche time histogram shows exponential decay. The crosstalk component can be extracted as a deviation from the exponential behavior. Fig. 3.6(b) shows the histogram of measured inter-avalanche time for two adjacent pixels with D_a = 3 μm and L_{a-a} = 1 μm, taken under low light condition. The time bin width for the histogram is set at 2 ns. A central peak indicates the time-correlated crosstalk events. The tails are fitted with exponential functions (red lines). Fig. 3.6(c) shows the crosstalk map for a 4×4 guard-ring-shared pixel array with D_a = 3 μm and L_{a-a} = 3.4 μm at V_{ex} = 6 V. Each crosstalk probability is independently estimated from the corresponding inter-avalanche time histogram. The crosstalk probability is higher for pixels closer to a reference pixel (shown in white). Fig. 3.6(d) shows the calculated crosstalk probability as a function of the excess bias. The plotted crosstalk is the median of 4 pixels adjacent to a single reference pixel. No critical difference is observed between the well sharing and the guard-ring-sharing devices. This result strongly suggests that guard-ring sharing leads to no significant degradation of electrical crosstalk with respect to the conventional well-sharing technique. Detailed analysis of the crosstalk in the guard-ring-shared SPAD arrays is performed. Fig. 3.7(a) shows the active-to-active distance dependence of the measured crosstalk probability for different excess biases. The crosstalk probability is increased with

decreasing L_{a-a} and increasing V_{ex}. The highest crosstalk probability of 3.57% is smaller than typical crosstalk probability of 5 to 10% in CISs. The level of crosstalk is acceptable for most of the applications mentioned in this thesis, while it could limit key performance in some specific applications, whereever correlated noise has a major impact [20]. Fig. 3.7(b) shows the crosstalk probability as a function of the excess bias for different active diameters. Analogous to the trend for DCR, the lowest crosstalk probability is obtained at $D_a = 2$ μm. Degradation of crosstalk is observed at $D_a = 1.2$ μm. The non-monotonic dependence of the crosstalk probability on D_a is unexpected because it is against the implication of the scaling laws in Chapter 2. This could be due to a negative impact of the premature edge breakdown on the crosstalk, in contrast to the scaling law analysis where the premature edge breakdown is assumed to be suppressed. Quadratic increase of the crosstalk probability stems from two factors: linear increase of photon emission for V_{ex}, and linear increase of PDP for V_{ex}. Note that the crosstalk probability can be further suppressed by employing 3D-stacking processes, where reduction of parasitic capacitance for the anode regulates the avalanche-induced light emission.

3.5.4   Afterpulsing and timing jitter

Fig. 3.8(a) shows the measured inter-avalanche histogram for a single pixel with $D_a = 3$ μm and $L_{a-a} = 1$ μm, taken under low light condition. In contrast to the crosstalk measurement, where inter-avalanche time for two different pixels is extracted, the inter-avalanche time between two adjacent output pulses in the same SPAD is monitored to evaluate afterpulsing probability. The SPAD dead time is fixed at 10 ns. The histogram shows exponential decay in the population. The red line shows the fitting result with the ideal case. Magnified view for shorter inter-avalanche time is shown in the inset. No positive deviation of the histogram from the ideal trend indicates that the afterpulsing is suppressed in the measured device. Calculated afterpulsing probability is 0.21%, close to the measurement limit. Similar afterpulsing probability was observed for different active diameter and active-to-active distance.

Fig. 3.8(b) shows the measured histogram for timing jitter. A 785 nm laser pulsed at 25 MHz (average power: 5 mW, optical pulse width: 80 ps, ALS GmbH, Berlin, Germany) illuminated the SPAD array through ND-filter, so as to observe the timing jitter of output pulses. The obtained histogram was fitted by Gaussian distribution as shown in a red curve. The raw timing jitter was 147 ps full width at half maximum (FWHM) at $V_{ex} = 6$ V, which comprises SPAD jitter and circuit-induced jitter. The latter component can be independently measured by monitoring an identical SPAD signal through two different paths towards OUT1 and OUT2. The readout circuits are designed symmetric for the two paths, and the timing jitter between OUT1 and OUT2 gives $\sqrt{2}$ times the circuit jitter. The circuit jitter component is estimated to be 117.7 ps. As a result, SPAD timing jitter for $D_a = 1.2$, 2, and 3 μm is calculated to be 72, 70, and 88 ps, respectively.

The measured afterpulsing probability and timing jitter are similar or better than the state-of-the-art SPAD devices, suggesting that the guard-ring sharing has no impact on those performance measures.
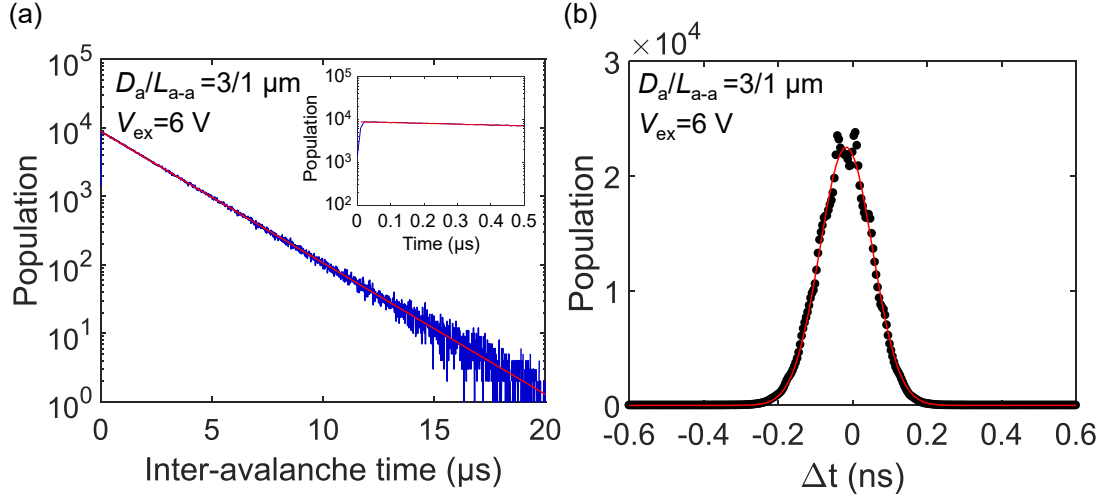
Figure 3.8 – Correlated noise analysis for guard-ring-shared SPAD with $D_a = 3$ μm and $L_{a-a} = 1$ μm. (a) Inter-avalanche histogram for extraction of afterpulsing probability. Red curve is the exponential fit. (b) Timing jitter histogram. Red curve is the Gaussian fit.

3.6 Conclusion

We proposed a novel guard-ring-sharing technique enabling high fill factor, miniaturized SPAD pixels, thus paving the way to very large format SPAD imagers. Feasibility of the guard-ring-sharing technique is verified by device simulation, preliminary I-V test and detailed measurement of 4×4 SPAD arrays. A pixel pitch of 2.2 μm was demonstrated, the smallest ever reported for a SPAD. Table 3.1 shows state-of-the-art comparison of performance and specifications in miniaturized SPAD arrays. In spite of an extremely small pixel pitch of 2.2, 3, and 4 μm, the guard-ring-shared SPAD arrays showed excellent fill factor, PDP, PDE, DCR, afterpulsing, and timing jitter performance. Crosstalk probability is relatively high due to aggressive miniaturization of active-to-active distance and much higher excess bias. Note that the concept of guard-ring sharing is not limited to p-i-n SPAD, but also applicable to other SPAD structures such as a p+/n-well SPAD, a p-well/buried-n-well SPAD, and a SPAD for enhanced near-infrared PDP [21, 22, 23].

Note that a large array of guard-ring-shared SPADs could suffer from voltage drop through a single shared deep-well. One of the scalable solutions for this issue is a hybrid approach: guard-ring shared in vertical and horizontal directions and not shared in diagonal directions. In such a case, the SPADs are well-shared in diagonal directions, and thus the cathode contacts can be placed in every corner between pixels. Another solution is to make a small SPAD sub-array (e.g. 8×8 pixels) to form a large-scale array, where every small sub-array can be surrounded by well-ring for cathode contact. This solution effectively reduces the average pixel pitch, while changing the periodicity of the pixels.

The guard-ring sharing technique is even more effective when combined with other miniaturization techniques, such as on-chip microlens and 3D-stacking. For 3D-stacked SPAD sensors,

the dimension of pixel circuitry is predicted to be under 1 μm by employing advanced logic processes, such as 11 nm-CMOS [24], and thus the limiting factor of SPAD pixel scaling is considered to reside in the SPAD arrays rather than the circuits. In such a small SPAD pixels whenever implemented in an advanced logic process, a cascode transistor or poly-resistor could be necessary for level-shifting photon count signals, so as to ensure high excess bias and low voltage operation of logic circuit. The guard-ring sharing can go beyond the limit of the conventional SPAD technologies, and could play a critical role for achieving compact multi-megapixel SPAD sensors, and ultimately 1 μm-pitch SPAD arrays towards a next generation of quanta image sensors [25] in the near future.

Table 3.1 – State-of-the-art comparison of performance and specifications in miniaturized SPAD arrays.

	[8]	[26]	[2]	This work		
Process technology	65/40 nm 3D-BSI CMOS	90 nm CMOS	130 nm CIS	180 nm CMOS		
Pixel pitch (μm)	7.83	5	3	2.2	3	4
Active diameter (μm)	-	2	1	1.2	2	3
Drawn fill factor (%)	45	12.5	14	19.5	32.3	42.4
Sensor resolution	128×120	3×3	4×4	4×4	4×4	4×4
Breakdown voltage (V)	12	10.3	15.8	32.35	23.6	22.1
Max. PDP (%)	27.5 (V_{ex} = 3 V)	36 (V_{ex} = 0.6 V)	15 (V_{ex} = 3.2 V)	10.3 (V_{ex} = 4 V)	17.3 (V_{ex} = 6 V)	33.5 (V_{ex} = 6 V)
Max. PDE (%)	12.4 (V_{ex} = 3 V)	4.5 (V_{ex} = 0.6 V)	2.1 (V_{ex} = 3.2 V)	2.0 (V_{ex} = 4 V)	5.6 (V_{ex} = 6 V)	14.2 (V_{ex} = 6 V)
Median DCR (cps)	11,000 (V_{ex} = 3 V)	250 (V_{ex} = 0.6 V)	150 (V_{ex} = 1 V)	751 (V_{ex} = 4 V)	1.6 (V_{ex} = 6 V)	2.5 (V_{ex} = 6 V)
Crosstalk (%)	-	<0.1 (V_{ex} = 0.6 V)	0.13-0.22 (V_{ex} = 1 V)	2.97 (V_{ex} = 4 V)	2.75 (V_{ex} = 6 V)	3.57 (V_{ex} = 6 V)
Afterpulsing probability (%)	-	-	0.18 (V_{ex} = 1 V)	<0.20 (V_{ex} = 4 V)	0.2 (V_{ex} = 6 V)	0.21 (V_{ex} = 6 V)
Timing jitter (ps)	136 (V_{ex} = 3 V)	107 (V_{ex} = 0.6 V)	185 (V_{ex} = 3 V)	72 (V_{ex} = 4 V)	70 (V_{ex} = 6 V)	88 (V_{ex} = 6 V)

References

[1] Y. Hirose, S. Koyama, T. Okino, A. Inoue, S. Saito, Y. Nose, M. Ishii, S. Yamahira, S. Kasuga, M. Mori, T. Kabe, K. Nakanishi, M. Usuda, A. Odagawa, and T. Tanaka, "A 400×400-Pixel 6μm-Pitch Vertical Avalanche Photodiodes CMOS Image Sensor Based on 150ps-Fast Capacitive Relaxation Quenching in Geiger Mode for Synthesis of Arbitrary Gain Images," IEEE Int. Solid-State Circuits Conference (2019).

[2] Z. You, L. Parmesan, S. Pellegrini, and R. Henderson, "3μm Pitch, 1μm Active Diameter SPAD Arrays in 130nm CMOS Imaging Technology," Int. Image Sensor Workshop (2017).

[3] T. Al Abbas, N. Dutton, O. Almer, F. M. D. Rocca, S. Pellegrini, B. R. Rae, D Golanski, and R. K. Henderson, "8.25μm Pitch 66% Fill Factor Global Shared Well SPAD Image Sensor in 40nmCMOS FSI Technology," Int. Image Sensor Workshop (2017).

[4] J. M. Pavia, M. Wolf, and E. Charbon, "Measurement and modeling of microlenses fabricated on single-photon avalanche diode arrays for fill factor recovery," Opt. Express 22(4), 4202-4213 (2014).

[5] G. Intermite, A. McCarthy, R. E. Warburton, X. Ren, F. Villa, R. Lussana, A. J. Waddie, M. R. Taghizadeh, A. Tosi, F. Zappa, and G. S. Buller, "Fill-factor improvement of Si CMOS single-photon avalanche diode detector arrays by integration of diffractive microlens arrays," Opt. Express 23(26), 33777-33791 (2015).

[6] I. Gyongy, A. Davies, B. Gallinet, N. A. W. Dutton, R. Duncan, C. Rickman, R. K. Henderson, and P. A. Dalgarno, "Cylindrical microlensing for enhanced collection efficiency of small pixel SPAD arrays in single-molecule localisation microscopy," Opt. Express 26(3), 2280-2291 (2018).

[7] J. M. Pavia, M. Scandini, S. Lindner, M. Wolf, and E. Charbon, "A 1 × 400 Backside-Illuminated SPAD Sensor With 49.7 ps Resolution, 30 pJ/Sample TDCs Fabricated in 3D CMOS Technology for Near-Infrared Optical Tomography," IEEE J. Solid-State Circuits 50(10), 2406-2418 (2015).

[8] T. Al Abbas, N. A. W. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. K. Henderson, "Backside illuminated SPAD image sensor with 7.83 μm pitch in 3D-stacked CMOS technology," IEEE Int. Electron Devices Meeting, 811-814 (2016).

References

[9] A. R. Ximenes, P. Padmanabhan, M.-J. Lee, Y. Yamashita, D.-N. Yaung, and E. Charbon, "A Modular, Direct Time-of-Flight Depth Sensor in 45/65-nm 3-D-Stacked CMOS Technology," IEEE J. Solid-State Circuits 54(11), 3203-3214 (2019).

[10] S. W. Hutchings, N. Johnston. I. Gyongy, T. Al Abbas, N. A. W. Dutton, M. Tyler, S. Chan, J. Leach, and R. K. Henderson, "A Reconfigurable 3-D-Stacked SPAD Imager With In-Pixel Histogramming for Flash LIDAR or High-Speed Time-of-Flight Imaging," IEEE J. Solid-State Circuits 54(11), 2947-2956 (2019).

[11] S. Burri, C. Bruschini, and E. Charbon, "LinoSPAD: A Compact Linear SPAD Camera System with 64 FPGA-Based TDC Modules for Versatile 50 ps Resolution Time-Resolved Imaging," Instruments 1(1), 6 (2017).

[12] C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, "A 100-m Range 10-Frame/s 340×96-Pixel Time-of-Flight Depth Sensor in 0.18μm CMOS," IEEE J. Solid-State Circuits 48(2), 559-572 (2013).

[13] I. M. Antolovic, C. Bruschini, and E. Charbon, "Dynamic range extension for photon counting arrays," Opt. Express 26(17), 22234-22248 (2018).

[14] C. Veerappan, and E. Charbon, "A Low Dark Count p-i-n Diode Based SPAD in CMOS Technology," IEEE Trans. Electron Devices 63(1), 65-71 (2015).

[15] H. Finkelstein, M. J. Hsu, and S. C. Esener, "STI-Bounded Single-Photon Avalanche Diode in a Deep-Submicrometer CMOS Technology," IEEE Electron Device Letters 27(11), 887-889 (2006).

[16] S. Lindner, S. Pellegrini, Y. Henrion, B. Rae, M. Wolf, and E. Charbon, "A High-PDE, Backside-Illuminated SPAD in 65/40-nm 3D IC CMOS Pixel With Cascoded Passive Quenching and Active Recharge," IEEE Electron Device Letters 38(11), 1547-1550 (2017).

[17] F. Acerbi, G. Paternoster, A. Gola, N. Zorzi, and C. Piemonte, "Silicon photomultipliers and single-photon avalanche diodes with enhanced NIR detection efficiency at FBK," Nuclear Inst. and Methods in Physics Research, A (2017).

[18] I. Rech, A. Ingargiola, R. Spinelli, I. Labanca, S. Marangoni, M. Ghioni, and S. Cova, "Optical crosstalk in single photon avalanche diode arrays: a new complete model," Opt. Express 16(12), 8381-8394 (2008).

[19] S. Jahromi, and J. Kostamovaara, "Timing and probability of crosstalk in a dense CMOS SPAD array in pulsed TOF applications," Opt. Express 26(16), 20622-20632 (2018).

[20] L. Gasparini, B. Bessire, M. Unternährer, A. Stefanov, D. Boiko, M. Perenzoni, and D. Stoppa, "SUPERTWIN: towards 100kpixel CMOS quantum image sensors for quantum optics applications," Proc. of Quantum Sensing and Nano Electronics and Photonics XIV 10111 (2017).

[21] C. Niclass, A. Rochas, P.-A. Besse, and E. Charbon, "Design and Characterization of a CMOS 3-D Image Sensor Based on Single Photon Avalanche Diodes," J. Solid-State Circuits 40(9), 1847-1854 (2005).

[22] J. A. Richardson, E. A. G. Webster, L. A. Grant, and R. K. Henderson, "Scaleable Single-Photon Avalanche Diode Structures in Nanometer CMOS Technology," IEEE Trans. Electron Devices 58(7), 2028-2035 (2011).

[23] C. Niclass, H. Matsubara, M. Soga, M. Ohta, M. Ogawa, and T. Yamashita, "A NIR-Sensitivity-Enhanced Single-Photon Avalanche Diode in 0.18μm CMOS," Int. Image Sensor Workshop (2015).

[24] N. A. W. Dutton, T. Al Abbas, I. Gyongy, F. M. D. Rocca, and R. K. Henderson, "High Dynamic Range Imaging at the Quantum Limit with Single Photon Avalanche Diode-Based Image Sensors," Sensors 18(4), 1166 (2018).

[25] E. R. Fossum, J. Ma, S. Masoodian, L. Anzagira, and R. Zizza, "The Quanta Image Sensor: Every Photon Counts," Sensors 16(8), 1260 (2016).

[26] R. K. Henderson, E. A. G. Webster, R. Walker, J. A. Richardson, and L. A. Grant, "A 3×3, 5μm pitch, 3-transistor single photon avalanche diode array with integrated 11V bias generation in 90nm CMOS technology," IEEE Int. Electron Devices Meeting, 336-339 (2010).

4 Charge focusing SPADs for low-light imaging applications

A number of photon counting image sensors based on SPADs [1, 2], vertical avalanche photo-diodes (VAPDs) [3, 4], and CMOS-based quanta image sensors [5, 6, 7] have been proposed to achieve the shot-noise-limited imaging under low light illumination. Those devices are advantageous to capture images and videos under extremely low-light conditions owing to their single-photon sensitivity. The SPAD device operating in the Geiger-mode is more robust to process, voltage and temperature (PVT) variations in comparison with the linear-mode APD, and is a potential candidate for low cost image sensors in next-generation security cameras, as well as consumer cameras and camcorders. The digital nature of SPAD pixels eliminates some of the noise sources dominant in conventional CISs, such as kTC noise and random telegraph noise (RTN).

In recent years SPAD detectors have appeared in the market for ToF ranging and medical imaging [8, 9], as well as specialized scientific imaging. However, little attempt has been made to develop the SPAD-based 2D image sensors to compete with CCDs and CISs of image quality. Light detection efficiency is limited by relatively low fill factor and PDP in SPADs with respect to the CISs, where near-100% fill factor and over 70% maximum quantum efficiency are achieved. In addition, highly concentrated electric field at the p-n junction in the SPAD induces the excess leakage current, or dark counts, degrading the image quality particularly under long exposure or high temperature conditions. In this chapter, we propose an image sensor with a charge focusing SPAD to further reduce DCR and temporal noise so as to detect a single-photon signal at a standard video frame rate. On-chip color filter array is implemented on the sensor to demonstrate color imaging under photon-starved conditions.

4.1 Principle of charge focusing SPAD

One of the fundamental issues in SPAD device is a native tradeoff between PDE and DCR. As discussed in the previous chapters, both PDE and DCR are positively correlated with the active area: a region where high electric field is formed to induce avalanche multiplication. Smaller active area is desirable for suppressing DCR while sacrificing the PDE, and vice versa. Fig. 4.1(a)
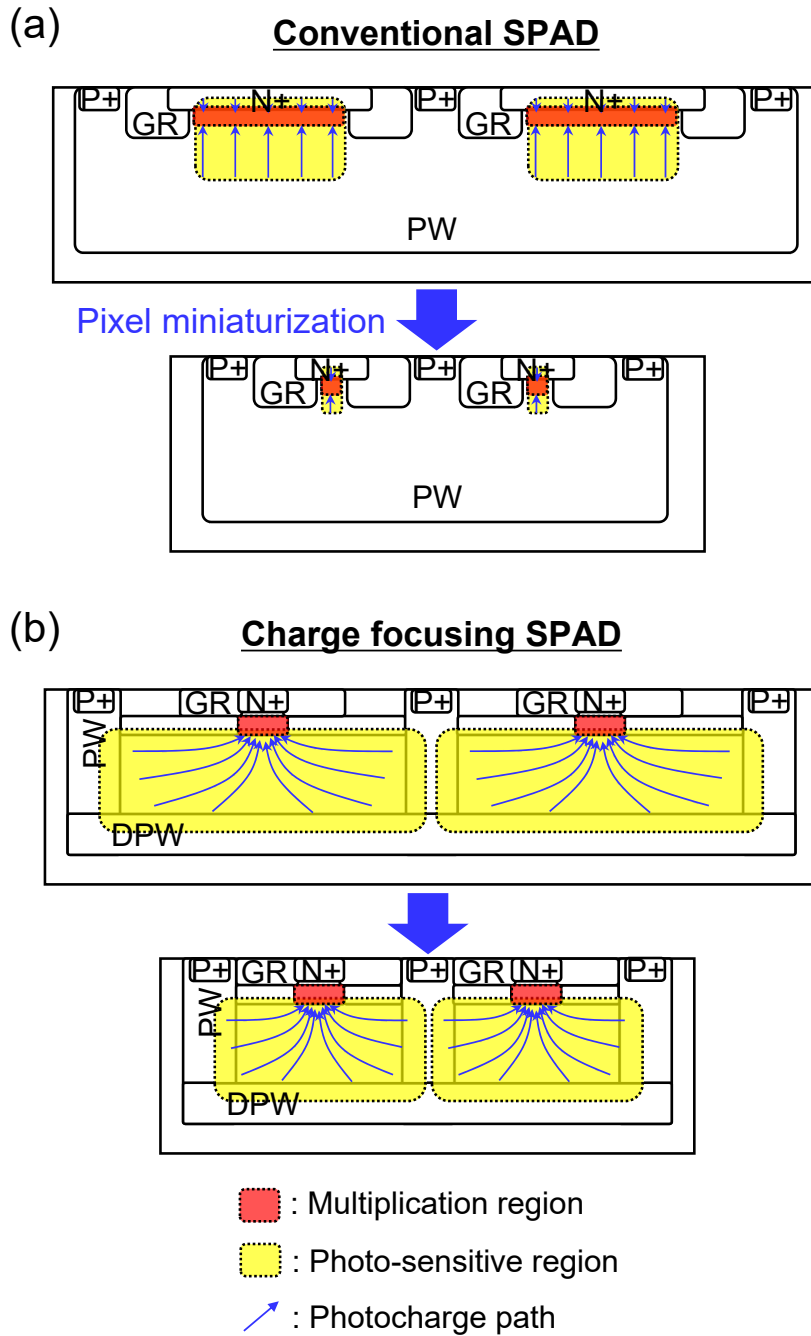
Figure 4.1 – Principle of charge focusing SPAD. (a) Schematic cross-section of conventional SPAD arrays with large pixel pitch (top) and small pixel pitch (bottom). (b) Schematic cross-section of charge focusing SPAD arrays with large pixel pitch (top) and small pixel pitch (bottom). Red and yellow regions depict multiplication region and photo-sensitive region, respectively. Blue arrows show photocharge paths.

shows a schematic cross-sections of conventional SPAD arrays. The red area indicates the multiplication region, while the yellow area indicates the photo-sensitive region. The blue arrows show photocharge paths determined by electrostatic potential distribution. The top panel in the figure shows the cross-section for larger pixel size. Photoelectrons generated inside or underneath the multiplication region are attracted to the surface N+, and trigger avalanche multiplication events. In contrast, photocharges generated underneath or external to the GR are attracted to the GR region and reach N+ without passing the multiplication region. Those charges are not detectable in this device configuration. A similar analysis can be applied to photoholes; the SPAD can only detect photoholes generated just above the multiplication region. As a result, the photo-sensitive area from the top view is almost identical to the active area. When considering the pixel miniaturization, the GR width and pixel isolation region do not scale with the pixel size due to process limitations and the necessity to suppress premature edge breakdown. As shown in the bottom panel of the figure, the pixel miniaturization leads to significant reduction of the active area, and hence fill factor and PDE are severely degraded. An on-chip microlens optically concentrates incident photons to the active area and can recover part of the fill factor [10, 11, 12]. However, there are several limitations. First, the microlens is less effective when the active area is too small, or when the f-number of the objective lens is relatively small, e.g. 5.6 or less. Second, the incident light angle dependence of PDE can be increased by the microlens, thereby degrading optical nonideality called vignetting. Third, misalignment and process variation of the microlens can degrade the photoresponse nonuniformity (PRNU).

Here we propose a novel charge focusing SPAD to overcome the basic tradeoff between PDE and DCR, and potentially break the lower bound of SPAD pixel size determined by PDE degradation. Fig. 4.1(b) shows cross-sectional views of the proposed charge focusing SPAD arrays. The top panel in the figure, again, shows the cross-section for larger pixel. The essence of the charge focusing SPAD lies in the electrostatic potential design to minimize the multiplication region, while maximizing the photo-sensitive region. The area of the surface N+ is designed to be as small as possible, so that the electric field is concentrated just below the miniaturized N+ region. The N+ region is surrounded by the GR region to avoid premature edge breakdown. A buried p-type layer below the GR and a p-type isolation (PW) are designed to create lateral electric field to collect photoelectrons into the central multiplication region. A concept of detecting photocharges outside its active area using lateral avalanche propagation has been proposed in [13], where the PDE is increased at the expense of the DCR performance. In our approach, the optimized layout and doping profile enhance the PDE by efficiently collecting the photoelectrons from an electrostatically confined deep photo-conversion region to the multiplication region while further suppressing the DCR. The bottom panel of the figure indicates that, in contrast to conventional SPADs, the charge focusing SPAD with proper potential design enables pixel miniaturization without sacrificing the PDE; theoretically, near-100% fill factor can be achieved at 2 μm-pitch or below. Thus, the charge focusing SPAD can break the lower limit of SPAD pixel pitch while achieving ultralow DCR.

In addition to the high PDE, low DCR, and miniaturization possibilities, the charge focusing

SPAD has a number of additional advantages. First, the extremely small multiplication region reduces the probability of having defects or traps around the high electric field region. This significantly reduces the population of hot pixels. Second, the miniaturized area of the main p-n junction results in small parasitic capacitance for the cathode. This leads to the reduction of afterpulsing probability, crosstalk probability, dead time, and SPAD-induced power consumption, all of which are critical for megapixel implementation of SPADs for a wide range of applications. One concern of the charge focusing SPAD is a reduced PDP for shorter wavelengths in the FSI configuration; photoelectrons generated by UV to blue light above the buried p-type layer are absorbed by the N+ region without reaching the multiplication region. The BSI integration will circumvent this issue by exposing the photo-conversion region directly to the light incident surface. Another potential question to the charge focusing SPAD is timing jitter performance. This will be discussed later in this chapter.

Note that the concept of charge focusing has been investigated, although in a different context, in high-speed CMOS image sensors [14]. Just recently, a similar concept is introduced in SPAD [15], where the need of a static bias current fundamentally limits its scalability. Moreover, it is not clear what effect it could have on overall power consumption. In addition, the measured DCR is the order of 1 Mcps at room temperature, which needs to be greatly improved for practical implementation.

4.2 Proof-of-concept sensor design

To verify the feasibility of the proposed solution, a proof-of-concept SPAD image sensor was developed. Fig. 4.2 illustrates a block diagram of the SPAD image sensor. The chip comprises a 128×128 SPAD pixel array, a vertical shift register, a horizontal shift register and a multiplexer. Each SPAD pixel consists of a Geiger-mode APD with V_B = 28.1 V and V_{ex} = 2.5 V, a passive quenching pMOS transistor, an inverter, a 16-bit digital counter and output selectors. The pixel pitch is 84 μm, and the photo-sensitive area is 200.96 μm^2. The larger pixel size and the lower fill factor with respect to the typical CIS are due to the FSI monolithic integration, and can potentially be compensated by employing 3D-stacked BSI configuration. The sensor is fabricated in 0.18 μm 1P3M CIS process with 3.3 V-analog and 1.8 V-digital power supply. The on-chip Bayer color filter enables the RGB color imaging down to the single-photon signal level. Fig. 4.3 is the photograph of the proof-of-concept sensor chip.

Fig. 4.4 shows the simulated electrostatic potential and the electric field distributions, and the cross-sectional scanning electron microscope (SEM) image of the proof-of-concept SPAD pixel. The simulated electrostatic potential in Fig. 4.4(a) shows a high gradient at the center in horizontal direction to induce avalanche multiplication. The lateral gradient seen in the deeper part of the photodiode is designed for charge collection. The deep photo-conversion region is mostly surrounded by a potential barrier to suppress the loss of the signal electrons. Fig. 4.4(b) is the magnified view of the corresponding electric field distribution. The central part of the region underneath the surface N+ region shows the highest electric field, while the lateral electric field around the N+ is well-suppressed by the guard ring. Fig. 4.4(c) is the
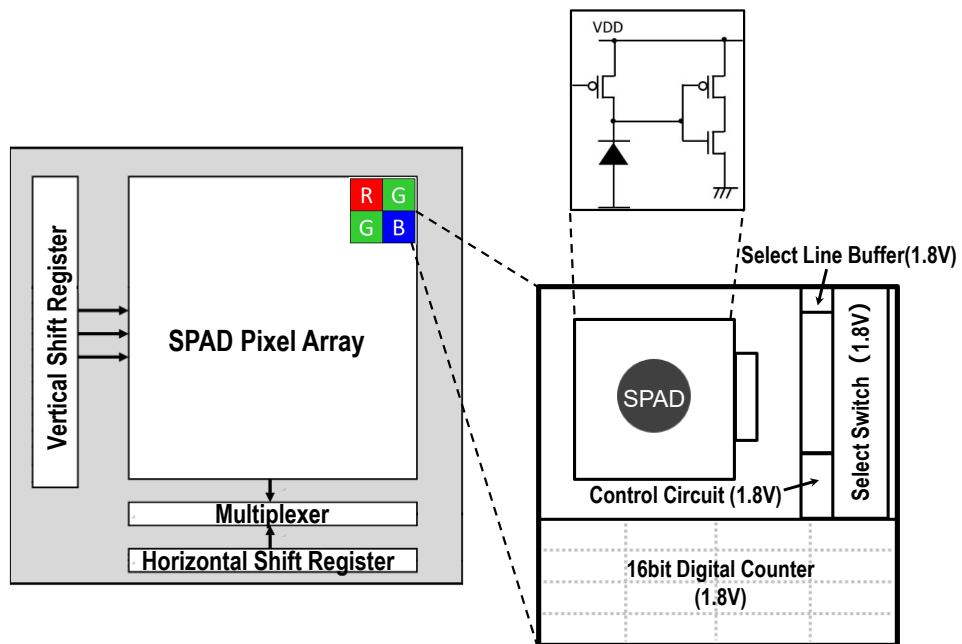
Figure 4.2 – Proof-of-concept sensor architecture. Sensor block diagram is shown in left, and pixel circuit diagram is shown in right.
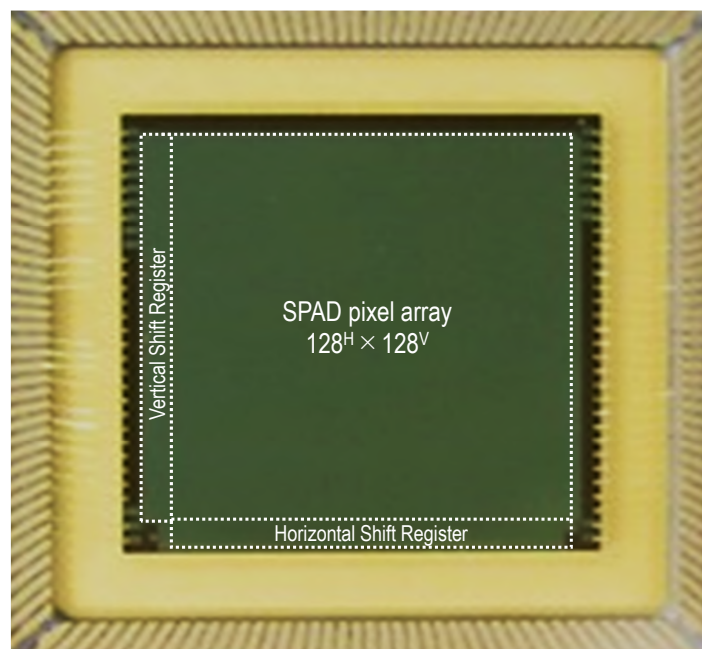
Figure 4.3 – Chip micrograph of proof-of-concept sensor.
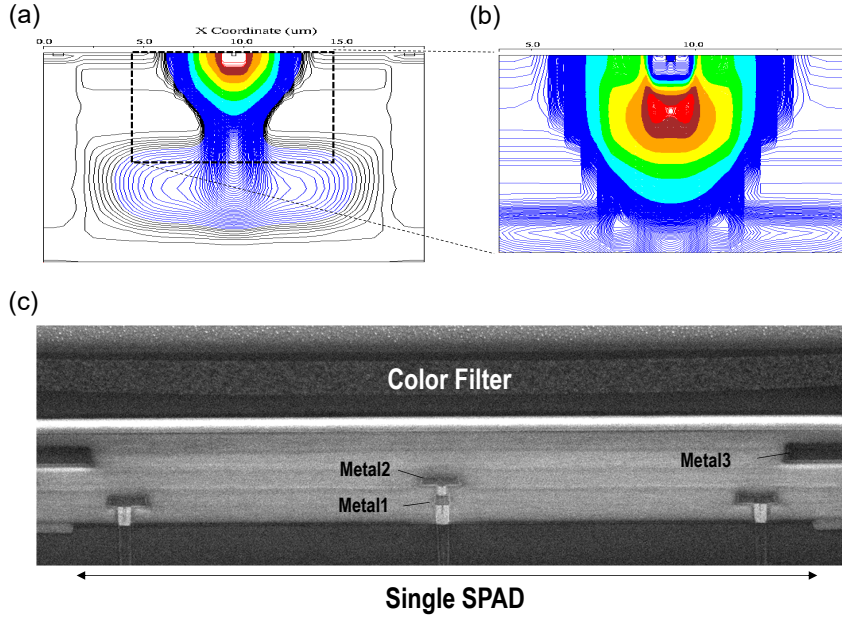
Figure 4.4 – Device simulation and cross-sectional SEM image. (a) 2D cross-section of electrostatic potential profile. Red and blue regions correspond to higher and lower potential, respectively. (b) 2D cross-section of electric field profile. Red and blue regions correspond to higher and lower electric field, respectively. (c) SEM image of pixel cross-section.

cross-sectional SEM image of the fabricated SPAD pixel. Central metal and via are connected to the cathode, while the right and left metal and via supply the anode voltage. Color filter is formed above the pixel region.

4.3 Experimental results

4.3.1 Photoresponse characteristics

Fig. 4.5 shows a schematic diagram of in-pixel output clipping circuit and the measured photoresponse characteristics of the SPAD sensors. In the passive quenching SPAD, output photon counts saturate and then drop with increasing the illumination power due to the increased count loss of photons impinging during the dead time [16]. In 2D imaging, this non-monotonic behavior of the photoresponse causes an undesirable phenomenon called black spot, where extremely bright objects are clipped to black in a captured image. The black spot observed in CIS, though caused by a different mechanism from that in SPAD, has been considered as a critical issue in consumer camera applications, and various solutions have been proposed [17, 18]. In the SPAD sensor, the issue can be avoided by introducing active recharge, which ensures that a maximum photon flux is reached and maintained at a saturation level. We developed a novel in-pixel output clipping circuit to address this issue with minimal side effect to the image quality. Fig. 4.5(a) and (b) are the pixel circuit diagram and
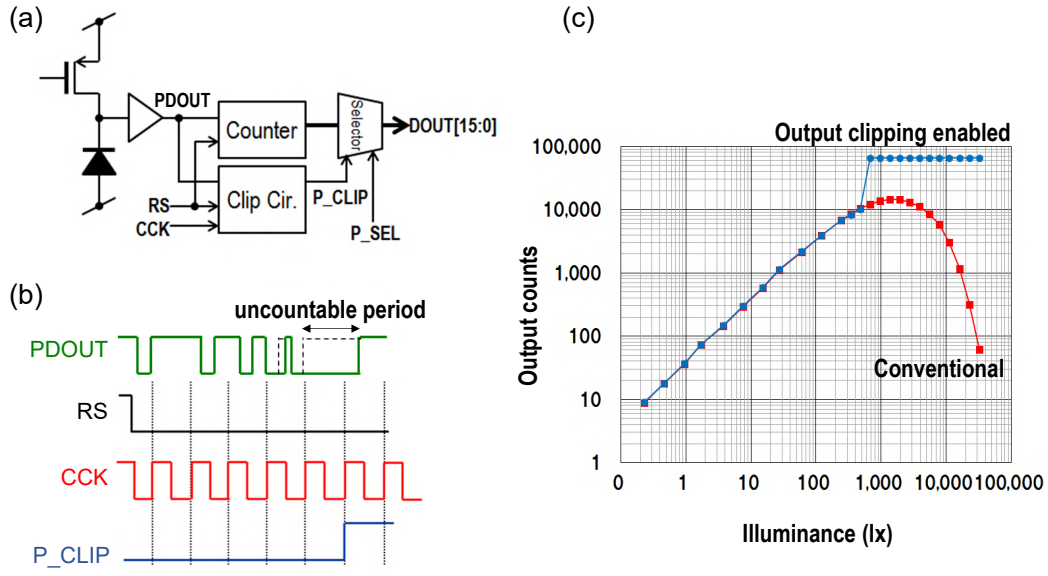
Figure 4.5 – Principle and measured results for pixel-wise output clipping circuit. (a) Pixel circuit diagram. (b) Timing chart. (c) Measured photoresponse characteristics, with and without the output clipping function.

the corresponding timing diagram, respectively. A pulsed output signal from SPAD (PDOUT) is compared with a reference clock (CCK) to measure the duration of ON-state of the photon counting. Under low light condition, photon counting pulses are well-separated in time domain, and the duration of ON-state for each pulse is fixed at a dead time. Under high light condition, photon counting pulses are observed more frequently, and some of the adjacent pulses can be merged into a single long pulse. Hence, pixel-wise measurement of the duration of ON-state gives supportive information on whether the light intensity for the pixel is higher than a certain threshold or not. When PDOUT is kept ON for N consecutive clock cycles of CCK (N=2 in the figure), the pixel recognizes that the light intensity is too high to perform reliable photon counting, and a signal P_{CLIP} is activated to clip the 16-bit pixel output (DOUT) to all high state corresponding 65,535 counts, so as to indicate that the pixel is "saturated". As shown in Fig. 4.5(c), the output signal is clipped to the saturation level when the incident light power exceeds a certain threshold. When the incident light power is below the threshold, the pixel performs photon counting based on passive quenching. The measured photoresponse characteristics shows monotonic behavior with the output clipping enabled (shown in blue). This result suggests that the proposed output clipping circuit is effective to suppress the black spot.

4.3.2 DCR and temporal noise

Fig. 4.6(a) shows the measured cumulative histograms of hot pixels at 60 °C for the charge focusing SPAD and a reference SPAD with conventional n+/p-well junction. For fair comparison, diameters of the inner circle for isolation well are designed to be equivalent for both structures. The population of hot pixels in the charge focusing SPAD is one to two orders of magnitude lower than that of the reference SPAD device due to the reduced volume of the high-electric-field region. Fig. 4.6(b) shows the temperature dependence of the measured median DCR for the charge focusing SPAD. At room temperature, the measured median DCR for our SPAD sensor is 3 cps. The corresponding DCR density, defined as the median DCR divided by photo-sensitive area, is 0.015 cps/μm^2. The DCR reduces by 100× at T = -20 °C, corresponding to a DCR density of 0.00015 cps/μm^2. The activation energy of the median DCR is 1 eV over the temperature range from -20 °C to 60 °C, indicating that the diffusion current dominates over the trap-assisted generation and tunneling for the majority of the pixels [19]. Unlike CIS where kTC noise and RTN are the major noise factors, the temporal noise in the SPAD image sensor is determined solely by the dark count shot noise and is strongly temperature dependent even with short exposure. Fig. 4.7(a) shows the temporal noise as a function of exposure time. The temporal noise N is calculated as:

$$N = \sqrt{DCR \cdot t_{exp}}, \tag{4.1}$$

where t_{exp} is the exposure time. The unit of the temporal noise, e$^-_{rms}$, depicts the input referred noise for incident photons (or photo-excited electrons), which is introduced for systematic noise comparison with that of CISs. The measured temporal noise at room temperature is 0.16 e$^-_{rms}$ at t_{exp} = 1/120 second, and is further reduced to 0.05 e$^-_{rms}$ at t_{exp} = 1/1000 second. At t_{exp} = 1 second, the temporal noise is still below 2 e$^-_{rms}$, comparable to or lower than the typical temporal noise in CISs. Fig. 4.7(b) shows the temperature dependence of the temporal noise at t_{exp} = 1/120 second. In the measured temperature range, the lowest temporal noise of 0.016 e$^-_{rms}$ is obtained at -20 °C, while the exponential behavior without saturation in lower temperature implies the further reduction towards the temperature below -20 °C.

4.3.3 PDP and sensitivity

Fig. 4.8(a) shows the measured PDP as a function of incident light wavelength for the charge focusing SPAD at V_{ex} = 2.5 V. The maximum PDP of 40% is obtained at λ = 550 nm. The PDP at λ = 900 nm is measured as 6.4%, which could be further enhanced by introducing deeper p-well. Relatively low PDP for λ < 500 nm stems from the insensitive region above the photo-sensitive region. Again, dramatic improvement of the UV to blue light sensitivity is expected in BSI configurations. Fig. 4.8(b) shows the measured normalized sensitivity as a function of photodiode size (pixel pitch) for the conventional and the charge focusing SPADs. The normalized sensitivity is defined, analogously to the sensitivity for CIS, as a number of
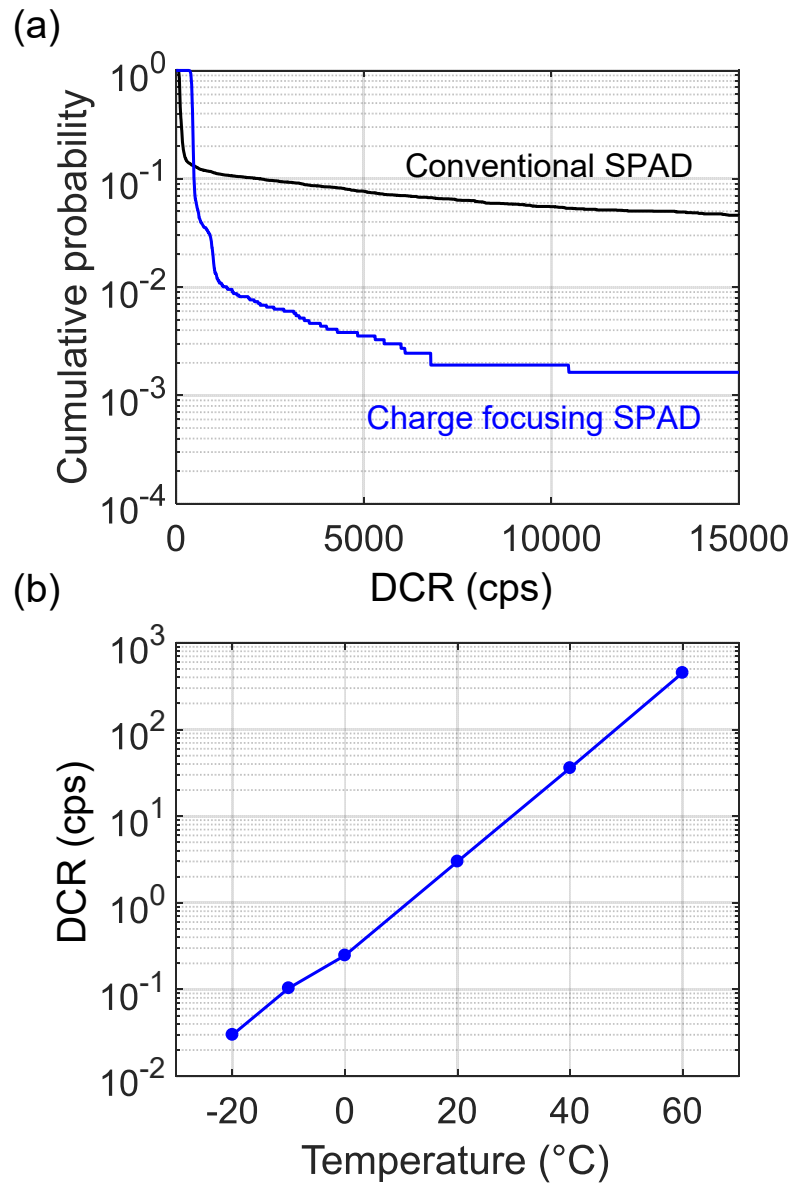
Figure 4.6 – Measured DCR characteristics. (a) Cumulative histogram of DCR for conventional SPAD and charge focusing SPAD. (b) Temperature dependence of measured DCR for charge focusing SPAD.

(a)

(b)
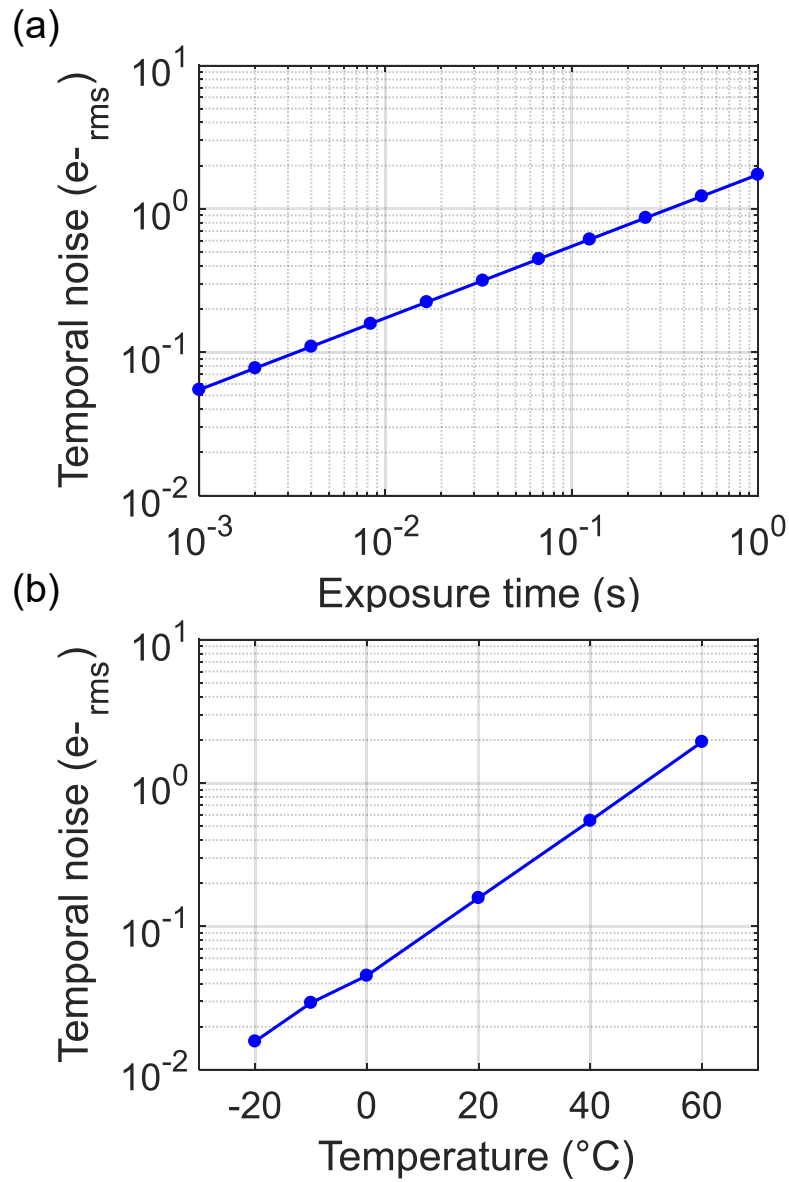
Figure 4.7 – Measured temporal noise characteristics. (a) Temporal noise at 1/120 second exposure time as a function of exposure time. (b) Temporal noise as a function of temperature at 1/120 second.

(a)

(b)
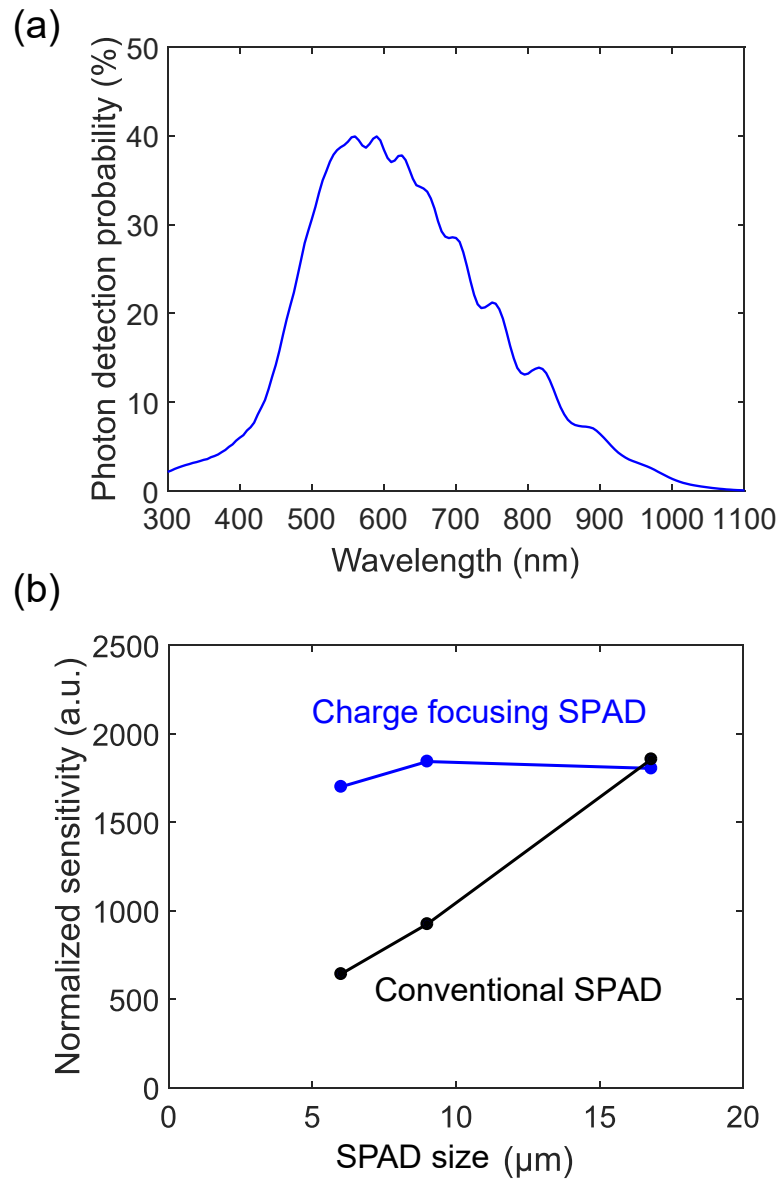
Figure 4.8 – Measured PDP and sensitivity characteristics. (a) PDP as a function of wavelength. (b) Normalized sensitivity under 2850 K for conventional and charge focusing SPADs.

photoelectrons detected in a unit photo-sensitive area (1 μm^2) during 1 second under 2850 K lamp through an IR-cut filter, with illuminance of 1 lux. The conventional SPAD shows a degradation of normalized sensitivity for smaller pixel size due to the border effect [20]. In contrast, the normalized sensitivity for charge focusing SPAD shows constant behavior up to a pitch of 6 μm, indicating the efficient charge collection from the edge of the pixel. The result demonstrates that the proposed charge focusing SPAD has a potential advantage over conventional SPAD devices for pixel miniaturization below 10 μm-pitch.

4.3.4   2D imaging

Fig. 4.9 shows captured intensity images of a moving object (fan) and a still color object based on the pixel-parallel photon counting. In Fig. 4.9(a), the SPAD sensor is operated in two readout modes: rolling shutter (RS) and global shutter (GS). The rolling shutter distortion observed in the RS mode is suppressed in the GS mode. In CIS-based GS sensors, significant portion of a GS pixel area is devoted for a signal storage node, limiting its sensitivity and full well capacity. In contrast, the SPAD sensor features in-pixel digital memories which do not consume extra pixel area for GS and are highly immune to parasitic light sensitivity. Fig. 4.9(b) shows color images acquired under sufficient light condition above 100 lux (left) and extremely low-light conditions (center and right). For all the color images, raw signals in a 2×2 pixel block are used to generate a single RGB pixel signal with proper tuning of color tone. This non-interpolating method is helpful to visualize the color imaging capabilities under low light conditions with minimal influence of post-processing, while the image resolution is reduced by a half from the raw data. The color tone is well reproduced to recognize the objects even under the signal level corresponding to or below a single photon per pixel per frame, indicating the shot-noise-limited color imaging. No noticeable impact of DSNU, PRNU, hot pixels, afterpulsing and crosstalk is observed in the obtained images. Those images are compared with the captured color images by a high sensitivity CIS with temporal noise of 2.1 e$^-_{rms}$ (Fig. 4.9(c)). The image contrast is tuned so as to avoid signal clipping to white and black caused by the dark noise. When signal levels, image resolution and field-of-view are tuned for fair comparison, the image quality for the SPAD image sensor is much better than that of the CIS, where the letters (Canon) on the target and RGB colors of the square boxes are hard to recognize due to the excessive noise. The results are reasonable because the signal level of 0.5 or 1 photon is sufficiently higher than the temporal noise level (rms) in the charge focusing SPAD, but considerably lower than that in the CIS.

4.4   State-of-the-art comparison

Table 4.1 presents a summary of the chip specifications and a comparative table of the state-of-the-art photon counting image sensors. The DCR per unit of photo-sensitive area at room temperature is 0.015 cps/μm^2, considerably lower than the prior art based on the SPAD [21, 22] and CIS [6]. The temporal noise of 0.16 e$^-_{rms}$ ensures the single-photon-resolved imaging

(a)

Rolling shutter   Global shutter

(b)

RGB Color SPAD

(c)

High sensitivity CMOS imager

Highlight   Signal = 1 photon   Signal = 0.5 photon
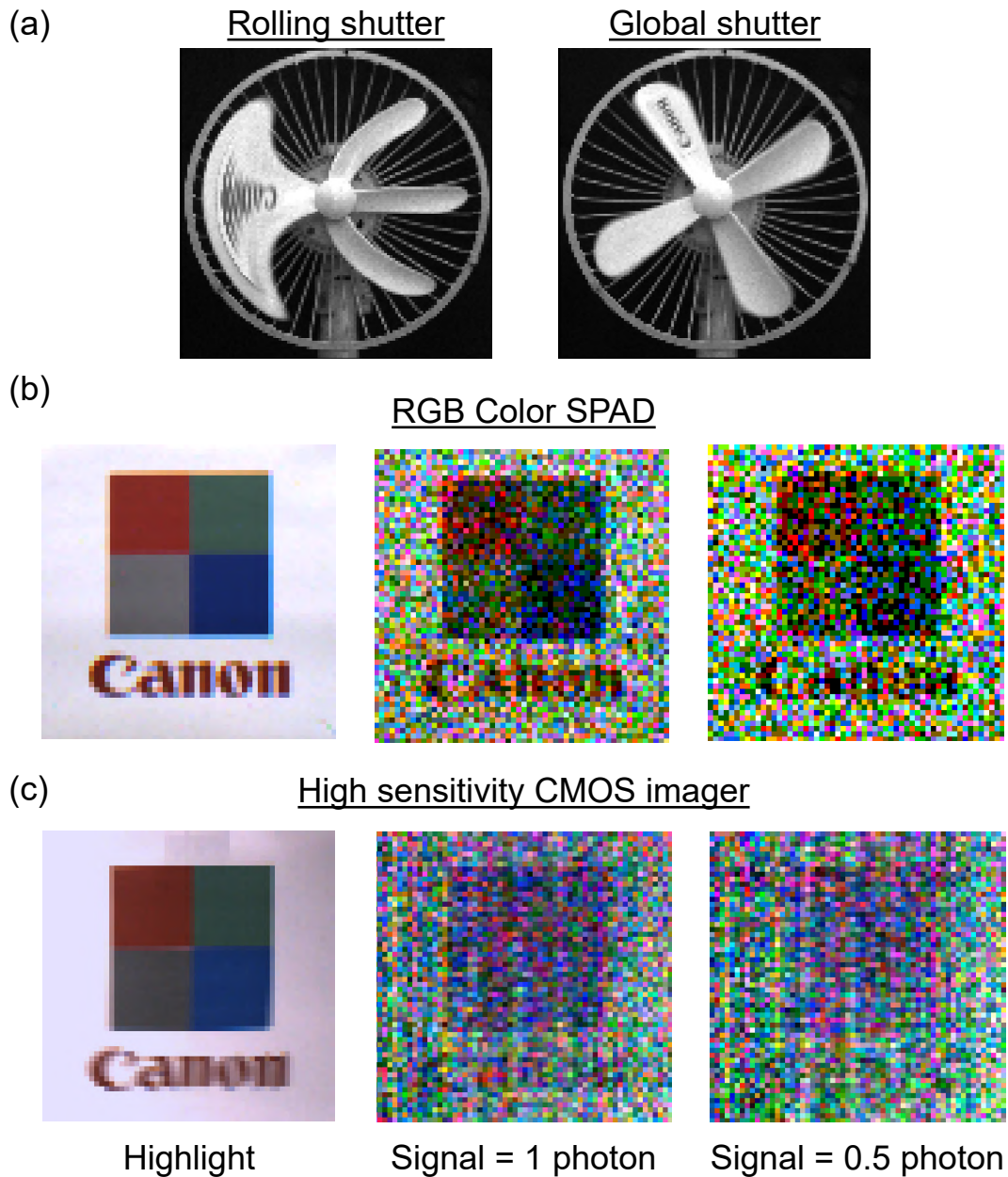
Figure 4.9 – 2D images captured with the developed monochrome and RGB-color SPADs. (a) 2D intensity images captured with the monochrome SPAD sensor in RS mode (left) and GS mode (right). (b) 2D intensity images with the RGB-color SPAD sensor under sufficient light above 100 lux (left), signal level corresponding to 1 photon for white part (center), and signal level corresponding to 0.5 photon for white part (right). (c) 2D intensity images with the RGB-color CIS under equivalent input signal level to (b).

at video-rate. The measured afterpulsing probability is less than 1% at 5.3 ns dead time (not shown in the table). We note that photon-counting-based GS imaging is one of the striking features of the SPAD-based approach. Another advantage of our approach is the low PRNU; in the CMOS-based approach, the PRNU is affected by conversion-gain variation. The variation could be degraded further when increasing the absolute value of the conversion gain to suppress the temporal noise, potentially limiting the image quality under mid- to high-illumination. In contrast, the major source of the PRNU in the SPAD image sensor is a pixel-to-pixel variation of avalanche triggering probability. This can be suppressed by increasing the excess bias further. With optimized pixel circuit and biasing, the PRNU in our device is expected to be below 1%.

Table 4.1 – State-of-the-art comparison of performance and specifications in photon counting image sensors.

	This work	[2]	[3]	[7]	[5]	[23]
Type	SPAD	SPAD	VAPD	CMOS	CMOS	CMOS
Process technology	180 nm FSI	180 nm FSI	65 nm FSI	45/60 nm BSI	110 nm FSI	180 nm FSI
Pixel pitch (μm)	84	16.38	6.0	1.1	11.2×5.6	5.5
Pixel array size (H×V)	128×128	512×512	400×400	1,024×1,024	312×512	360×1,680
Pixel output (bit)	16	1	-	3	-	-
Operation voltage (V)	3.3/-26.5	-	3.3/-29	-	3/25	-
Photodiode area (μm^2)	200.96	28.2	25.2	1.21	-	14.76
Max. PDP or QE (%)	40	50	-	80	-	-
DCR density (cps/μm^2)	0.015	0.26	2.77	<0.16	-	14.76
Temporal noise at RT (e$^-_{rms}$)	0.16	-	-	0.21	0.27	0.46
DSNU at RT (e$^-_{rms}$)	0.2	-	-	-	-	-
PRNU (%)	1.7	-	-	>2.6	-	-
Electronic shutter	GS/RS	GS/RS	RS	RS	RS	RS
On-chip color filter	w/	w/o	w/o	w/o	w/o	w/o

4.5 Conclusion

In this chapter we proposed a novel charge focusing SPAD to overcome the native tradeoff between PDE and DCR in small pixels, and to push the limit of pixel miniaturization. We demonstrated the feasibility of the proposed device in the proof-of-concept 128×128 2D image sensor for low-light imaging applications. We observed the best-in-class performance in DCR, temporal noise, DSNU, PRNU, and afterpulsing probability. The performance is confirmed in SPADs with a pitch as low as 6 μm. Relatively larger pixel pitch and lower PDP

in UV to blue range originate from the FSI monolithic approach, however we expect that it will be significantly improved by introducing the BSI 3D-stacked configurations combined with advanced logic processes. The proof-of-concept image sensor featured single-photon-sensitive GS function and black spot-free image capturing function, as well as RGB color imaging under photon-starved conditions.

We emphasize that while the charge focusing SPAD concept is studied in this chapter for low-light imaging applications, this device can be applied to any other applications where high PDE or low DCR is required. In addition, the unique features of near-100% fill factor and high PDP up to 80% in miniaturized pixels ($< 10\ \mu$m) will realize unrivaled sensing capabilities in the compact sensor size.

One potential question of the charge focusing SPAD is the timing jitter; lateral charge focusing process could degrade the temporal variation of photon detection. Qualitatively, smaller SPAD size can reduce the travel path length of photocharges, and is desirable for improved timing jitter. In a preliminary simulation for 6 μm SPAD size, we confirmed that careful design of 3D electrostatic potential distribution achieves a timing jitter component from lateral and vertical drift-diffusion of photocharges below 50 ps FWHM. This indicates that the miniaturized charge focusing SPAD pixels can achieve timing jitter comparable to that of conventional SPADs, and can be applied to a wide variety of time-resolved applications requiring sub-100 ps timing resolution.

References

[1] N. A. W. Dutton, I. Gyongy, L. Parmesan, S. Gnecchi, N. Calder, B. R. Rae, S. Pellegrini, L. A. Grant, and R. K. Henderson, "A SPAD-Based QVGA Image Sensor for Single-Photon Counting and Quanta Imaging," IEEE Trans. Electron Devices 63(1), 189-196 (2016).

[2] A. C. Ulku, C. Bruschini, I. M. Antolovic, Y. Kuo, R. Ankri, S. Weiss, X. Michalet, and E. Charbon, "A 512×512 SPAD Image Sensor with Integrated Gating for Widefield FLIM," IEEE J. Selected Topics in Quantum Electronics 25(1), 1-12 (2018).

[3] Y. Hirose, S. Koyama, T. Okino, A. Inoue, S. Saito, Y. Nose, M. Ishii, S. Yamahira, S. Kasuga, M. Mori, T. Kabe, K. Nakanishi, M. Usuda, A. Odagawa, and T. Tanaka, "A 400×400-Pixel 6μm-Pitch Vertical Avalanche Photodiodes CMOS Image Sensor Based on 150ps-Fast Capacitive Relaxation Quenching in Geiger Mode for Synthesis of Arbitrary Gain Images," IEEE Int. Solid-State Circuits Conference (2019).

[4] T. Okino, S. Yamada, Y. Sakata, S. Kasuga, M. Takemoto, Y. Nose, H. Koshida, M. Tamaru, Y. Sugiura, S. Saito, S. Koyama, M. Mori, Y. Hirose, M. Sawada, A. Odagawa, and T. Tanaka, "A 1200×900 6μm 450fps Geiger-Mode Vertical Avalanche Photodiodes CMOS Image Sensor for a 250m Time-of-Flight Ranging System Using Direct-Indirect-Mixed Frame Synthesis with Configurable-Depth-Resolution Down to 10cm, " IEEE Int. Solid-State Circuits Conference (2020).

[5] M.-W. Seo, S. Kawahito, K. Kagawa, and K. Yasutomi, "A 0.27e-rms Read Noise 220-μV/e-Conversion Gain Reset-Gate-Less CMOS Image Sensor With 0.11-μm CIS Process," IEEE Electron Devices Letters 36(12), 1344-1347 (2015).

[6] E. R. Fossum, J. Ma, S. Masoodian, L. Anzagira, and R. Zizza, "The quanta image sensor: every photon counts," Sensors 16, 1260 (2016).

[7] J. Ma, S. Masoodian, D. A. Starkey, and E. R. Fossum, "Photon-number-resolving megapixel image sensor at room temperature without avalanche gain," Optica 4(12), 1474-1481 (2017).

[8] S. Pellegrini, and B. Rae, "Fully industrialised single photon avalanche diodes," Proc. SPIE 10212, Advanced Photon Counting Techniques XI, 102120D (2017).

References

[9] C. Jackson, K. O'Neill, L. Wall, B. McGarvey, "High-volume silicon photomultiplier production, performance, and reliability," Optical Engineering 53(8), 081909 (2014).

[10] J. M. Pavia, M. Wolf, and E. Charbon, "Measurement and modeling of microlenses fabricated on single-photon avalanche diode arrays for fill factor recovery," Opt. Express 22(4), 4202-4213 (2014).

[11] I. Gyongy, A. Davies, B. Gallinet, N. A. W. Dutton, R. Duncan, C. Rickman, R. K. Henderson, and P. A. Dalgarno, "Cylindrical microlensing for enhanced collection efficiency of small pixel SPAD arrays in single-molecule localisation microscopy," Opt. Express 26(3), 2280-2291 (2018).

[12] I. M. Antolovic, A.C. Ulku, E. Kizilkan, S. Lindner, F. Zanella, R. Ferrini, M. Schnieper, E. Charbon, and C. Bruschini, "Optical-stack optimization for improved SPAD photon detection efficiency," in Proc. SPIE, Quantum Sensing and Nano Electronics and Photonics XVI, San Francisco, CA, USA (2019).

[13] C. Veerappan, and E. Charbon, "Silicon Integrated Electrical Micro-lens for CMOS SPADs Based on Avalanche Propagation Phenomenon," Int. Image Sensor Workshop (2013).

[14] V. T. S. Dao, N. Ngo, A. Q. Nguyen, K. Morimoto, K. Shimonoura, P. Goetschalckx, L. Haspeslagh, P. De Moor, K. Takehara, and T. G. Etoh, "An Image Signal Accumulation Multi-Collection-Gate Image Sensor Operating at 25 Mfps with 32 × 32 Pixels and 1220 In-Pixel Frame Memory," Sensors 18, 3112 (2018).

[15] G. Jegannathan, H. Ingelberts, and M. Kujik, "Current-Assisted Single Photon Avalanche Diode (CASPAD) Fabricated in 350 nm Conventional CMOS," Appl. Sci. 10, 2155 (2020).

[16] A. Eisele, R. Henderson, B. Schmidtke, T. Funk, L. Grant, J. Richardson, and W. Freud, "185 MHz Count Rate, 139 dB Dynamic Range Single-Photon Avalanche Diode with Active Quenching Circuit in 130 nm CMOS Technology," Int. Image Sensor Workshop (2011).

[17] M. Murakami, et al., "2.8μm-Pixel Image Sensor vMaicovicon," in Proc. of CCD & AIS Workshop (2005).

[18] S. -H. Lim, "Column analog-to-digital converter of a CMOS image sensor for preventing a sun black effect," U.S. Patent 7,218,260 (2007).

[19] N. Teranishi, "Dark current and white blemish in image sensors," IEEE Int. Symposium on VLSI technology, Systems and Applications (2013).

[20] F. Acerbi, G. Paternoster, A. Gola, N. Zorzi, and C. Piemonte, "Silicon photomultipliers and single-photon avalanche diodes with enhanced NIR detection efficiency at FBK," Nuclear Inst. And Methods in Physics Research, A (2017).

[21] D. Bronzi, F. Villa, S. Bellisai, B. Markovic, S. Tisa, A. Tosi, F. Zappa, S. Weyers, D. Durini, W. Brockherde, and U. Paschen., "Low-noise and large-area CMOS SPADs with timing

response free from slow tails," in Proc. Eur. Solid-State Device Res. Conf. (ESSDERC), 230–233 (2012).

[22] C. Veerappan, and E. Charbon, "A Low Dark Count p-i-n Diode Based SPAD in CMOS Technology," IEEE Trans. Electron Devices 63(1), 65-71 (2015).

[23] S. Wakashima, F. Kusuhara, R. Kuroda, and S. Sugawa, "A Linear Response Single Exposure CMOS Image Sensor with 0.5e- Readout Noise and 76ke- Full Well Capacity," IEEE Int. Symposium on VLSI Circuits (2015).

5 45 nm/22 nm CMOS BSI 3D-stacked SPAD image sensors

In 3D stacking, 2 chips are connected, usually, at wafer level, through hybrid bonding or equivalent technologies. This solution enables complete separation of SPADs, usually integrated in the top tier, from other circuit components, usually fabricated in the bottom tier. A photodiode wafer and a circuit wafer can be independently fabricated by dedicated processes, thus achieving high PDE and low noise SPAD arrays with compact, low power and functional pixel circuits. This chapter introduces ongoing developments for 3D-stacked BSI SPAD image sensors. SPAD sensors with several different architectures are all designed and fabricated based on 3D-integration technology (45 nm CIS process for PD and 22 nm logic process for circuit).

5.1 3D-stacked SPAD sensors for advanced time-gating approaches

In ToF ranging applications, time-resolved SPAD pixels based on TDC have been intensively studied so far [1, 2, 3, 4, 5]. The advantages of TDC-based d-ToF methods are: fast acqui-sition of object distance thanks to single-shot measurement, high depth resolution up to sub-centimeter, immunity to ambient light, and the capability of detecting multiple reflec-tions. On the other hand, this method has a difficulty in scaling pixel pitch and expanding spatial resolution due to area-consuming and power-hungry TDC circuits. Furthermore, recording and processing all the time stamps of incident photons involve large volumes of output data, which in particular cases severely limit frame rate. Alternatively, time-gating ap-proaches are a potential candidate for scalable SPAD-ToF sensor. Time-gated SPAD pixels with less than 15 pixel transistors have been demonstrated in a monolithic approach [6, 7, 8, 9, 10]. In contrast to TDC-based approaches where time stamps need to be acquired for every photon, time-gated SPAD sensors directly count the number of photons arriving to each pixel in a predetermined time-gating window. Thus, in-pixel analog or digital counters help reducing the total data rate from the pixel array.

Recent advances in 3D-stacked processes have enabled implementation of more sophisticated pixel circuits. A time-gated SPAD sensor, combined with 3D stacking, has been reported in

[12], where an in-pixel multi-bit reconfigurable digital counter achieves intensity imaging mode and i-ToF ranging mode. However, little attempts have been made to enhance the functionality of time-gated SPADs; one direction is to employ an arbitrarily coded time-gating window instead of a single pulsed window. Another direction is to optimize the time-gating window in pixel-wise manner. In the following sections, we propose several time-gated SPAD architectures based on 45 nm/22 nm 3D-stacked technology.

5.2 Sensor architecture and design

5.2.1 Floor plan

Fig. 5.1 shows a floor plan for SPAD sensor design based on 45 nm/22 nm 3D-stacked technology. An area of approximately 11 mm×7.5 mm has been segmented into small regions. Chip A-1 and A-2 are designed to operate independently, while sharing a single pad ring to deliver a full-resolution intensity image. Chip B, C, D, and E have pad rings for each, and those sensors have different pixel architectures, pixel pitches and array sizes. The rest (shown in gray) is devoted for test structures and other proof-of-concept sensors.
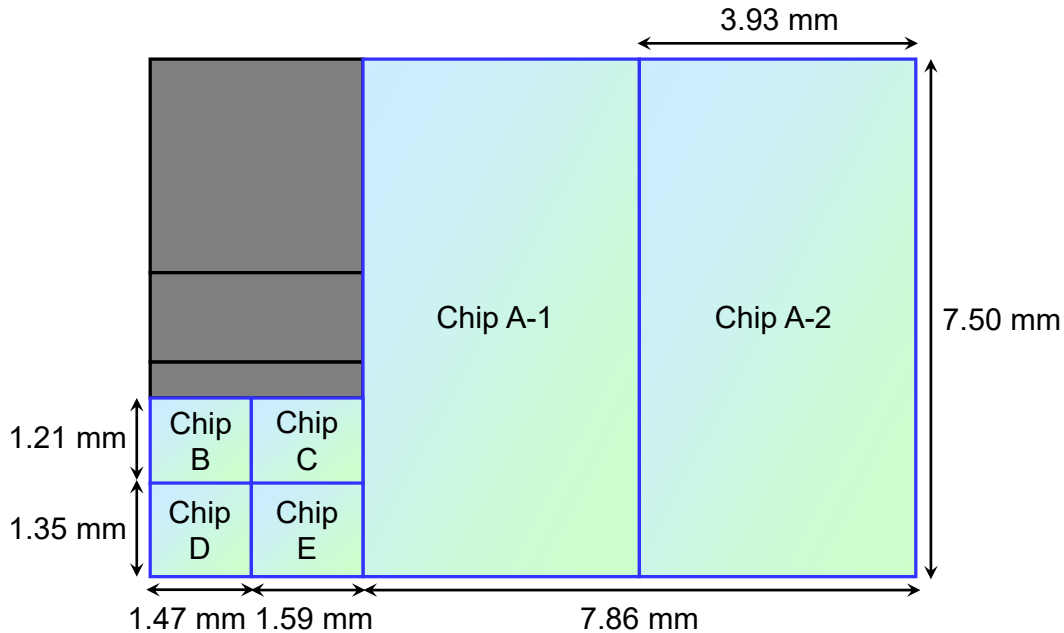
Figure 5.1 – Floor plan for sensor design based on 45 nm/22 nm 3D-stacked technology.

5.2.2 0.5 megapixel SPAD array for successive approximation ToF

As discussed above, time gating is a predominant approach towards scalable SPAD-ToF ranging. Still, there are several issues to be considered. First, this approach suffers from a tradeoff between distance range, depth resolution, and frame rate. Distance range and depth resolution are determined by gate scanning range and step, respectively. Fig. 5.2 (a) and (b) show a conventional pixel architecture for time-gated SPAD and the corresponding timing diagram, respectively. Relative timing of the gate window is scanned over the range of interest to perform linear-search of a reflection peak. Assuming fixed distance range, depth resolution is inversely proportional to frame rate. Similarly, with fixed depth resolution, measurable distance range (minimum to maximum) is inversely proportional to frame rate. In practice, it is rather difficult to apply the time-gate scanning method to long-distance and high-precision ranging while ensuring acceptable frame rate, e.g. ≥10 fps, to measure moving objects. The limited frame rate in the scanning method stems from its low efficiency of photon detection in time domain; signal photons are detected only when the time-gating window is aligned to the time-of-arrival of signal, whereas the signal is completely discarded otherwise.

To boost the frame rate of time-gated ToF, a novel successive approximation (SAR) algorithm is proposed. In Fig. 5.2 (c), typical circuit diagram for the SAR-ToF algorithm is demonstrated.
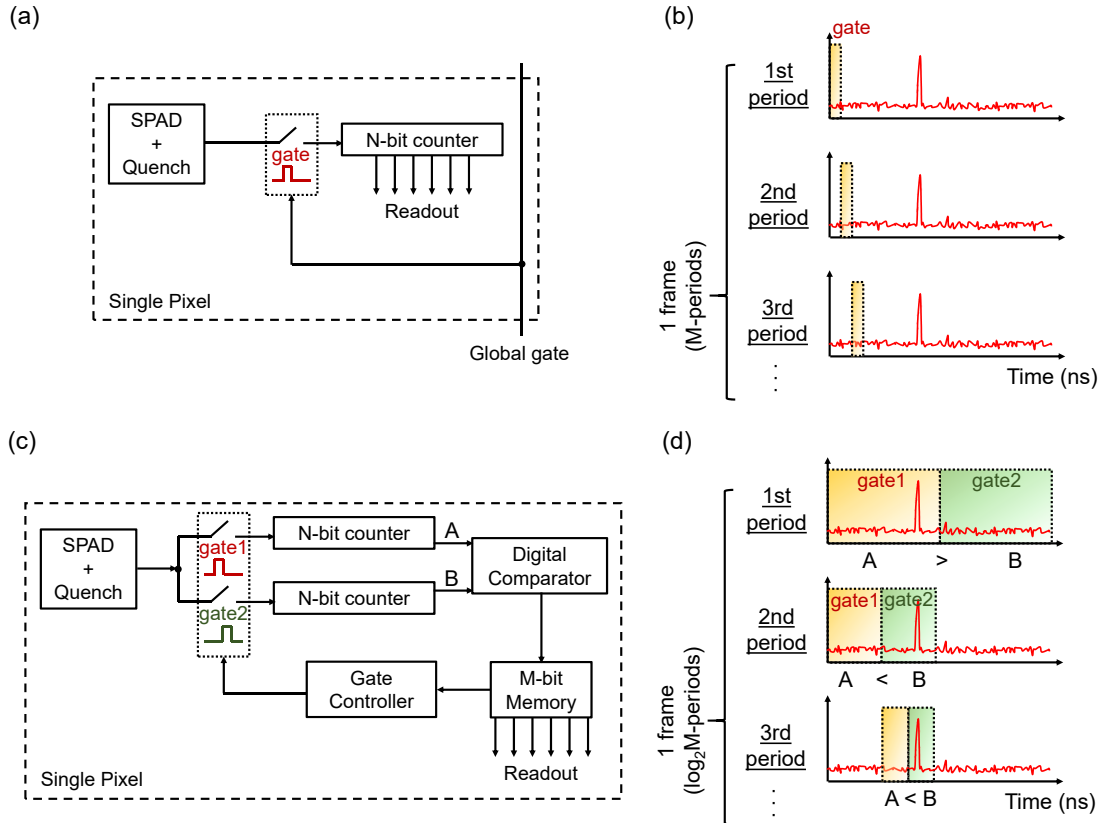
Figure 5.2 – Concept of proposed SAR-ToF. (a) Example circuit diagram. (b) Timing diagram.

The SPAD is connected to two parallel gate switches, followed by dual digital counters. The outputs of the two counters are fed to a digital comparator, and the comparison result is stored in a M-bit memory. Based on the stored data, a gate controller generates two gate pulses. Fig. 5.2 (d) is the corresponding timing diagram, where the red curve shows the ground-truth histogram. In the first period, gate 1 and gate 2 are complementary and have exactly the same time duration. Either one of them encloses the ToF peak. Based on the assumption that the background noise is time-uncorrelated, photon counts A and B in the repeated measurement in first period holds $A > B$ if the peak exists in gate 1, and vice versa. Taking the comparison result into account, timing offsets of gate window 1 and gate window 2 for the next period are determined pixel-wise. The two gate windows are kept with the same duration, but reduced by half with respect to the previous frame. By successively iterating the counting and comparison process, the signal peak position can be estimated. Unlike the scanning approach with the linear search, the SAR-ToF approach is based on a binary search, which is analogous to the relationship between slope- and SAR-ADC. In the linear scanning with the N-steps, the time duration of a single depth frame is proportional to N, whereas in the proposed SAR-ToF the time duration of a single depth frame is proportional to $\log_2 N$. This clearly suggests that the proposed SAR algorithm brings a remarkable improvement in the frame rate and realizes a dramatic compression of measured data.

The proposed SAR-ToF pixel has been implemented in chip A-2 in the floor plan. Fig. 5.3 (a) shows a schematic cross-section of the BSI SPADs with 6.93 μm pixel pitch, forming a half-megapixel (500×1024) pixel array. The avalanche multiplication junction is defined between NW and buried p-well (BPW). The SPADs are well-shared to achieve a high fill factor in a small pitch. The cathodes are connected to the circuit wafer via pixel-level bonding. Fig. 5.3 (b) and (c) are the single pixel layout and schematic in the circuit wafer, respectively. The pixel circuit consists of 13.86 μm-pitch unit to receive photon counting signals from 2×2 SPAD subpixels. The signals are combined in a coincidence detector, followed by dual gating circuit, dual counter, digital comparator, 8-bit register, clock and gate generator, and readout circuits. The coincidence detector, which detects more than one close events of photon arrival, is implemented to suppress influence of background noise. The background noise rejection is critical for this algorithm because too high noise level results in a misjudgement of the comparison process. In the readout process, faster operation is achieved by streaming out only the data stored in the register. By contrast, full readout of the data stored in the dual counters and the register provides richer information to estimate reliability of depth estimation, which is indispensable for constructing a more robust ranging system.

Note that the pitch can be further reduced by sharing components of the clock and gate generator, optimizing counter bit depth for the target application, and employing more advanced technology such as 14 nm-CMOS. Compared with the TDC methods, the SAR-based method is far more simple. Indeed, in the TDC method each pixel generates hundreds or even thousands of time stamps to get a single depth frame. By contrast, in the SAR method each pixel gives no more than 24-bit (at minimum 8-bit) information per frame. This data simplicity is beneficial for area- and power-efficient sensor design, as well as saving resources for data storage and signal processing.
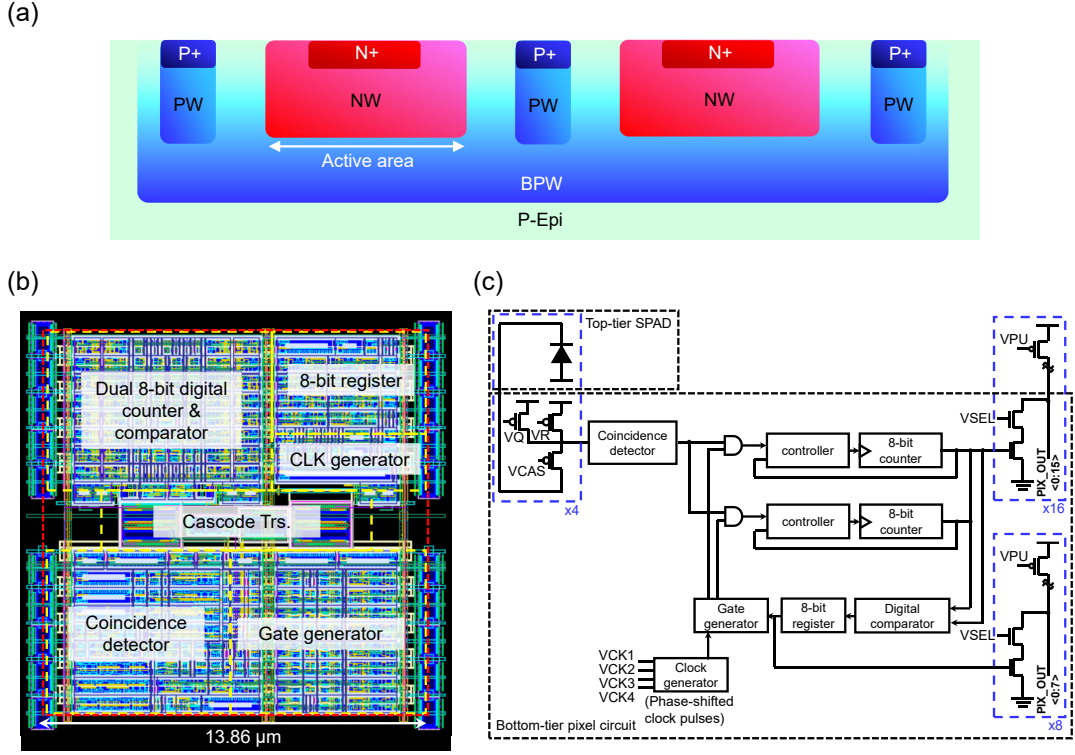
Figure 5.3 – 6.93 μm-pitch SPAD pixel design. (a) Schematic cross-section of SPADs. (b) Unit pixel circuit layout corresponding to 2×2 SPADs. (c) Schematic diagram of the unit pixel circuit.

5.2.3 Miniaturized time-gated SPAD array

Taking advantage of 3D integration, a possible implementation of time gating in a miniaturized SPAD pixel has been investigated. In chip B shown in the floor plan, we designed 2.97 μm-pitch time-gated SPAD pixel array, the smallest pitch ever reported in BSI SPADs. Fig. 5.4 (a) is the cross-section of designed SPADs. The avalanche multiplication junction is defined between PW and buried n-well (BNW). NW is shared with neighboring pixel to minimize the gap between adjacent active areas. As shown in Fig. 5.4(b) and (c), the pixel circuit consists of a cascode transistor [13], a quenching and recharge transistor, time-gating circuit (3-terminal AND), 4-bit asynchronous counter and readout. The cascode transistor helps enhancing the maximum applicable excess bias. The time-gating window is defined by the duration from the rising edge of $VG1$ to the rising edge of $VG2$. This architecture is suitable for achieving a sub-nanosecond gate window. The total number of transistors in a single pixel is 116. Thanks to 3D integration and area-efficient design, 10× transistors are realized with respect to a previously reported time-gated SPAD array in 180 nm FSI technology [10], while the pixel area is reduced by a factor of 32.

In the peripheral circuit, the gate pulses $VG1$, $VG2$ and the recharge pulse VQR are generated and homogeneously distributed to each column through balanced binary trees [10, 14, 15].

(a)

(b)

(c)
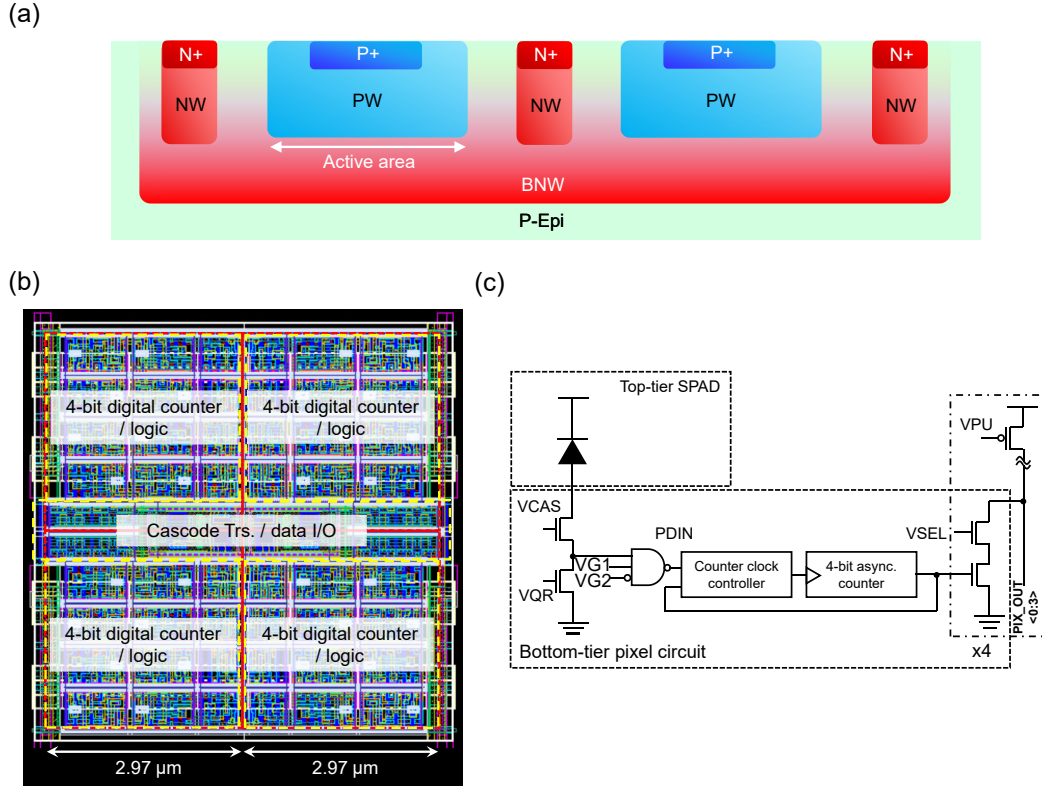
Figure 5.4 – 2.97 μm-pitch SPAD pixel design. (a) is the cross-section of designed well-shared SPAD in 45 nm CIS technology, (b) and (c) are the layout and the schematic diagram of pixel circuit designed in 22 nm CMOS logic process. The red dotted line in (b) corresponds to the pixel boundary.

Based on circuit simulation the horizontal timing skew of gate rising edge is estimated to be 20 ps, and the vertical skew 40 ps. The 4-bit pixel output lines are connected to 4 sets of D-flip flop chain, and the digital output is successively transferred to the output pad.

The designed pixel array size is 64×128, but can readily be extended to a megapixel in a compact sensor format, e.g. 4 mm×4 mm. This sensor can be potentially used for a wide variety of applications such as time-gated d-ToF ranging and FLIM, as well as Raman spectroscopy.

5.2.4 Dual time-gated SPAD array

A variant of the miniaturized time-gated SPAD sensor has been designed in chip C (see the floor plan). Fig. 5.5(a) and (b) are the layout and the schematic diagram of the pixel circuit. 4.95 μm-pitch pixels form a 128×128 array. A major difference with respect to the aforementioned 2.97 μm pixel design is that the pixel in this sensor has a dual parallel 4-bit counters. time-gating window can be individually defined for both gating circuits to enable the parallel acquisition of photons in different time duration, thus improving the light detection efficiency.
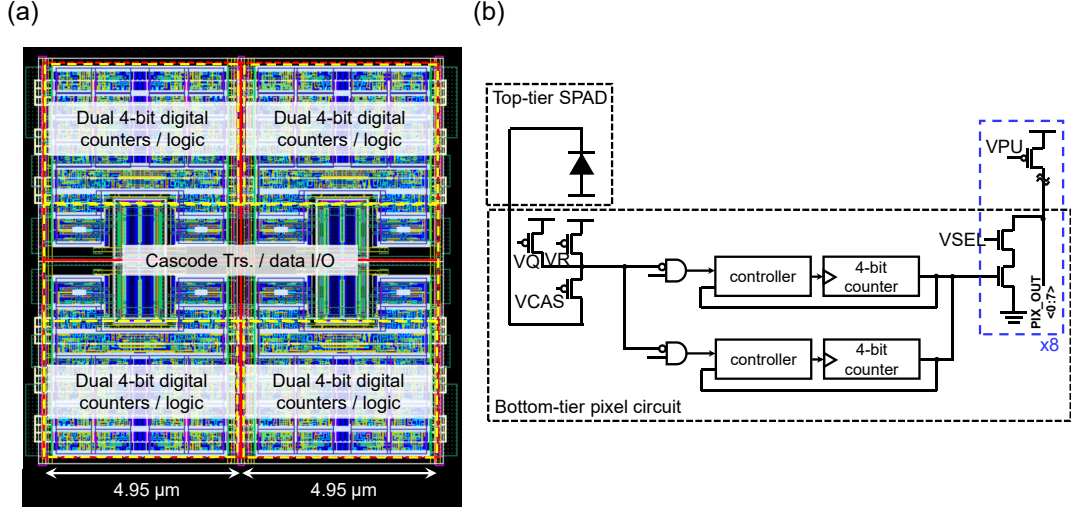
Figure 5.5 – 4.95 μm-pitch SPAD pixel design. (a) and (b) are the layout and the schematic diagram of pixel circuit designed in 22 nm CMOS logic process. Red dotted line in (a) corresponds to the pixel boundary.

5.2.5 Coded time-gating sensor

In addition to ToF ranging application, the time-gating approach is also suitable for FLIM. Among the various approaches, photon counting-based phasor FLIM is beneficial for intuitive visualization of complicated FLIM parameters [16, 17]. Fig. 5.6 (a) shows the conventional approach to estimate phasor parameter g and s, which are formulated as follows:

$$g \quad = \quad \frac{1}{N} \sum_{i=1}^{N} \cos\left(2\pi \tau_i / P\right), \tag{5.1}$$

$$s \quad = \quad \frac{1}{N} \sum_{i=1}^{N} \sin\left(2\pi \tau_i / P\right), \tag{5.2}$$

where, N is the total number of detected photons, τ_i is the i-th photon arrival time, and P is the period of laser pulse excitation. By defining a nanosecond time-gating window and by scanning its position, a fluorescence lifetime histogram is obtained. Based on the above equation, sinusoidal and cosinusoidal terms are weighted by photon counts at each gate position, and g and s are derived by arithmetic processing. However, when implementing a large-scale pixel array, this arithmetic processing necessitates considerable amount of memories and computation to achieve reasonable frame rate. Instead of scanning digital gate pulses, dynamic modulation of SPAD sensitivity is a possible solution for direct observation of
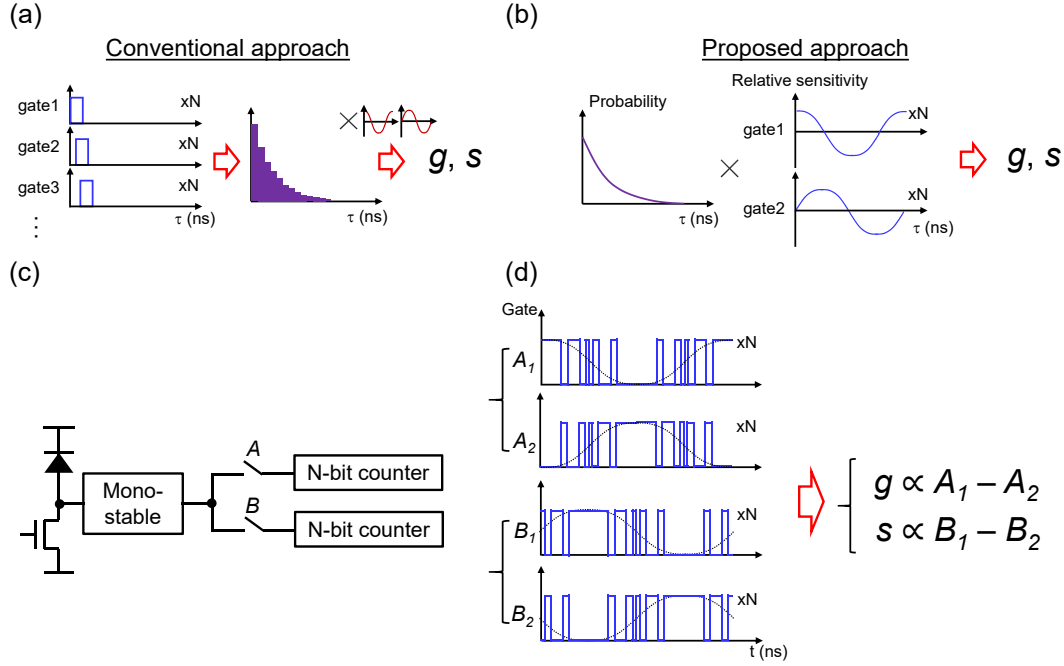
Figure 5.6 – Comparison between conventional and proposed phasor FLIM measurement schemes: (a) conventional approach with time-gate scanning; (b) proposed approach to measure phasor parameters using sinusoidal sensitivity modulation; (c) simplified circuit architecture for the proposed scheme; (d) timing diagram of digital gate pulses to reproduce pseudo-sinusoidal gate window.

phasor parameters without complex processing (see Fig. 5.6 (b)). In principle, SPAD sensitivity is a function of excess bias, and can be analogically modulated by providing dynamic signal to the SPAD bias. In particular, for the phasor application the sensitivity has to be modulated as a sinusoidal function in time-domain. In practice, this "analog" approach suffers from nonidealities such as pixel-to-pixel breakdown voltage variation. Figures 5.6 (c) and (d) describe more robust "digital" approach to realize pseudo-sinusoidal sensitivity modulation. The possible pixel configuration is a SPAD connected to a quenching transistor, followed by a monostable circuit, two parallel switches and dual counters. Switches A and B are turned on and off in accordance with a predetermined code sequence, which approximates sinusoidal and cosinusoidal curves in the lower frequency region. Two sets of measurements provide four different gate patterns with their photon counts A_1, A_2, B_1, and B_2, respectively. The phasor parameters can be estimated with significantly simplified equations:

$$g = \frac{(A_1 - A_2)}{(A_1 + A_2 + B_1 + B_2)/2}, \tag{5.3}$$

$$s = \frac{(B_1 - B_2)}{(A_1 + A_2 + B_1 + B_2)/2}. \tag{5.4}$$

These equations are much easier to calculate compared to the previous equations for gate scanning method, and thus a dramatic reduction of computational cost can be achieved. Targeting the aforementioned phasor FLIM application, a highly functional 9.9 μm-pitch 64×64 SPAD pixel is designed. The single and 4×4 pixel layouts are shown in Fig. 5.7 (a) and (b), respectively. The corresponding schematic is shown in Fig. 5.7 (c). Similar to the 4.95 μm-pitch pixel in the previous section, the proposed pixel contains two sets of counters with two parallel gating circuits. The gate pulses for the switches are generated by 64-bit memories and gate generator, which are implemented in 4×4 pixel unit. During the writing process, a 64-bit code to reproduce time-gating window pattern is stored in a 64-bit looped shift register. During the sensing process, 4 phase-shifted clock pulses are provided to a local clock generator, and the local clock with quadrupled frequency is generated. This clock is fed to the 64-bit
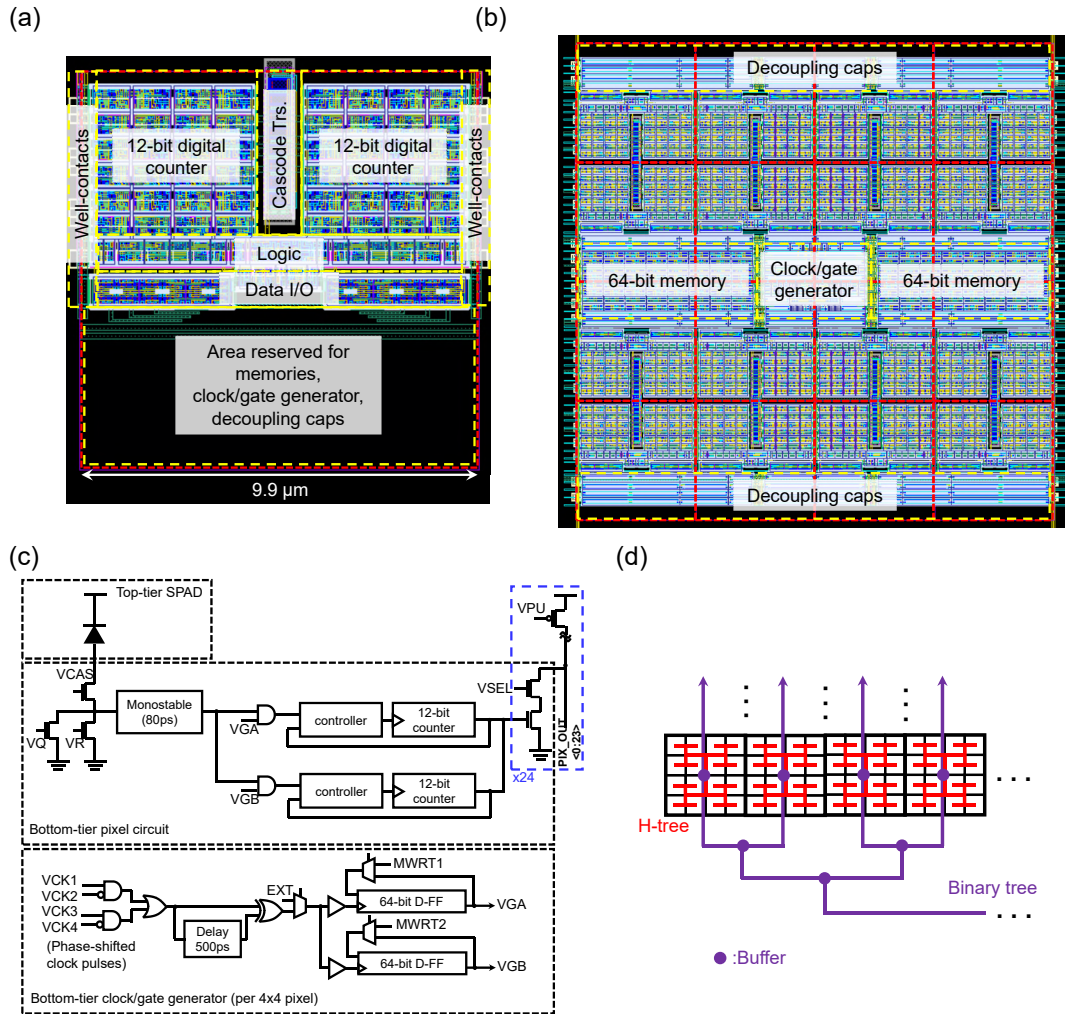
Figure 5.7 – 9.9 μm-pitch SPAD pixel design: (a) single pixel layout; (b) 4 × 4 pixel layout; (c) schematic diagram of pixel circuit; (d) conceptual diagram of H-binary-hybrid gate tree.

looped shift register to the generate a digital pulse sequence tracing the preprogrammed gating code pattern. When assuming the input clock frequency of 200 MHz, the local clock frequency goes up to 800 MHz by means of a cascade of two frequency doublers.

To achieve uniform GHz clock generation in each pixel block, it is critical to suppress the clock timing skew over the whole pixel array. The binary tree approach, as discussed in the previous section, is advantageous in reducing horizontal gate timing skew, whereas relatively worse vertical skew is expected due to the long propagation distance. Analogous to digital SiPMs [18], a full H-tree approach could be effective to improve both horizontal and vertical skew at the same time, while it suffers from its design complexity in a large-scale array. Here, a novel H-binary-hybrid approach to achieve reasonably low timing skew is proposed. As in Fig. 5.7 (d), the clock input is first distributed to each 2^N column groups using binary trees. The clock is propagated along the vertical direction towards a buffer placed at the center of each $2^N \times 2^N$ pixel block. The buffer generates the local clock pulses, which is distributed to all the pixels in the block through local H-tree. Compared to the conventional binary tree approach, the parasitic capacitance of a vertical clock line can be reduced by a factor of 2^N, and hence the propagation delay is suppressed accordingly.

5.3 Conclusion

In this chapter, various concepts of 3D-stacked time-gated SPAD sensors have been proposed. For ToF-ranging applications, an improvement in data acquisition speed of several orders of magnitude is expected. For phasor FLIM applications, dramatic simplification of signal processing can be achieved based on the proposed coded gating approach. The concepts are implemented in 45 nm/22 nm 3D-stacked BSI SPAD sensors with up to 0.5 megapixel array. The sensors are under fabrication during the writing of this thesis. Detailed measurement, analysis and applications will be continued by coauthors in AQUALab, EPFL in the near future.

References

[1] J. A. Richardson, E. A. G. Webster, L. A. Grant, and R. K. Henderson, "Scaleable Single-Photon Avalanche Diode Structures in Nanometer CMOS Technology," IEEE Trans. Electron Devices 58(7), 2028-2035 (2011).

[2] C. Veerappan, J. Richardson, R. Walker, D.-U. Li, M. W. Fishburn, Y, Maruyama, D. Stoppa, F. Borghetti, M. Gersbach, R. K. Henderson, and E. Charbon, "A 160×128 single-photon image sensor with on-pixel 55ps 10b time-to-digital converter," IEEE Int. Solid-State Circuits Conference (2011).

[3] A. Carimatto, S. Mandai, E. Venialgo, T. Gong, G. Borghi, D. R. Schaart, and E. Charbon, "A 67,392-SPAD PVTB-compensated multi-channel digital SiPM with 432 column-parallel 48ps 17b TDCs for endoscopic time-of-flight PET," IEEE Int. Solid-State Circuits Conference (2015).

[4] M. Perenzoni, D. Perenzoni, and D. Stoppa, "A 64×64-pixel digital silicon photomultiplier direct ToF sensor with 100Mphotons/s/pixel background rejection and imaging/altimeter mode with 0.14% precision up to 6km for spacecraft navigation and landing," IEEE Int. Solid-State Circuits Conference (2016).

[5] R. K. Henderson, N. Johnston, F. M. D. Rocca, H. Chen, D. D.-U. Li, G. Hungerford, R. Hirsch, D. Mcloskey, P. Yip, and D. J. S. Birch, "A 192×128 Time Correlated SPAD Image Sensor in 40-nm CMOS Technology," IEEE J. Solid-State Circuits 54(7), 1907-1916 (2019).

[6] N. A. W. Dutton, I. Gyongy, L. Parmesan, S. Gnecchi, N. Calder, B. R. Rae, S. Pellegrini, L. A. Grant, and R. K. Henderson, "A SPAD-Based QVGA Image Sensor for Single-Photon Counting and Quanta Imaging," IEEE Trans. Electron Devices 63(1), 189-196 (2016).

[7] M. Perenzoni, N. Massari, D. Perenzoni, L. Gasparini, and D. Stoppa, "A 160×120-Pixel Analog-Counting Single-Photon Imager with Time-Gating and Self-Referenced Column-Parallel A/D Conversion for Fluorescence Lifetime Imaging," IEEE J. Solid-State Circuits 51(1), 155-167 (2016).

[8] I. Gyongy, N. Calder, A. Davies, N. A. W. Dutton, R. R. Duncan, C. Rickman, P. Dalgamo, and R. K. Henderson, "A 256×256, 100-kfps, 61% Fill-Factor SPAD Image Sensor for Time-Resolved Microscopy Applications," IEEE Trans. Electron Devices 65(2), 547-554 (2018).

References

[9] Y. Hirose, S. Koyama, T. Okino, A. Inoue, S. Saito, Y. Nose, M. Ishii, S. Yamahira, S. Kasuga, M. Mori, T. Kabe, K. Nakanishi, M. Usuda, A. Odagawa, and T. Tanaka, "A 400×400-Pixel 6μm-Pitch Vertical Avalanche Photodiodes CMOS Image Sensor Based on 150ps-Fast Capacitive Relaxation Quenching in Geiger Mode for Synthesis of Arbitrary Gain Images," IEEE Int. Solid-State Circuits Conference (2019).

[10] A. C. Ulku, C. Bruschini, I. M. Antolovic, Y. Kuo, R. Ankri, S. Weiss, X. Michalet, and E. Charbon, "A 512×512 SPAD Image Sensor with Integrated Gating for Widefield FLIM," IEEE J. Selected Topics in Quantum Electronics 25(1), 1-12 (2018).

[11] J. M. Pavia, M. Scandini, S. Lindner, M. Wolf, and E. Charbon, "A 1 × 400 Backside-Illuminated SPAD Sensor With 49.7 ps Resolution, 30 pJ/Sample TDCs Fabricated in 3D CMOS Technology for Near-Infrared Optical Tomography," IEEE J. Solid-State Circuits 50(10), 2406-2418 (2015).

[12] T. Al Abbas, N. A. W. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. K. Henderson, "Backside illuminated SPAD image sensor with 7.83 μm pitch in 3D-stacked CMOS technology," IEEE Int. Electron Devices Meeting, 811-814 (2016).

[13] S. Lindner, S. Pellegrini, Y. Henrion, B. Rae, M. Wolf, and E. Charbon, "A High-PDE, Backside-Illuminated SPAD in 65/40-nm 3D IC CMOS Pixel With Cascoded Passive Quenching and Active Recharge," IEEE Electron Device Letters 38(11), 1547-1550 (2017).

[14] Y. Maruyama, J. Blacksberg, and E. Charbon, "A 1024×8, 700-ps Time-Gated SPAD Line Sensor for Planetary Surface Exploration With Laser Raman Spectroscopy and LIBS," IEEE J. Solid-State Circuits 49(1), 179–189 (2014).

[15] S. Burri, Y. Maruyama, X. Michalet, F. Regazzoni, C. Bruschini, and E. Charbon, "Architecture and applications of a high resolution gated SPAD image sensor," Opt. Express 22(14), 17573-17589 (2014).

[16] D. M. Jameson, E. Gratton, and R. D. Hall, "The measurement and analysis of heterogeneous emissions by multifrequency phase and modulation fluorometry," Appl. Spectrosc. Rev. 20(1), 55–106 (1984).

[17] R. A. Colyer, O. H. W. Siegmund, A. S. Tremsin, J. V. Vallerga S. Weiss, and X. Michalet, "Phasor imaging with a widefield photon-counting detector," Journal of Biomedical Optics 17(1), 016008 (2012).

[18] S. Mandai, and E. Charbon, "Timing optimization of a H-tree based digital silicon photomultiplier," Journal of Instrumentation 8(09) (2013).

6 Megapixel time-resolved SPAD image sensors

The quest for the first time-resolved megapixel SPAD sensor is not only a core objective of this thesis, but also a long-awaited goal after 17 years of intensive research and development in CMOS-compatible SPAD detectors and image sensors [1]. Major challenges lie in the chip design to optimize the sensor performance while minimizing timing skew, variation, noise, and power dissipation. This chapter is based on results presented in, K. Morimoto et al. "Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," Optica 7(4), 346–354 (2020), and the corresponding preprint, K. Morimoto et al. "A megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," arXiv:1912.12910 (2019).

6.1 Status of large-scale time-resolved SPAD array design

As discussed in Chapter 1, time-resolved image sensors are useful in many imaging applications. We mentioned VR/AR, LiDAR, FLIM, time-resolved Raman spectroscopy, and so on. Recently, a new class of applications in imaging has emerged, known collectively as quantum vision. We recognize as quantum vision techniques, such as ghost immaging, sub-shot-noise imaging, quantum LiDAR, quantum distillation, etc. [2, 3, 4]. Common to these applications is the need for single-photon detection and high timing resolution with low noise and high sensitivity. An important limitation in the majority of the implementations in quantum vision has been the image sensor, usually made of a single pixel or at most a 1kpixel array. Thus, a larger format picture requires 1D or 2D scanning, thus curtailing the frame rate, which is limited by the speed of the scanner. In addition, scanners may be bulky and add another level of complexity to the imaging system. To address these issues, researchers have recently created large-format cameras with a SPAD in each pixel and time gating or TDCs on chip [5, 6, 7, 8, 9]. Though, the crux of a large-format camera remains the pixel pitch and the amount of functionality per pixel. Researchers have thus resorted to 3D integration using backside-illuminated SPADs on the top tier and control/processing/readout electronics on the bottom tier [10, 11, 12, 13].

Recently, a novel photon-counting image sensor called QIS (see Chapter 1) has been demonstrated [14, 15, 16]. Single-photon sensitivity without avalanche gain is achieved in the QIS

with pixels having sub-0.5e- rms read noise as a result of a high conversion gain. QISs inherit several advantages of CISs such as a potentially small pixel size, high spatial resolution, low dark current, high quantum efficiency and low power consumption. A spatial resolution of up to 1 megapixel with 1.1 μm pixels has been reported in a QIS [17], enabling low noise and high dynamic range imaging for scientific, space, and security applications. Later, the first color-filtered 1 megapixel QIS has been demonstrated [18]. A limitation of the QIS technology, though, is timing resolution. Indeed, the finite time required for charge transfer in the pixels and sequential scanning readout prevent QISs from detecting timing information below 1 μs. SPADs, in contrast, enable single-photon detection with a timing resolution of up to few tens of picoseconds owing to the fast avalanche multiplication process. While a TDC-based approach enables precise time stamping of the detected photons in a SPAD array, it is not suited for scaling due to large circuit area and high power dissipation [19, 20, 21, 22, 23]. Our time-gating approach, in contrast to [5, 6, 7, 8], entails less than 8 transistors, and is promising for scalable photon counting image sensors towards sub-100-picosecond timing resolution and megapixel sensor resolution. In this chapter, we advocate the use of this approach to achieve large-format time-gated SPAD sensors capable of high timing resolution and small pixel pitch.

6.2 In-pixel time gating

In-pixel time gating approaches involve compact pixel circuits. They are therefore well-suited for the implementation of large-scale time-resolved SPAD sensors with low power dissipation. Fig. 6.1(a) illustrates the operation principle of time-gated ToF ranging. Laser pulses are repeatedly sent towards the target; the reflected photons are detected at the sensor with a delay of Δt. Typical time-gating measurements involve consecutive frames with a finely shifted gate window, each of which performs photon counting integrated over N sub-frames. Finer gate scanning improves timing resolution, while sacrificing depth measurement rate or range. From these measurements, a histogram may be derived, as shown in Fig. 6.1(b). Photon counts in a histogram bin include background photon counts and dark counts when the reflected laser pulse is outside of the gate window, whereas the photon counts are offset by the reflected laser signal when the reflected pulse is captured in the gate window. When the peak intensity of the reflected signal is higher than the background photon counts, the resulting photon count profile for each pixel forms a rectangular distribution with its width corresponding to the gate window length. The delay time Δt can be extracted from either rising or falling edge of this profile, while distance L from the detector to the target is estimated by:

$$L = \frac{c\Delta t}{2},$$ (6.1)
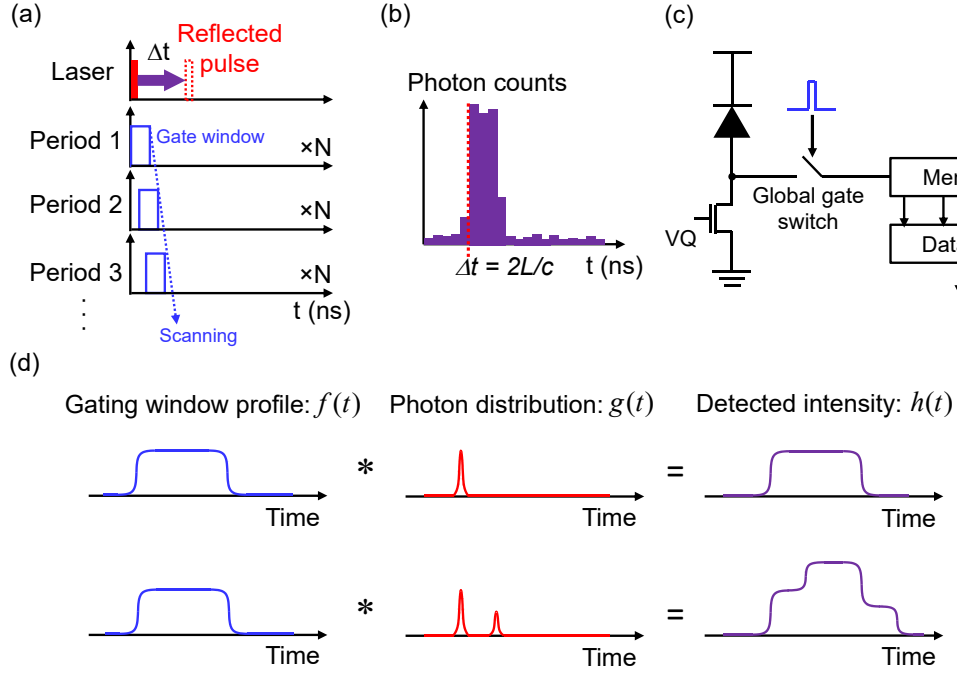
Figure 6.1 – Conceptual views of time-gated ToF ranging. (a) Timing diagram of ToF ranging based on time gate scanning, where N is the number of repeated measurement in a single period. (b) Expected photon count distribution as a function of gate position. (c) Pixel circuit architecture of time-gated SPAD sensor. (d) Schematic views of gating window profile, photon distribution and measured intensity over time with single reflective object (top) and double reflective objects (bottom).

where c is the speed of light. Fig. 6.1(c) shows a simplified schematic of a time-gated SPAD pixel. The SPAD is connected to a quenching transistor to limit self-sustained avalanche breakdown, and its output signal is selectively fed to an in-pixel memory when a global gate switch is activated. The gate control pulse can be as short as a few nanoseconds, and is synchronized with repetitive laser pulse emissions towards target objects. The stored signal is read out through a fast data I/O circuit.

More generally, the detected intensity profile $h(t)$ in a given measurement time frame is formulated by the convolution of two functions:

$$h(t) = f(t) * g(t), \tag{6.2}$$

where $f(t)$ is the gating window profile, and $g(t)$ the photon probability density function. Note that $h(t)$ yields a, when integrated from $-\infty$ to $+\infty$, where a is the total detected photon count in the measurement time frame. Fig. 6.1(d) shows the detected intensity profile of photons captured by the detector characterized by $f(t)$. When the photon probability density function

can be approximated by a single Gaussian distribution with a sufficiently small standard deviation, if compared to the gate length, the intensity profile can be expressed as:

$$h(t) \approx f(t) * a\delta(t - \Delta t) = af(t - \Delta t), \tag{6.3}$$

where a and Δt are the photon count and delay time of the Gaussian peak, respectively, and δ the Dirac delta function as an approximated form of a narrow Gaussian distribution. In practice, the detected intensity profile can take more complicated forms. For instance, when the target object is imaged through semi-transparent (semi-reflective) materials such as glass, plastics or liquids, the photon distribution can be expressed as a superposition of multiple Gaussian functions with different peak heights and positions. Assuming again negligible standard deviation, the detected intensity profile is:

$$h(t) \approx f(t) * [\sum_i a_i \delta(t - \Delta t_i)] = \sum_i a_i f(t - \Delta t_i), \tag{6.4}$$

where a_i and Δt_i are the photon count and delay time of the i-th Gaussian peak, respectively. Eq. (6.4) suggests that the multiple reflection results in a superposition of multiple gating window functions, each having different height and delay. An example with two reflective peaks in the photon distribution is shown in the bottom of Fig. 6.1(d).

Note that when $h(t)$ is measured and $f(t)$ is known, a full profile of $g(t)$ can thus be obtained by deconvolution. In a real situation, $h(t)$ can be distorted by non-ideal effects such as photon-shot noise, ambient light, dark counts, afterpulsing, crosstalk, timing jitter, etc. Those effects can introduce noise in the deconvolution. In ToF ranging, however, the assumption for reflected laser pulses to have negligibly narrow widths is valid in most cases. This assumption simplifies the process of distance calculation, where the time-of-arrival information can be readily extracted by finding the rising or falling edges in the measured intensity profile.

6.3 Sensor architecture and simulation

For the 1 Mpixel camera based on the SPAD pixel described above, with a pitch of 9.4 μm, we propose two pixel architectures: 7T (7 transistors per pixel) for pixel A and 5.75T (5.75 transistors per pixel on average) for pixel B without and with readout transistor sharing, respectively. The pixels achieve a fill factor of 7.0% and 13.4%, respectively; both pixels use a dynamic memory to store single-photon events generated by the SPAD. Binary photon counting images are captured and streamed out at 24,000 fps (maximum), corresponding to a total data rate of 25 Gbps.

Fig. 6.2(a) and (b) show the schematics and timing diagrams of both pixels. Upon detection of a photon, the SPAD generates an avalanche current pulse that is converted to a voltage through quenching transistor M_Q, which is controlled by V_{QR}. In pixel A, the voltage pulse is transferred to the gate of a transistor acting as memory through a gating transistor M_G,
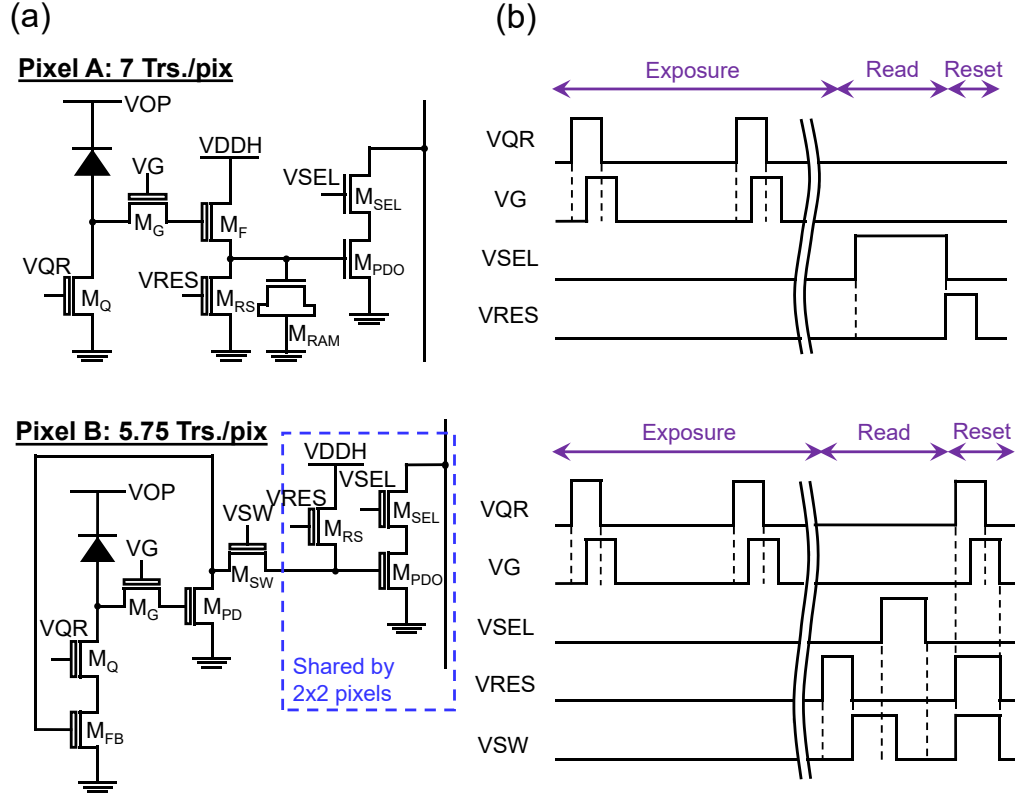
Figure 6.2 – Schematic views of designed SPAD pixels. (a) Pixel circuit schematics for pixel A and pixel B. Pixel A consists of thick- and thin-oxide transistors, whereas pixel B consists only of thick-oxide transistors. (b) Timing charts for pixel circuit operation.

controlled by gating signal V_G, and follower M_F, with a loss of 1 V_{TH}. M_{RS}, controlled by V_{RES}, is used to reset the dynamic memory, implemented by M_{RAM}, to GND. In pixel B, the voltage pulse is transferred to a pulldown transistor M_{PD} via gating transistor M_G, which is controlled by gating signal V_G. As a result, feedback transistor M_FB is set OFF, thus disconnecting the source of M_Q and disabling quenching in the SPAD. The voltage at the drain of M_PD is kept near GND for a sufficiently long time until the entire chip is read out and the node is charged again to V_{DDH}-V_{TH}-V_{DSAT} for the next detection, thus turning M_FB back ON. Transistor MSW, controlled by V_SW, connects the drain of M_PD to the source of M_RS, which is pre-charged to V_{DD}-V_{TH} via signal V_{RES}. In both pixel types, transistor M_{PDO} is used to pull down the entire column when M_{SEL} is turned ON by V_{SEL}, i.e. the row is selected in a sequential fashion. The difference is that in pixel B M_{PDO}–M_{SEL} are shared among a 2×2-pixel array, while in pixel A the dynamic memory and the readout transistors are thin-oxide devices. All the other transistors are thick-oxide devices to enable operation at 3.3 V. V_{DDH}, V_{DD}, and all the controls (when high) are 3.3 V. The feedback loop in pixel B prevents any subsequent avalanches within a frame; this is advantageous in very large arrays, since it reduces the current drawn from the cathode voltage node V_{OP} and thus the power dissipation from that node, which, given the
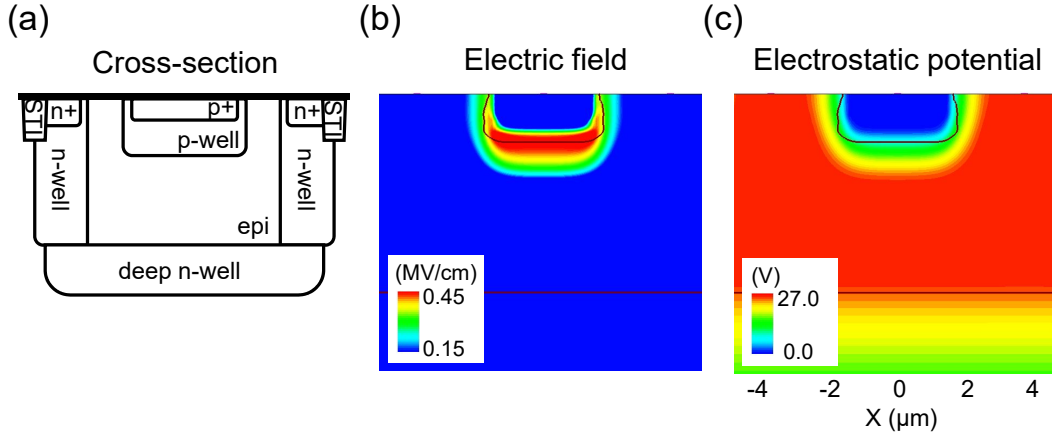
Figure 6.3 – Cross-section and TCAD simulation results of the SPAD. (a) Schematic cross-section. (b) Simulated electric field distribution. (c) Simulated Electrostatic potential distribution. The device dimension is set for pixel B.

high voltages used, can be significant for pixel counts above 100,000. In our chip, under strong illumination, the current drawn from V_{OP} by pixel A is over 400× that drawn from pixel B.

The SPAD, whose cross-section is shown in Fig. 6.3(a), was implemented as a p-i-n structure [24], whereas the avalanche region is surrounded by a buried implant at the bottom, to enable uniform field, and a circular guard ring on the sides, to suppress premature edge breakdown, as shown in the cross-section in the figure. All the layers employed in this design are standard in the 180 nm CIS process we used in the chip. A simulation of the electric field and of the potential is shown in Fig. 6.3(b) and (c), demonstrating the location of the avalanche region under the drawn active area, wherever the electric field exceeds the critical field for sustained impact ionization in silicon. The wider depletion region in z-direction with respect to other SPAD structures [25, 26] leads to the lower tunneling-induced DCR. The crosstalk was measured in both pixels reaching a mean of 0.17% for pixel A and 0.39% for pixel B. A higher crosstalk in pixel B is expected due to higher proximity to neighbors.

The measured maximum PDP for pixel A was lower than that of pixel B because the impact of the border effect is more significant when the active area is smaller. Based on the comparison with measured PDP for different active diameters in the same device structure, we estimated that the inactive distance from the border of the active area is around 0.6 μm. The camera block diagram is shown in Fig. 6.4(a); it comprises two independent sections of 1024×500 pixels, where the bottom half array is based on pixel A and the top half on pixel B. A dual binary tree controls the time gate, which reaches a minimum length of 3.8 ns (tunable up to 9.6 ns), and its variation of 120 ps (FWHM). Each row is read out in 83 ns and stored in a 1024-bit and a 512-bit output register for pixel A and pixel B, respectively, at the chip bottom and top. A multiplexer (MUX) scans it in 128-bit words, which are then transferred off-chip via 128 I/O pins (each half) by way of a dual parallel bus, thus achieving a frame rate of 24 kfps. The micrograph of the image sensor is shown in Fig. 6.4(b). The active area is drawn circular, and

(a)

(b)
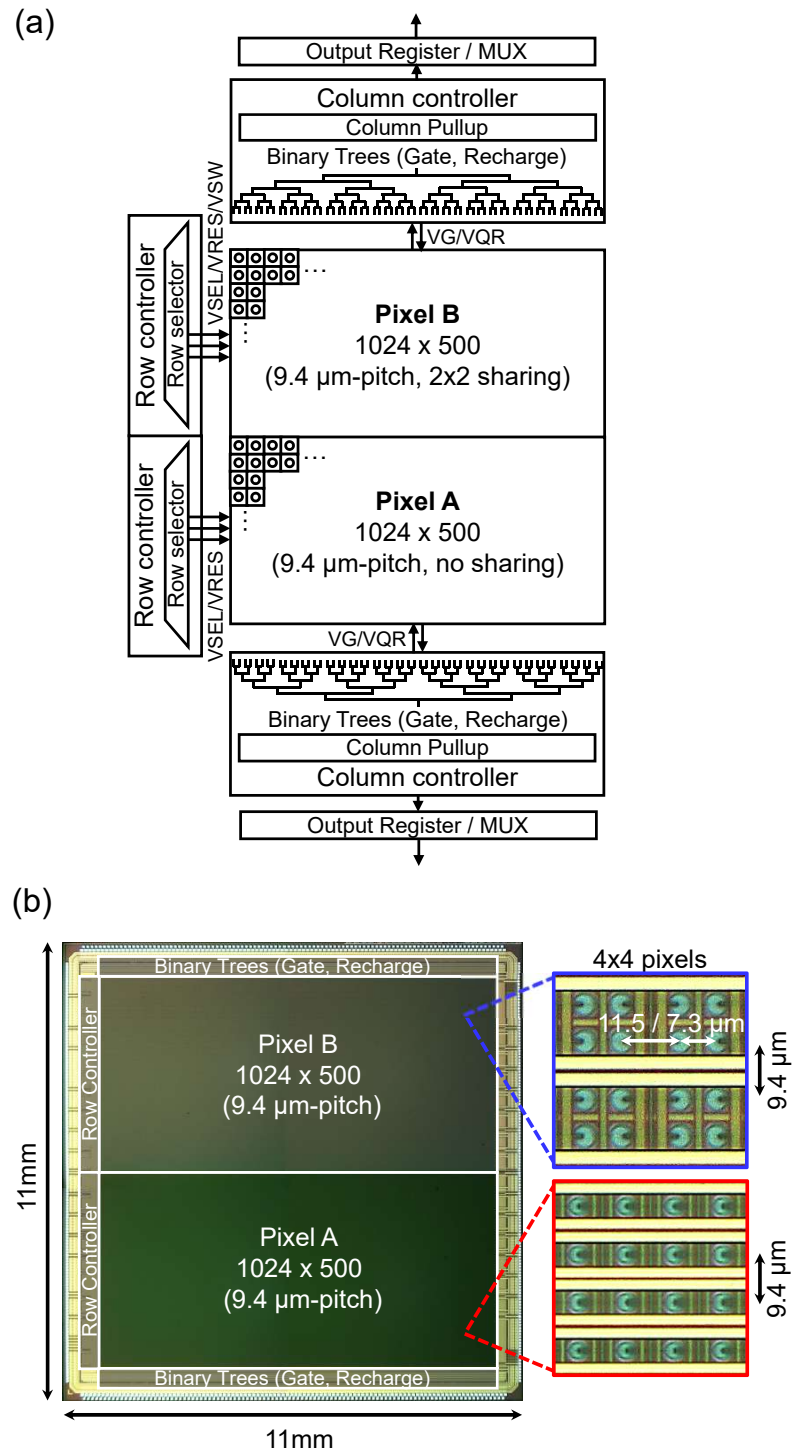
Figure 6.4 – 1 Mpixel time-gated SPAD image sensor architecture. (a) Sensor block diagram. (b) Chip micrograph and magnified views of pixel arrays.

the drawn active diameters for pixel A and B are 2.8 μm and 3.88 μm, respectively.
Unless otherwise noted, all following experiments are performed in GS mode with 42 μs readout per binary frame.

6.4 Camera system

Fig. 6.5(a) shows photographs of the developed 1 megapixel SPAD camera system. The boards consist of three layers; a sensor die is glued to the daughter board. A motherboard supplies power and biases to the sensor die, and interfaces the daughter board and two field-programmable gate array (FPGA) boards mounted on the backside. A 3D-printed lens mount, designed to attach C-mount lenses, is fixed on the daughter board. Fig. 6.5(b) is a photograph of chip-on-board bonded sensor die. The sensor has more than 700 pads for wire bonding; 256 pads are used for data output, and others for power, biasing, digital/analog inputs, and testing. The pads are arrayed with 50 μm-pitch in the top and bottom sides, and 70 μm-pitch in the right and left sides. Each pad on the die is wire-bonded to the corresponding pad on the daughter board. The pad positions are designed so that the Wire length does not exceed 5 mm to meet the requirement for the bonding process.

(a)　　　　　　　　　　　　　　　　　　　　(b)
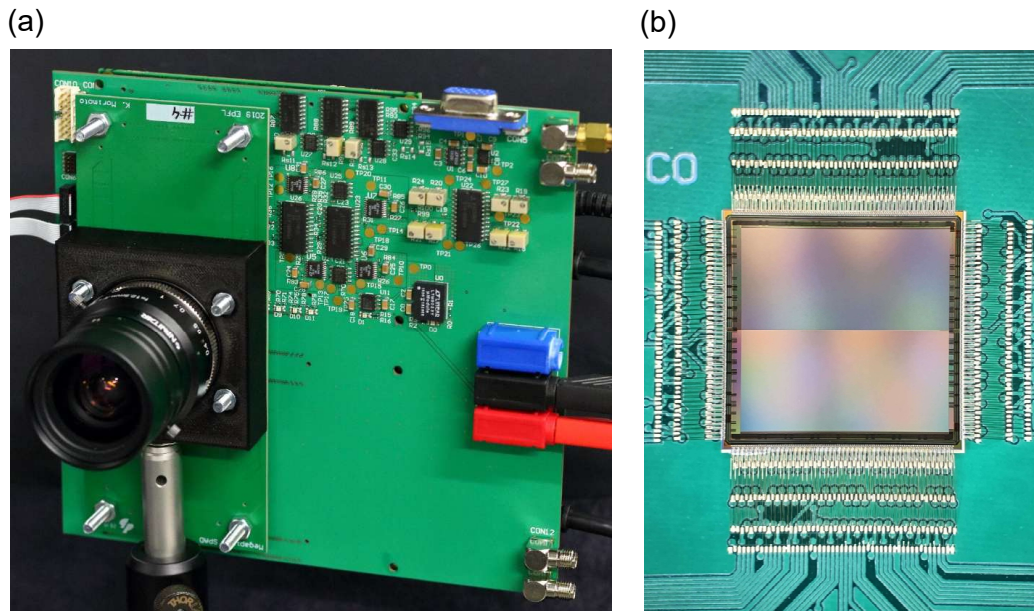
Figure 6.5 – Camera system of 1 Mpixel time-gated SPAD image sensor. (a) Photograph of 1 megapixel SPAD camera system. (b) Photograph of a chip-on-board bonded sensor die.

6.5 Experimental results

6.5.1 DCR and PDP

Fig. 6.6(a) shows the room temperature cumulative DCR probability distribution of the SPADs throughout the chip, with a median of 0.4 cps (pixel A) and 2.0 cps (pixel B) at an excess bias of 3.3 V. The corresponding DCR per unit drawn active area is 0.065 cps/μm^2 for pixel A and 0.17 cps/μm^2 for pixel B. These DCR density metrics are equal to or better than state-of-the-art SPAD devices [24, 29]. Fig. 6.6(b) shows the measured median DCR as a function of excess bias at room temperature. Fig. 6.6(c) shows the measured PDP as a function of wavelength. A maximum PDP of 10.5% (pixel A) and 26.7% (pixel B) is reached at 520 nm at the same excess bias of 3.3 V, while the PDP non-uniformity is better than 1.4% (pixel A) and 3.2% (pixel B) at room temperature. Lower PDP compared to the previous work based on p-i-n SPAD [24] is
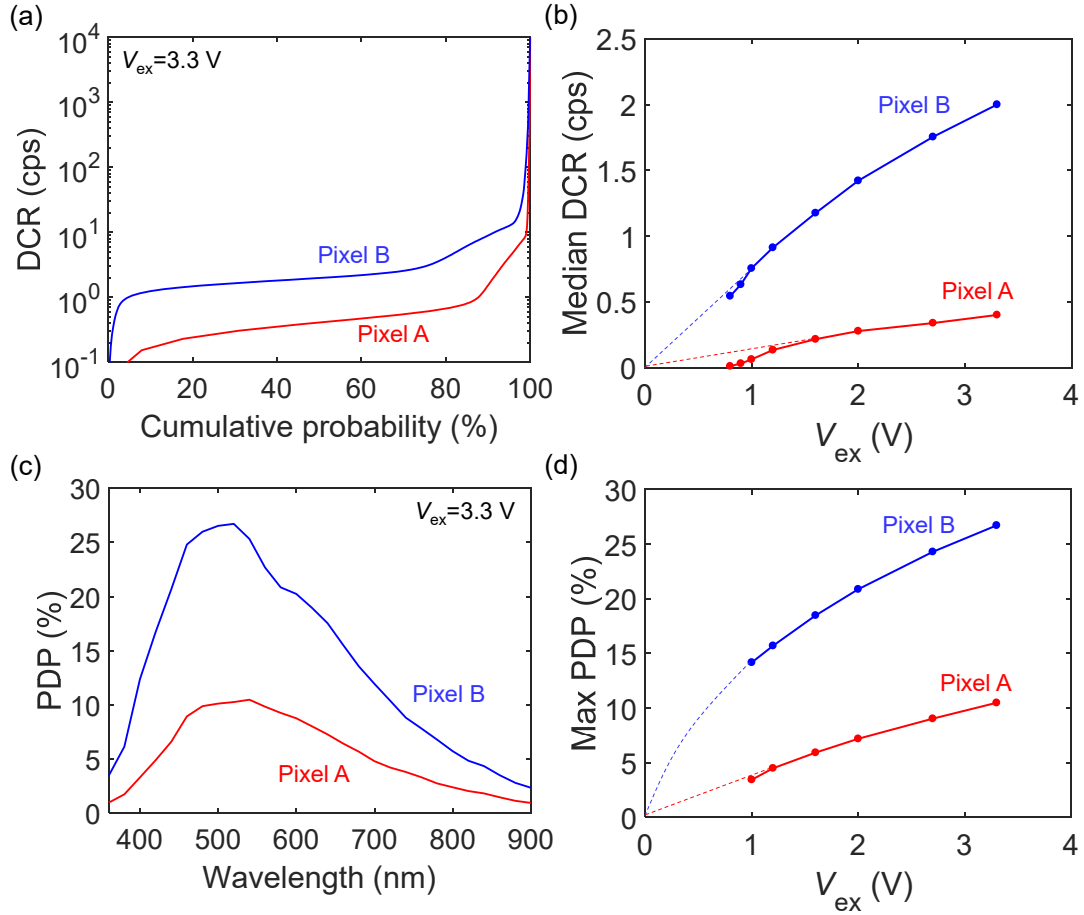
Figure 6.6 – Measured DCR and PDP for pixel A and B. (a) Room temperature cumulative histogram of DCR at excess bias of 3.3 V. (b) Excess bias dependence of median DCR at room temperature. (c) Wavelength dependence of PDP at an excess bias of 3.3 V. (d) Excess bias dependence of maximum PDP at room temperature.
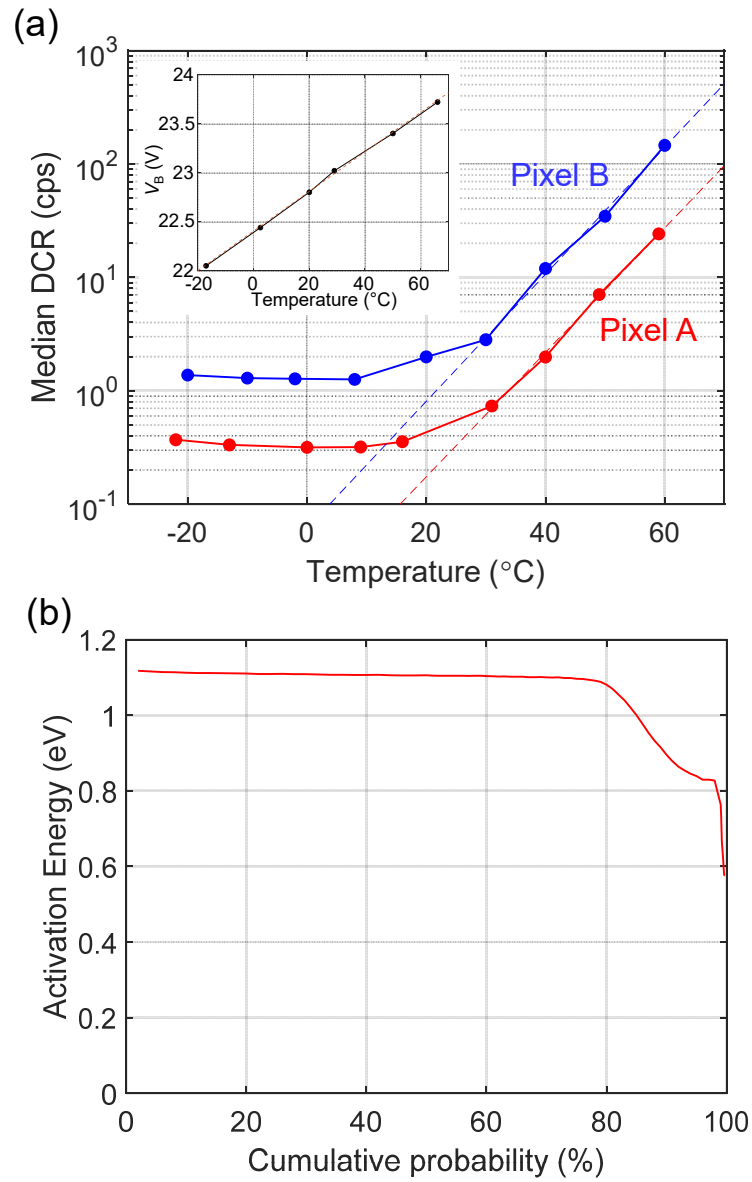
Figure 6.7 – Temperature analysis of DCR. (a) Temperature dependence of measured median DCR for pixel A and B. Temperature dependence of breakdown voltage is shown in the inset. (b) Activation energy distribution in pixel A, where horizontal axis shows pixel population in ascending order of DCR.

caused by typical border effects [30]. The border effects are more significant when the active diameter is smaller than 5 μm. Fig. 6.6(d) shows the maximum PDP as a function of the excess bias, whereas the dotted lines are guides for the eye.

Fig. 6.7(a) shows the temperature dependence of DCR for pixel A and B. Based on the temperature dependence of breakdown voltage shown in the inset, VOP is adjusted to keep V_{ex} = 3.3 V over measured temperature range. Median DCR for both pixels shows almost no temperature dependence at T < 10 °C, whereas it increases exponentially at T > 30 °C. For both pixels, the activation energy for T > 30 °C is extracted to be 1.1 eV, equivalent to the band gap of silicon. The result indicates that DCR in the majority of pixels is dominated by tunneling at low temperature and by diffusion current at high temperature [27, 28].

Fig. 6.7(b) is the distribution of the activation energy for pixel A, where the horizontal axis corresponds to that of Fig. 4(a) in the main text. Approximately 80% of the pixels in the array show an activation energy of 1.1 eV, whereas the remaining 20% with higher DCR exhibits between 1.1 eV and 0.55 eV. This indicates that those 'hot' pixels have mixed DCR sources from diffusion current and Schockley-Read-Hall (SRH) generation-recombination, which typically lead to the activation energies of 1.1 eV and 0.55 eV, respectively.

6.5.2 Time-gating performance

The timing performance of pixel A was characterized in Fig. 6.8. A 785 nm laser pulsed at 25 MHz (average power: 5 mW, optical pulse width: 80 ps FWHM, ALS GmbH, Berlin, Germany) illuminates the whole array, while the time gate window is continuously shifted with respect to the laser trigger by steps of 36 ps over a range of 10 ns. For each gate position, 255 binary frames are acquired and summed in a Kintex™ 7 FPGA (Xilinx Inc., San Jose, CA) to generate an 8-bit image. Fig. 6.8(a) shows the gate window profiles for 160 pixels uniformly sampled from the bottom-left to the top-right of the pixel array. Broadening of rising and falling edges indicates the non-uniformity of gate signal propagation over the pixel array. Figs. 6.8(b) and (c) demonstrate the spatial uniformity of gate position and gate length. The gate and recharge signals injected from the bottom side of the array require more time to propagate and activate the gate of the pixels in the top of the array. Horizontal skew of the gate position in the top side of the array stems from the asymmetry of power routing where the power and ground are supplied from the left, right and bottom side, but not from the top side of the array. The gate length distribution shows better uniformity than the gate position distribution. Note that white pixels in the color plots indicate that the measured timing is out of range with respect to the color bar. Fig. 6.8(d) shows the histograms of gate position, gate length, rise time and fall time. The gate position skews and variation in gate length were measured as 410 ps and 120 ps (FWHM), respectively, while an average gate length of 3.8 ns was achieved. These values include the laser pulse width of 80 ps FWHM.
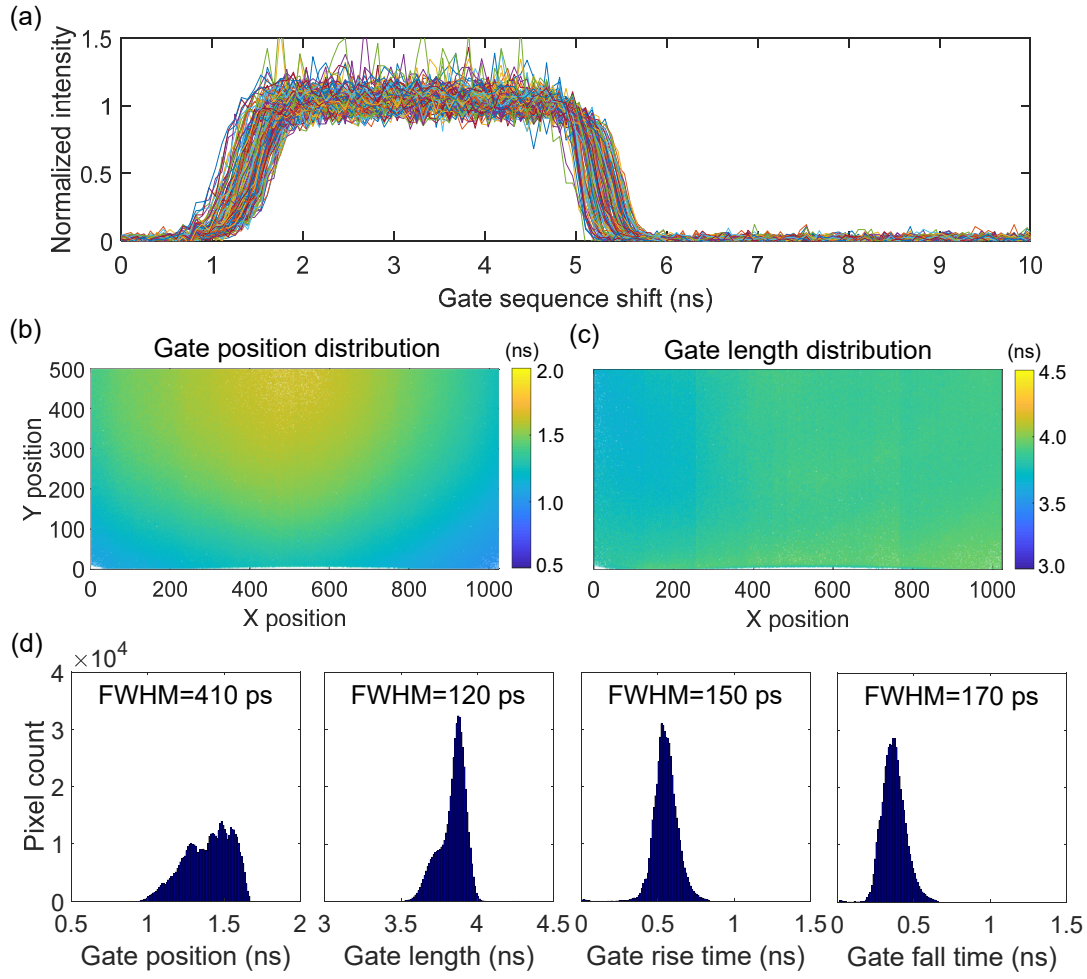
Figure 6.8 – Measured time-gating performance for pixel A. (a) Gate window profiles for uniformly sampled 160 pixels. (b) Color plot of gate position distribution over 1024×500 pixels. (c) Color plot of gate length distribution over 1024×500 pixels. (d) Histograms for gate position, gate length, rise time and fall time.

6.5.3 Power consumption analysis

Fig. 6.9 shows the measured power consumption of the megapixel SPAD sensor as a function of incident photon flux. Fig. 6.9(a) is the power consumption for pixel A consisting of VOP, 3.3 V core, 1.8 V core and 3.3 V I/O components, respectively. The power consumption at VOP is dominated by avalanche-induced current, and it proportionally increases to the incident photon counts. The power consumption component at VOP becomes dominant in the total power consumption under high light condition and reaches 9.118 W with SPAD saturation, where every gate/recharge cycle in all the pixels detects a photon. The power consumption is saturated at higher incident photon counts because no more than one photon detection event can take place in a single detection cycle. 3.3 V core component is dominated by generation of pixel control signals for VG and VR, and is independent of incident light intensity. 1.8 V core originates from pull-up and pull-down of vertical signal lines in the pixel array. The power consumption at 1.8 V core is proportional to the average number of photon detection events over pixel array, and hence it increases linearly in the lower light and saturates at higher light intensity. 3.3 V I/O increases when output binary signal switches frequently between '0' and '1'. It shows the peak at intermediate light intensity where the output binary signal varies randomly. The total power consumption, shown in black curve, exhibits complex behavior due to the mixture of multiple components with different behavior.

Fig. 6.9(b) is the power consumption for pixel B. Critical difference with respect to pixel A is observed in VOP. In pixel B, the power consumption through VOP saturates at the intermediate incident photon counts and the maximum consumption reaches only 0.021 W, approximately 400 times smaller than in pixel A. This stems from the fact that the feedback loop in pixel B closes the recharging path once first photon is detected. This suppresses any extra avalanche multiplication for photons which do not contribute to actual photon counting signal, whereas those extra photons can trigger an avalanche in pixel A due to the different pixel architecture. Another difference is the reversed behavior of 1.8 V core caused by the inverted signal output scheme from the pixels. In contrast to pixel A, the power consumption at 1.8 V core in pixel B is proportional to the ratio of '0' in output data stream, which shows monotonic decrease as a function of incident photon counts.

6.5.4 2D imaging

The chip was tested as an intensity image sensor with a standard chart. Fig. 6.10(a) shows a schematic view of the experimental setup. Fig. 6.10(b) shows a 1 Mpixel monochrome image obtained at 24 kfps with a uniform illumination of 50 lux (indoors). For each half, 16,320 binary images are taken in RS mode and are summed to acquire a 14-bit intensity image. Exposure time is set at 2 s to compensate the smaller aperture size for wide depth of field. The image contrast for top and bottom half is tuned independently to compensate the difference in the PDE. On the right side of Fig. 6.10(b), the magnified images show that the line patterns are well-resolved, up to number 10 in the chart, indicating the spatial resolution of 1000 dots within the horizontal and vertical field of view.

(a)

(b)

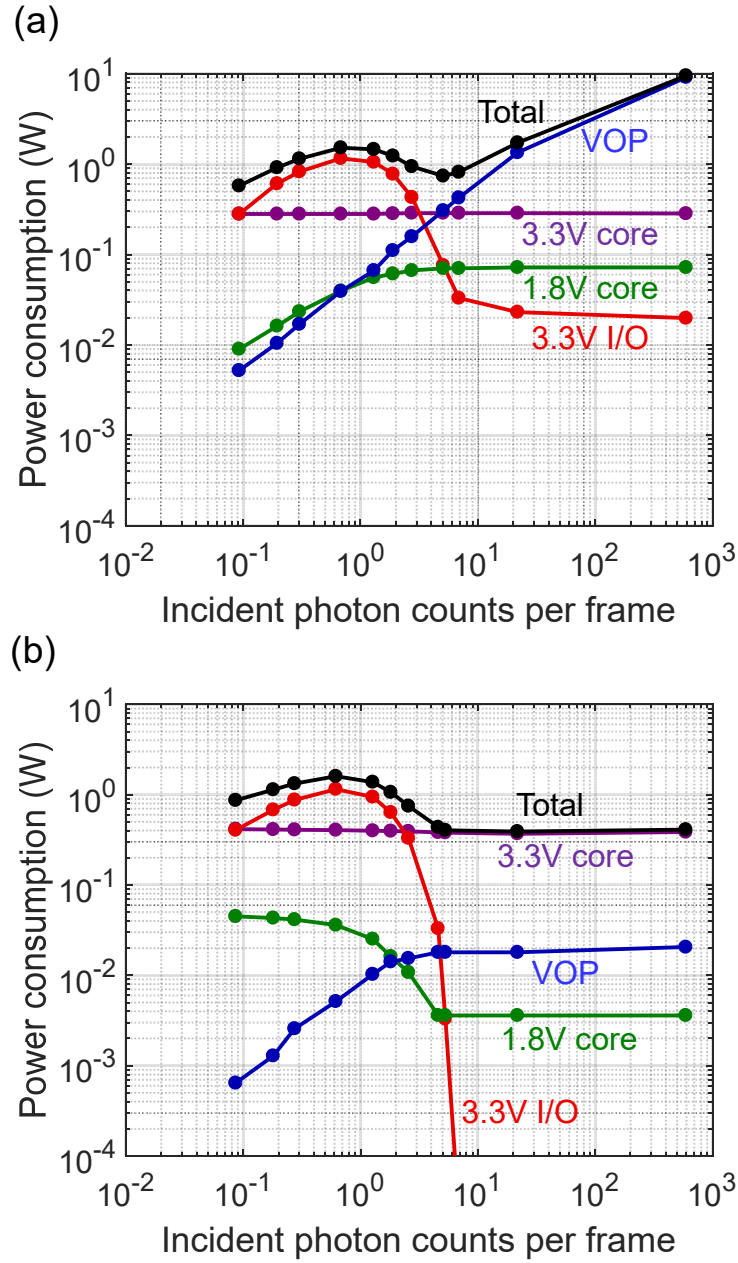Figure 6.9 – Measured power consumption as a function of incident photon counts per frame. (a) Power consumption for pixel A. (b) Power consumption for pixel B.
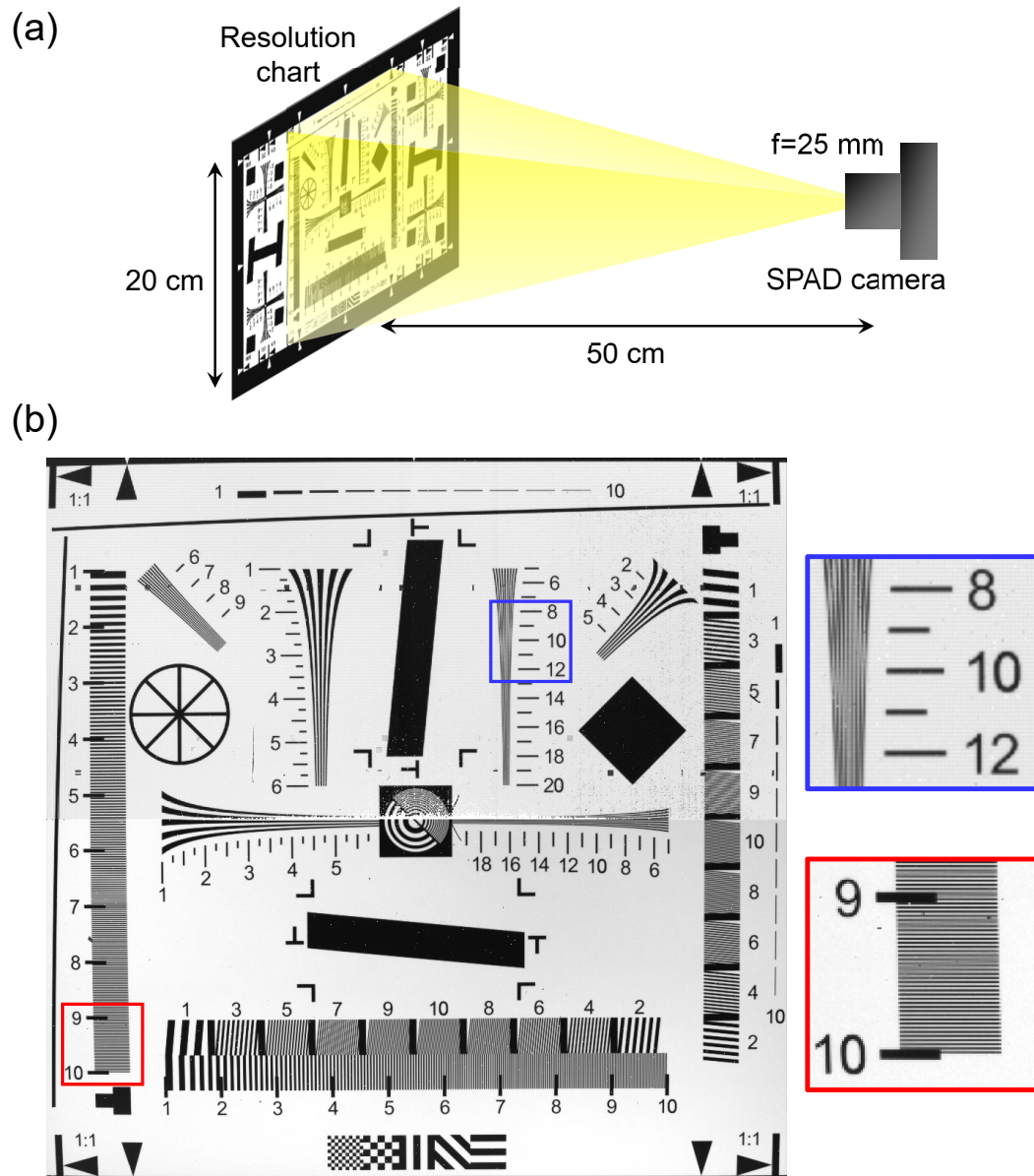
Figure 6.10 – A 2D intensity imaging of standard test chart with 1 Mpixel resolution. (a) Experimental setup. (b) A 14-bit image obtained by summing 16,320 binary images. Magnified views of two small areas, indicated by blue and red squares are shown on the right.

The dynamic range of a 2D image sensor is critical for a wide range of applications. Recently, a method to extend the dynamic range by mixing multiple different exposure times in a single frame has been proposed and analyzed for CMOS-based QIS [31], and reported for SPAD-based binary image sensors [32]. Compared to the case with fixed single exposure time for all the binary frames, mixing multiple exposure times results in slower saturation of the output counts when increasing the incident photon flux, giving richer tones for high illumination conditions. Yet, dynamic range extension based on interleaved multiple exposures in a SPAD sensor is reported only for a limited sensor resolution of 96×40 pixels. In addition, incident photon count dependencies of output signal, noise and SNR have not yet been systematically compared between single and multiple exposure modes under equalized total exposure conditions.

Fig. 6.11(a) shows the timing sequences of single and dual exposure modes in a time-gated SPAD sensor. The sensor is operated in GS mode; each shaded region in the figure represents a global exposure, followed by sequential readout of a full-resolution binary frame. A set of streamed binary frames is integrated in the FPGA to construct one N-bit image. In single exposure mode, the global exposure time is fixed over one N-bit frame. In dual exposure mode, short and long global exposures are staggered to form one N-bit image. In this experiment, the ratio between short exposure time τ_S and long exposure time τ_L for dual exposure mode is set at 1 to 8, whereas the exposure time τ_M for single exposure mode is set at $4.5\tau_S$. For systematic comparison of the two operation modes, maximum photon counts and total exposure time in a single N-bit frame are set equal for the two modes; $2\tau_M = \tau_S + \tau_L$.

Fig. 6.11(b) shows the measured output photon counts as a function of incident photon counts for single and dual exposure modes. 4080 binary frames are summed to form a 12-bit image. τ_S is set to 121 μs, whereas the readout of a binary frame takes 42 μs. The total exposure time is 2.2 s for both single and dual exposure modes. Dotted lines are the fitted curves for each mode based on the following equations:

$$N_{OUT}^S = N_{sat} \times (1 - e^{-\frac{N_{in}}{N_{sat}}}), \tag{6.5}$$

$$N_{OUT}^D = \frac{N_{sat}}{2} \times [(1 - e^{-\frac{2\tau_L}{(\tau_L+\tau_S)} \cdot \frac{N_{in}}{N_{sat}}}) + (1 - e^{-\frac{2\tau_S}{(\tau_L+\tau_S)} \cdot \frac{N_{in}}{N_{sat}}})], \tag{6.6}$$

where N_{out}^S and N_{out}^D are the output counts in the single and dual exposure modes, respectively, N_{sat} is 4080, and N_{in} is the incident photon count per one N-bit frame. The fitted curves are in good agreement with the trends of measured output counts. The output counts of dual exposure mode saturate later than those of single exposure mode, indicating the extended dynamic range.

Fig. 6.11(c) shows the standard deviation of measured outputs as a function of incident photon counts. Raw output counts of 100 pixels in the center of the array are used to calculate the standard deviation, whereas the photon-shot noise limit is also shown. In the lower
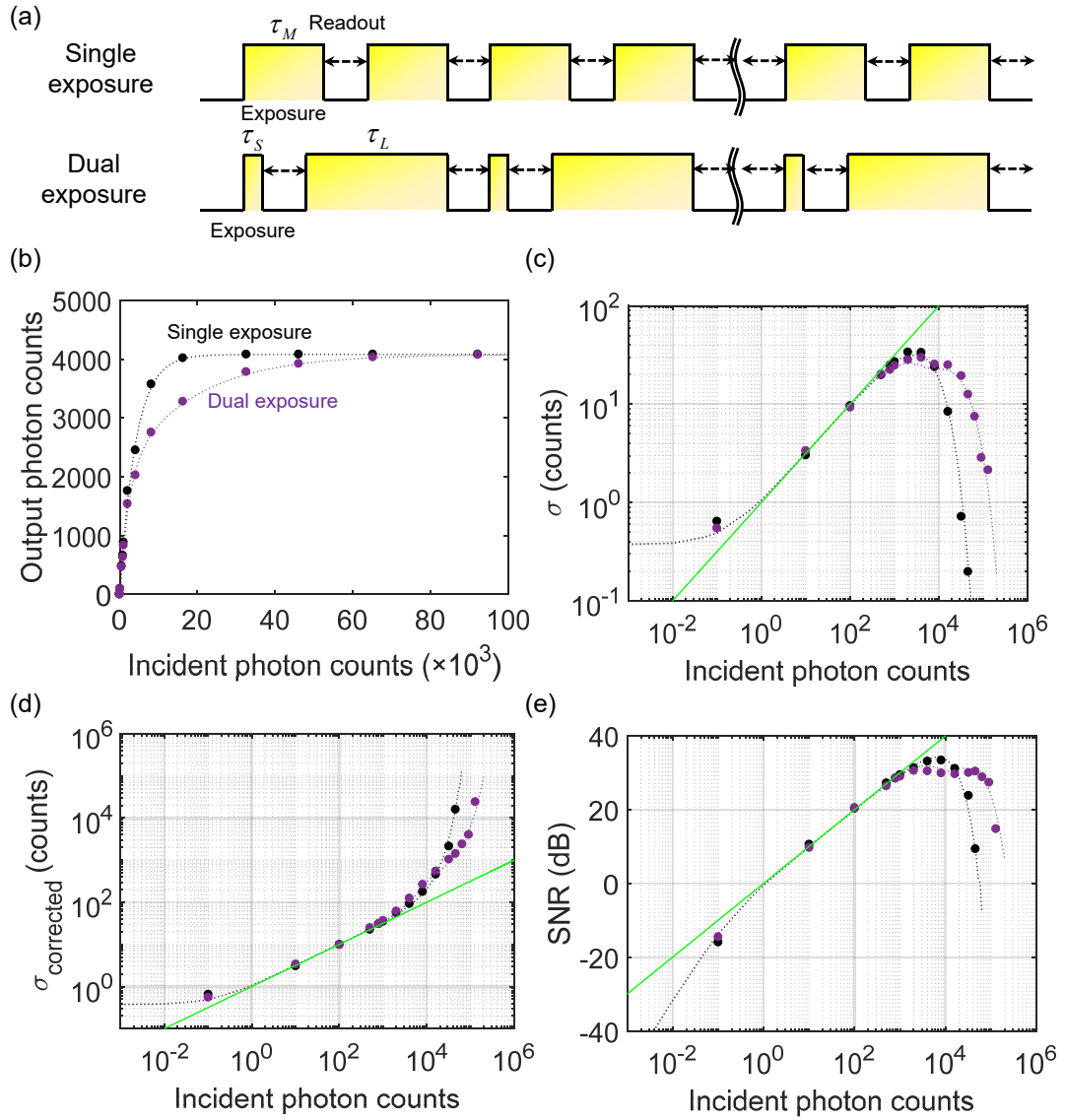
Figure 6.11 – Conceptual view and measured or simulated results for the dynamic range extension technique. (a) Timing diagrams of single and dual exposure modes. (b) Measured (markers) and fitted (dotted lines) output photon counts as a function of incident photon counts for pixel A. (c) Measured (markers) and Monte Carlo-simulated (dotted lines) standard deviation. (d) Measured and simulated standard deviation after linearity correction. (e) Measured and simulated SNR. Green lines indicate the photon-shot noise limit.

(a)

(b)
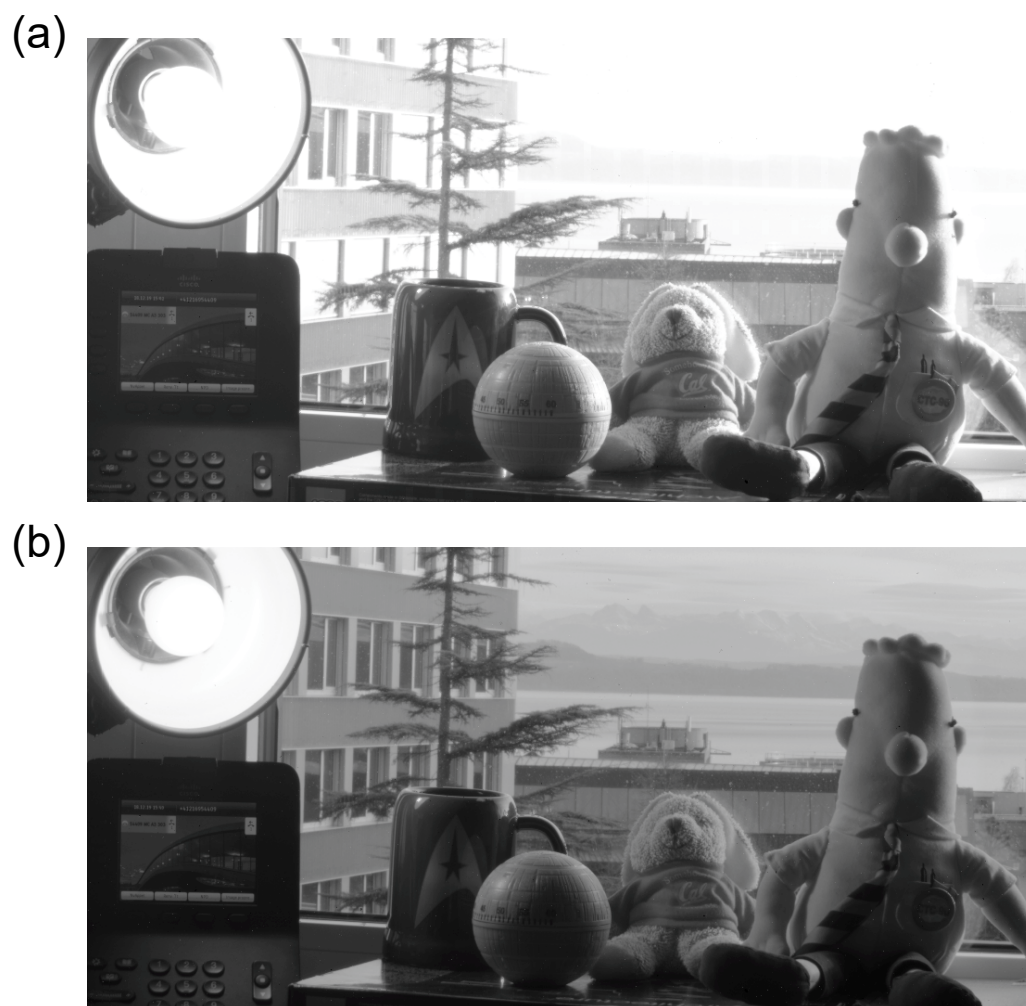
Figure 6.12 – 2D images of a real-life scene captured with pixel A: (a) 18-bit image taken in single exposure mode; (b) 18-bit image taken in dual exposure mode.

incident photon counts, the measured standard deviation is higher than the shot noise limit due to the contribution from DCR non-uniformity. Under intermediate photon counts, the measured standard deviation follows the shot noise limit. For higher incident photon counts, the measured standard deviation is lower than the shot noise limit due to the compression of the output signal when a saturation of 4080 counts is reached [33]. To reproduce the output characteristics, a Monte Carlo simulation was performed based on Poissonian statistics of the incident photons and dark counts for each binary frame. The simulation results for two modes are shown as dotted curves, which are highly consistent with the measured trends.

In real situations, nonlinear output characteristics in Fig. 6.11(b) has to be corrected to ensure the natural contrast for human eyes. Fig. 6.11(d) shows the measured standard deviation based on corrected output counts. Similar to the trends in Fig. 6.11(c), the deviation from photon-shot noise limit is observed in the lower incident photons due to the DCR non-uniformity, and the deviation is suppressed for intermediate photon counts. For the higher incident photons, the measured trends go above the shot noise limit. The difference with respect to the uncorrected curves of Fig. 6.11(c) arises from the amplification of photon-shot noise in the linearity correction process. Again, Monte Carlo simulations (dotted lines) precisely reproduce the measured results for both operation modes. The noise increase in the dual exposure mode is observed later than that of the single exposure mode, which is a direct consequence of dynamic range extension.

Fig. 6.11(e) shows the measured SNR plots for the two exposure modes. The dynamic range, defined as a ratio of input-referred photon counts for 99% of saturation counts to dark counts, was measured at 96.3 dB for single exposure and 108.1 dB for dual exposure. A 11.8 dB improvement is demonstrated with equal maximum counts and total exposure time. The highest SNR for single and dual exposure is 33.3 dB and 30.5 dB, respectively.

In Fig. 6.12, the effect of the dynamic range extension is investigated in a real-life scene. In single exposure mode, the background scene is overexposed (Fig. 6.12(a)), while the gray-scale tone of the scene is clearly visible in dual exposure mode (Fig. 6.12(b)). The difference in the maximum SNR is too small to be recognized.

6.5.5 3D imaging

Fig. 6.13(a) and (b) show a 2D and a color-coded 3D pictures obtained by illuminating a scene with a 637 nm-laser pulsed at 40 MHz (average power: 2 mW, optical pulse width: 80 ps FWHM, ALS GmbH, Berlin, Germany) and captured on the half-resolution image sensor (pixel A). The gate window with its length of 3.8 ns is shifted from 0.6 ns to 13.2 ns by steps of 36 ps to acquire full photon intensity profiles as a function of the gate position. The distance LSB in this measurement corresponds to 5.4 mm. The intensity profile for each pixel is smoothed by taking the moving average over gate positions to suppress the effect of photon-shot noise. The depth information is reconstructed by detecting the rising edge position of the smoothed intensity profile for each pixel, corresponding to the time-of-arrival of the reflected laser pulse. The gate timing skew over the array is compensated by subtracting the background distribution shown in Fig. 6.8(b) from the measured time-of-arrival distribution. In Fig. 6.13(b),

(a)

(b)

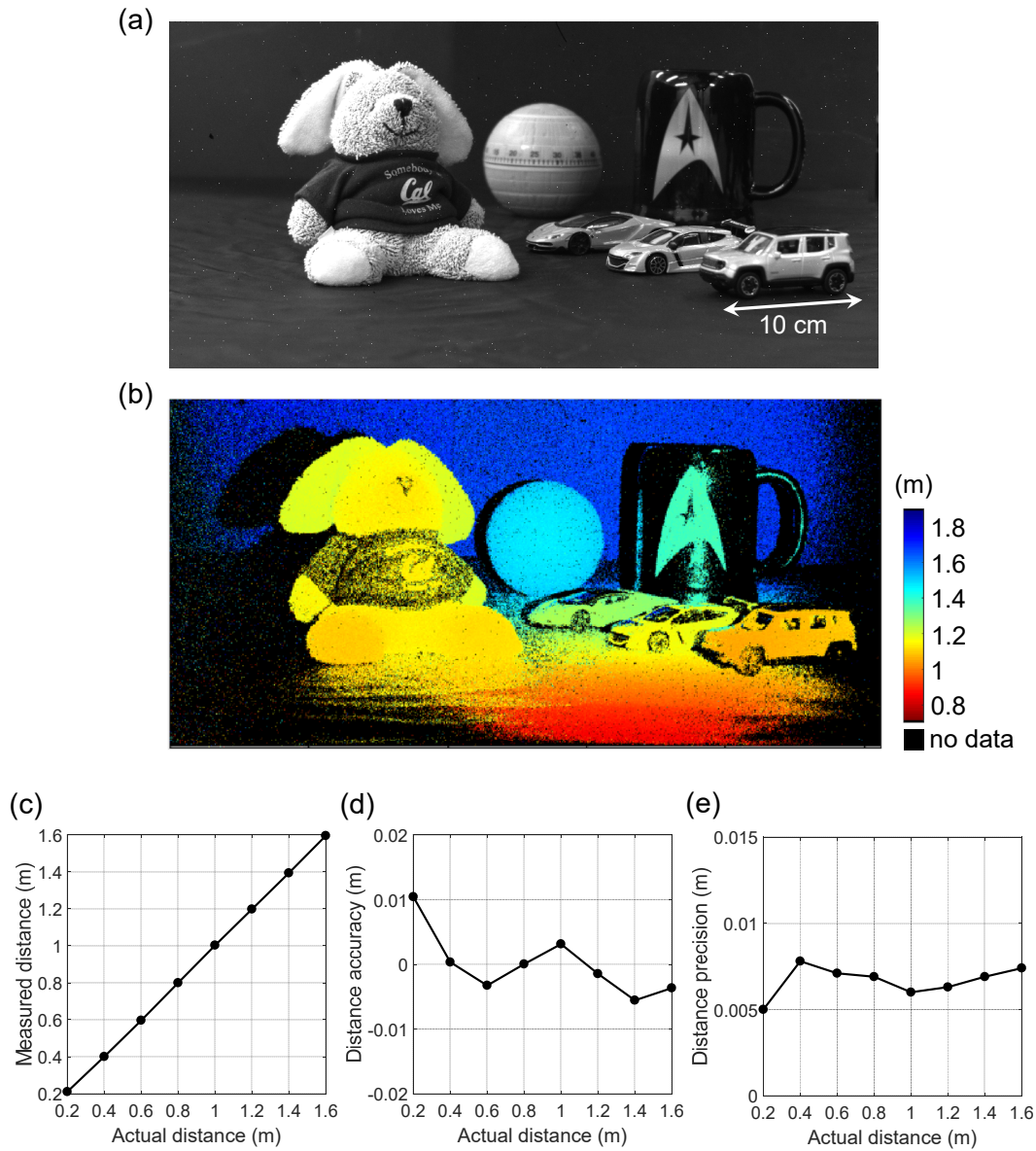(m)

1.8
1.6
1.4
1.2
1
0.8
no data

(c)

(d)

(e)

Figure 6.13 – Measured results for time-gated ToF ranging: (a) real-life 2D intensity image; (b) color-coded 3D image of the same scene obtained with time-gated ToF; (c) measured distance vs actual distance; (d) measured distance accuracy vs actual distance; (e) measured distance precision vs actual distance.

red color denotes higher proximity to the SPAD camera, whereas blue color corresponds to higher distance. The maximum depth range for this measurement was set to 2 m, but it can be extended to tens of meters by lowering the laser repetition frequency and increasing the gate step. Note that some pixels in the scene show "no data" (black) due to laser speckle patterns, leading to locally insufficient photon counts for distance estimation. The fine gate scanning pitch and long exposure are used to achieve high depth precision, and the resulting data acquisition time for this measurement was few tens of seconds. This is considerably longer than that of other ranging methods such as i-ToF, but it can be readily reduced by increasing the gate scanning pitch, reducing the scanning range, and increasing the laser power to reduce the exposure time. In addition, further improvement is expected by implementing on-chip microlens to boost the sensitivity.

Fig. 6.13(c) shows the measured distance as a function of the actual object distance. In Fig. 6.13(c), (d) and (e), a flat object covered with white paper (reflectance around 60%) is used to evaluate the measured distance, accuracy and precision. In Fig. 6.13(c), the measured distance is extracted by taking the average of the single pixel distance over 20×20 pixels at the center of the array. A very good agreement with the actual distance is observed within the measured range from 0.2 to 1.6 m. In Fig. 6.13(d), the distance accuracy is calculated as the averaged measured distance subtracted by the actual distance. For the measured distance range, the accuracy is always better than 1cm. In Fig. 6.13(e), distance precision is exploited as a standard deviation of the single pixel distance over 20×20 pixels in the center of the array. The precision is better than 7.8 mm (rms) for all the measured points up to 1.6 m.

6.5.6   Multi-object 3D imaging

Compared to i-ToF [34, 35, 36], d-ToF has the advantage that spatially overlapped multiple reflective objects can be imaged individually and accurately. Multi-object detection has been experimentally demonstrated in SPAD-based d-ToF sensors [37, 38], where power- and area-consuming TDC circuits and large computational cost for histogramming severely limited the size of the detector. Multi-object detection has also been demonstrated by either coding temporal illumination or exposure patterns [39, 40], which involves a large computational cost to recover 3D images.

A time-gated ToF sensor provides an alternative, scalable solution by means of compact pixel circuitry and less complicated computation. A CMOS-based time-gating scheme has been adopted for multi-object detection in a 160×120-pixel array [41], where, however, the readout noise limits the lower bound of detectable signal level for each pixel. The readout noise represents a critical issue for scaling the array size because smaller pixel size and larger pixel array size result in the reduced number of reflected photons per pixel, severely limiting SNR. Our gated SPAD pixel enables scalable and readout-noise-free single-photon time gating for multi-object detection.

Fig. 6.14(a) shows the experimental setup: 510 nm-laser beam pulsed at 40 MHz (average power: 2 mW, optical pulse width: 130 ps FWHM, PicoQuant GmbH, Berlin, Germany) is spread by a diffuser and used to illuminate a spherical target. The SPAD camera is synchro-
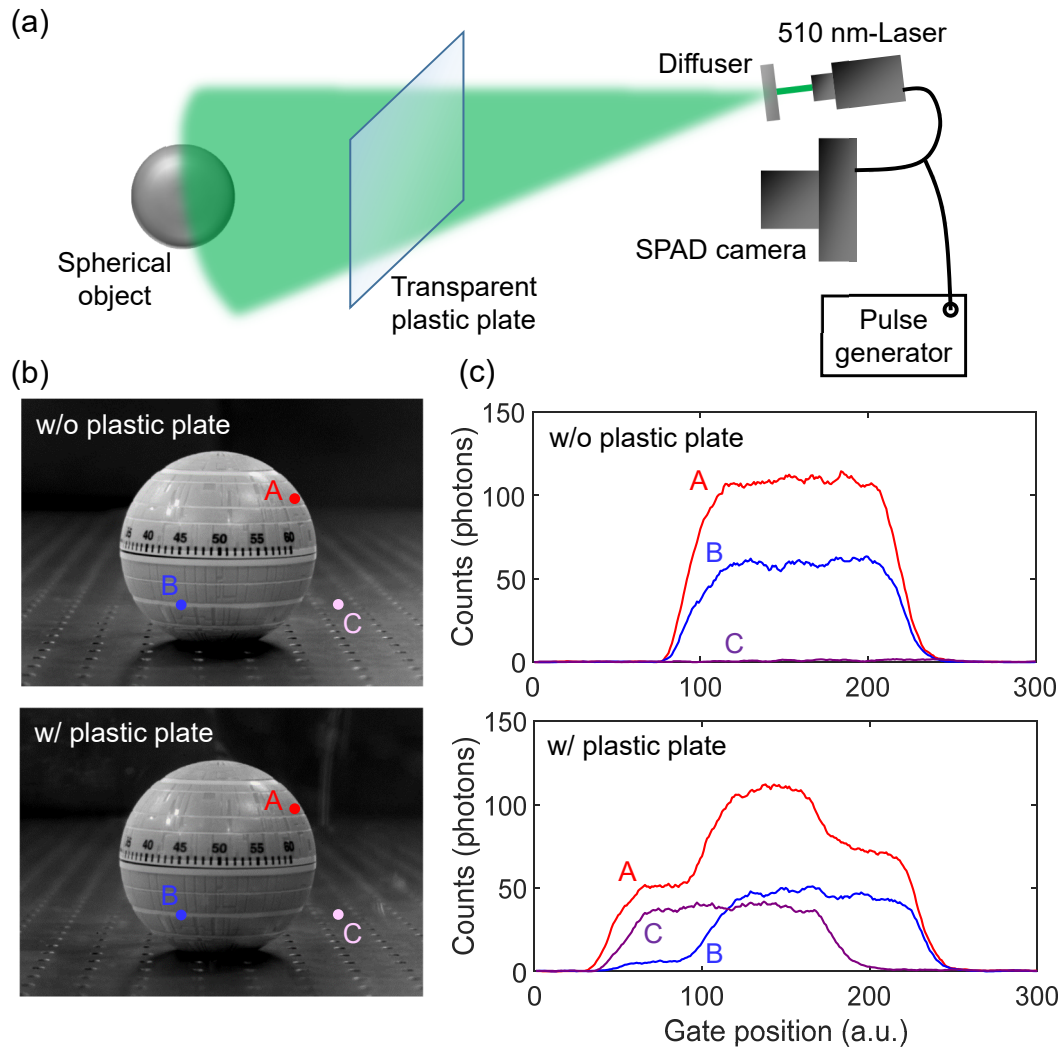
Figure 6.14 – Experimental setup and measured results for time-gated ToF under multiple reflections. (a) Experimental setup to perform the multi-object detection. (b) Captured 2D images with and without plastic plate. (c) Measured photon count profiles for three different pixels, with and without plastic plate.
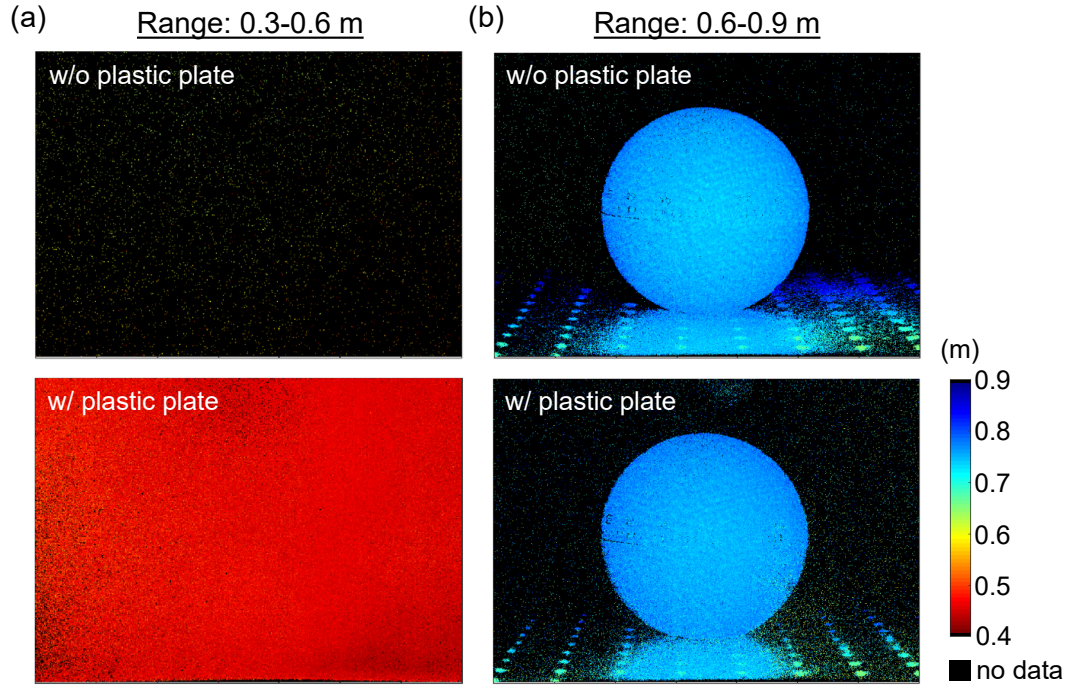
Figure 6.15 – Reconstructed 3D images in the multi-object detection experiment. (a) 3D images reconstructed based on the distance range of 0.3-0.6 m (central 700×500 pixels cropped). Black color indicates that no laser reflection is detected in the measured range. (b) 3D images reconstructed based on the distance range of 0.6-0.9 m.

nized with the laser triggering signal, and a transparent plastic plate is inserted between the camera and the object. The distances from the camera to the plastic plate and the object are 0.45 m, and 0.75 m, respectively. Fig. 6.14(b) shows 2D intensity images under indoor lighting with and without the plastic plate inserted. Since the plate is almost transparent, no significant difference is observed in the 2D images for those two cases.

The measured time-gating profiles for three representative points (A, B and C) are plotted in Fig. 6.14(c). Without the plate, the time-gating profiles for point A and B show only a single smoothed rectangular function waveform with its rising edge around gate position 100 (one step of the position corresponding to 36 ps). For point C, the photon count stays close to zero over the measured gate position range, indicating no reflective object is detected at this pixel. With the plastic plate, by contrast, the profile at point A shows two-step rising edges around gate positions 40 and 100. Given that the measured profile of photon counts is a convolution of a single smoothed rectangular function and the reflected photon intensity distribution, the two-step profile is a convincing evidence of double reflection from the plastic plate and the spherical object. Similar behavior is observed at point B, where the slope of the first rising edge around gate position 40 is milder than that of point A. The profile at point C shows only single rising edge around gate position 40, corresponding to the reflection from the plastic plate. The variation of the slope for the rising edge around gate position 40 between different

points is induced by the non-uniform reflection from the surface of the plastic plate.

Fig. 6.15 shows the reconstructed 3D images based on time-gated ToF. The photon counting profile for each pixel is analyzed to extract the position of rising edges. The rising edge is searched by defining a virtual gate window containing 60 data points of the measured intensity profile. The window is scanned over the whole gate position in a non-overlapping fashion, and the existence or non-existence of a rising edge in the virtual window is determined for each scanning position. Fig. 6.15(a) shows the estimated local distance within the range of 0.3 to 0.6 m. Black pixels represent no object detected within the range. Without plastic plate, the vast majority of the pixels shows no detection (black), while the majority of the pixels indicates a reflection at 0.45 m (dark red) with the plastic plate, which is consistent with its actual position. Fig. 6.15(b) shows the estimated distance within the range of 0.6 to 0.9 m. For both cases, the distance map of the spherical target object is reconstructed precisely. The measured target object distance is approximately 0.75 m, which is also consistent with the actual distance.

The results demonstrate the capability of our time-gated SPAD camera to perform spatially overlapped multi-object detection. Note that the proposed scheme can be applied to the detection of more than two reflection peaks. Finer scanning of the virtual gate window in post-processing enables systematic detection of multiple peaks. The minimum resolvable distance between two neighboring reflective materials is fundamentally limited by the finite rising or falling time of the gate window profile, corresponding to 5-10 cm in this SPAD sensor.

6.6 Conclusion

In this chapter, a 1 Mpixel time-gated SPAD image sensor is reported for the first time. The sensor is applied to high dynamic range 2D imaging and high spatio-temporal resolution 3D imaging. To the best of our knowledge, the spatially overlapped multi-object detection with single-photon time-gating scheme has been experimentally demonstrated for the first time. Fig. 6.16 shows a state-of-the-art comparison of SPAD pixel pitch and array size. The array size of our sensor is the largest, almost 4 times higher than that of the state-of-the-art sensor [5], while the pixel pitch is one of the smallest. A more detailed comparison is summarized in Table 6.1. Median DCR is the lowest among other works thanks to the optimized process and miniaturized active size. The lower fill factor and PDE compared to the prior art and other single-photon detector technologies (EMCCD, ICCD, QIS) are due to the FSI configuration; these figures of merit can be further improved, typically by a factor of 2 to 10, by introducing on-chip microlenses [42, 43, 44]. Owing to its noise and dynamic range performance, the proposed sensor will be useful in a wide variety of industrial applications such as security, automotive, robotic, biomedical, and scientific applications, including quantum imaging and ultra-high-speed imaging.
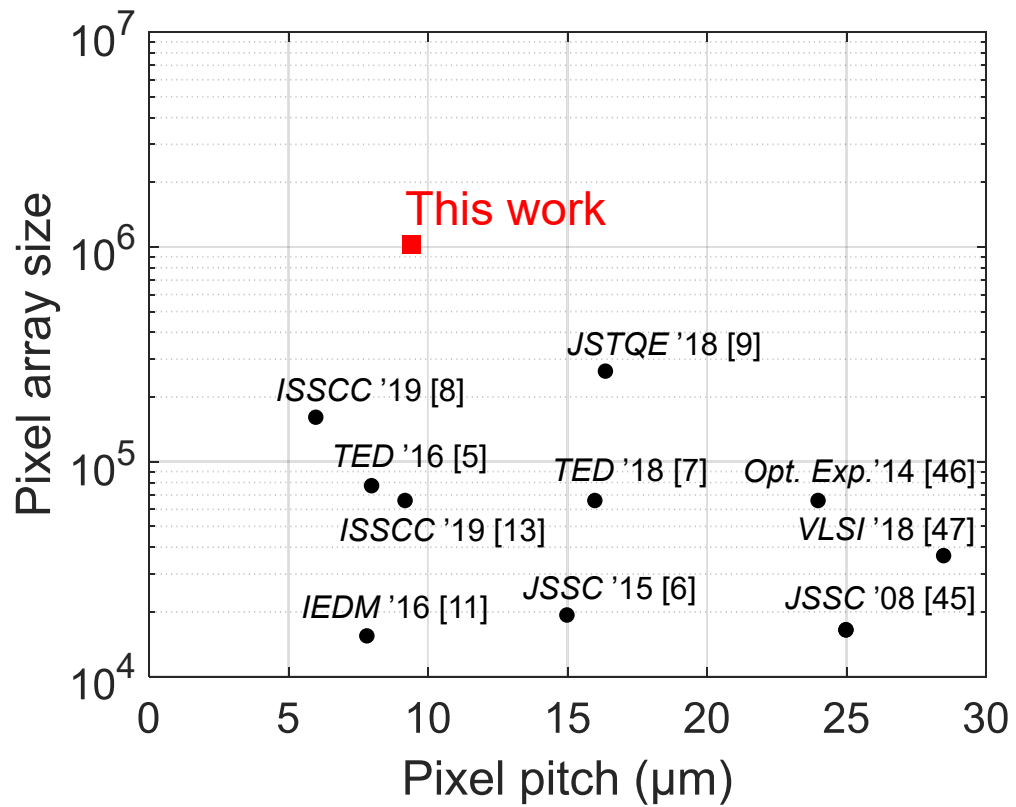
Figure 6.16 – State-of-the-art comparison of pixel array size and pixel pitch in SPAD sensors, based on published works as of December, 2019.

Table 6.1 – State-of-the-art comparison of performance and specifications in large-scale SPAD arrays, based on published works as of December, 2019.

	[6]	[5]	[13]	[9]	[8]	This work (pixel A/B)
Process technology	350 nm HV CMOS	130 nm CIS	40/90 nm 3D-BSI	180 nm CMOS	65 nm CIS	180 nm CMOS
Chip size (mm^2)	3.42×3.55	3.4×3.1	-	9.5×9.6	-	11×11
Sensor resolution	160×120	320×240	256×256	512×512	400×400	1024×1000
Pixel pitch (μm)	15	8	9.2	16.38	6	9.4
Fill factor (%)	21	26.8	51	10.5	70	7.0/13.4
Pixel output bit depth	5.4b	1b	14b	1b	1b	1b
No. of pixel transistors	8	9	>600	11	4	7/5.75
Median DCR (cps)	580 (V_{ex} = 3 V)	47 (V_{ex} = 1.5 V)	20 (V_{ex} = 1.5 V)	7.5 (V_{ex} = 6.5 V)	100 (-)	0.4/2.0 (V_{ex} = 3.3 V)
Max. PDP (%)	-	39.5 (V_{ex} = 1.5 V)	23 (V_{ex} = 3 V)	50 (V_{ex} = 6.5 V)		10.5/26.7 (V_{ex} = 3.3 V)
Max. PDE (%)	-	10.6 (V_{ex} = 1.5 V)	11.7 (V_{ex} = 3 V)	5.25 (V_{ex} = 6.5 V)		0.7/3.6 (V_{ex} = 3.3 V)
Crosstalk (%)	-	-	-	-	-	0.17/0.39 (V_{ex} = 3.3 V)
Min. gate length (ns)	0.75	-	-	5.75	-	3.8
Frame rate (fps)	486 (5.4b)	16,000 (1b)	-	97,700 (1b)	60 (-)	24,000 (1b)
Power dissipation (W)	0.1567	-	0.0776	0.0267	-	0.284/0.535

References

[1] E. Charbon, "Will Avalanche Photodiode Arrays Ever Reach 1 Megapixel?," Int. Image Sensor Workshop (2007).

[2] M. O'Toole, D.B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on light-cone transform," Nature 555(7696), 338 (2018).

[3] H. Defienne, M. Reichert, J. W. Fleischer, and D. Faccio, "Quantum image distillation," Science Advances 5(10), eaax0307 (2019).

[4] A. Lyons, F. Tonolini, A. Boccolini, A. Repetti, R. Henderson, Y. Wiaux, and D. Faccio, "Computational time-of-flight diffuse optical tomography," Nat. Photonics 13, 575-579 (2019).

[5] N. A. W. Dutton, I. Gyongy, L. Parmesan, S. Gnecchi, N. Calder, B. R. Rae, S. Pellegrini, L. A. Grant, and R. K. Henderson, "A SPAD-Based QVGA Image Sensor for Single-Photon Counting and Quanta Imaging," IEEE Trans. Electron Devices 63(1), 189-196 (2016).

[6] M. Perenzoni, N. Massari, D. Perenzoni, L. Gasparini, and D. Stoppa, "A 160×120-Pixel Analog-Counting Single-Photon Imager with Time-Gating and Self-Referenced Column-Parallel A/D Conversion for Fluorescence Lifetime Imaging," IEEE J. Solid-State Circuits 51(1), 155-167 (2016).

[7] I. Gyongy, N. Calder, A. Davies, N. A. W. Dutton, R. R. Duncan, C. Rickman, P. Dalgamo, and R. K. Henderson, "A 256×256, 100-kfps, 61% Fill-Factor SPAD Image Sensor for Time-Resolved Microscopy Applications," IEEE Trans. Electron Devices 65(2), 547-554 (2018).

[8] Y. Hirose, S. Koyama, T. Okino, A. Inoue, S. Saito, Y. Nose, M. Ishii, S. Yamahira, S. Kasuga, M. Mori, T. Kabe, K. Nakanishi, M. Usuda, A. Odagawa, and T. Tanaka, "A 400×400-Pixel 6μm-Pitch Vertical Avalanche Photodiodes CMOS Image Sensor Based on 150ps-Fast Capacitive Relaxation Quenching in Geiger Mode for Synthesis of Arbitrary Gain Images," IEEE Int. Solid-State Circuits Conference (2019).

[9] A. C. Ulku, C. Bruschini, I. M. Antolovic, Y. Kuo, R. Ankri, S. Weiss, X. Michalet, and E. Charbon, "A 512×512 SPAD Image Sensor with Integrated Gating for Widefield FLIM," IEEE J. Selected Topics in Quantum Electronics 25(1), 1-12 (2018).

References

[10] J. M. Pavia, M. Scandini, S. Lindner, M. Wolf, and E. Charbon, "A 1 × 400 Backside-Illuminated SPAD Sensor With 49.7 ps Resolution, 30 pJ/Sample TDCs Fabricated in 3D CMOS Technology for Near-Infrared Optical Tomography," IEEE J. Solid-State Circuits 50(10), 2406-2418 (2015).

[11] T. Al Abbas, N. A. W. Dutton, O. Almer, S. Pellegrini, Y. Henrion, and R. K. Henderson, "Backside illuminated SPAD image sensor with 7.83 μm pitch in 3D-stacked CMOS technology," IEEE Int. Electron Devices Meeting, 811-814 (2016).

[12] A. R. Ximenes, P. Padmanabhan, M.-J. Lee, Y. Yamashita, D.-N. Yaung, and E. Charbon, "A Modular, Direct Time-of-Flight Depth Sensor in 45/65-nm 3-D-Stacked CMOS Technology," IEEE J. Solid-State Circuits 54(11), 3203-3214 (2019).

[13] R. K. Henderson, N. Johnston, S. W. Hutchings, I. Gyongy, T. Al Abbas, N. Dutton, M. Tyler, S. Chan, and J. Leach, "A 256×256 40nm/90nm CMOS 3D-Stacked 120dB Dynamic-Range Reconfigurable Time-Resolved SPAD Imager," IEEE Int. Solid-State Circuits Conference (2019).

[14] E. R. Fossum, "What to do with sub-diffraction-limit (SDL) pixels? A proposal for a gigapixel digital film sensor (DFS)," in IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors, 214-217 (2005).

[15] E. R. Fossum, J. Ma, S. Masoodian, L. Anzagira, and R. Zizza, "The quanta image sensor: every photon counts," Sensors 16, 1260 (2016).

[16] F. Yang, Y. M. Lu, L. Sbaiz, and M. Vetterli, "Bits From Photons: Oversampled Image Acquisition Using Binary Poisson Statistics," IEEE Trans. Image Processing, 21(4), 1421-1436 (2012).

[17] J. Ma, S. Masoodian, D. A. Starkey, and E. R. Fossum, "Photon-number-resolving megapixel image sensor at room temperature without avalanche gain," Optica 4(12), 1474-1481 (2017).

[18] A. Gnanasambandam, O. Elgendy, J. Ma, and S. H. Chan, "Megapixel photon-counting color imaging using quanta image sensor," Opt. Express 27(12), 17298 (2019).

[19] J. Richardson, R. Walker, L. Grant, D. Stoppa, F. Borghetti, E. Charbon, M. Gersbach, and R. K. Henderson, "A 32×32 50ps Resolution 10 bit Time to Digital Converter Array in 130nm CMOS for Time Correlated Imaging," IEEE Custom Integrated Circuits Conference (2009).

[20] C. Veerappan, J. Richardson, R. Walker, D.-U. Li, M. W. Fishburn, Y, Maruyama, D. Stoppa, F. Borghetti, M. Gersbach, R. K. Henderson, and E. Charbon, "A 160×128 single-photon image sensor with on-pixel 55ps 10b time-to-digital converter," IEEE Int. Solid-State Circuits Conference (2011).

[21] A. Carimatto, S. Mandai, E. Venialgo, T. Gong, G. Borghi, D. R. Schaart, and E. Charbon, "A 67,392-SPAD PVTB-compensated multi-channel digital SiPM with 432 column-parallel

48ps 17b TDCs for endoscopic time-of-flight PET," IEEE Int. Solid-State Circuits Conference (2015).

[22] M. Perenzoni, D. Perenzoni, and D. Stoppa, "A 64×64-pixel digital silicon photomultiplier direct ToF sensor with 100Mphotons/s/pixel background rejection and imaging/altimeter mode with 0.14% precision up to 6km for spacecraft navigation and landing," IEEE Int. Solid-State Circuits Conference (2016).

[23] R. K. Henderson, N. Johnston, F. M. D. Rocca, H. Chen, D. D.-U. Li, G. Hungerford, R. Hirsch, D. Mcloskey, P. Yip, and D. J. S. Birch, "A 192×128 Time Correlated SPAD Image Sensor in 40-nm CMOS Technology," IEEE J. Solid-State Circuits 54(7), 1907-1916 (2019).

[24] C. Veerappan, and E. Charbon, "A Low Dark Count p-i-n Diode Based SPAD in CMOS Technology," IEEE Trans. Electron Devices 63(1), 65-71 (2015).

[25] C. Niclass, A. Rochas, P.-A. Besse, and E. Charbon, "Design and Characterization of a CMOS 3-D Image Sensor Based on Single Photon Avalanche Diodes," IEEE J. Solid-State Circuits 40(9), 1847-1854 (2005).

[26] J. A. Richardson, E. A. G. Webster, L. A. Grant, and R. K. Henderson, "Scaleable Single-Photon Avalanche Diode Structures in Nanometer CMOS Technology," IEEE Trans. Electron Devices 58(7), 2028-2035 (2011).

[27] N. Teranishi, "Dark current and white blemish in image sensors," IEEE Int. Symposium on VLSI technology, Systems and Applications (2013).

[28] Y. Xu, P. Xiang, and X. Xie, "Comprehensive understanding of dark count mechanisms of single-photon avalanche diodes fabricated in deep sub-micron CMOS technologies," Solid-State Electronics 129, 168-174 (2017).

[29] D. Bronzi, F. Villa, S. Bellisai, B. Markovic, S. Tisa, A. Tosi, F. Zappa, S. Weyers, D. Durini, W. Brockherde, and U. Paschen., "Low-noise and large-area CMOS SPADs with timing response free from slow tails," in Proc. Eur. Solid-State Device Res. Conf. (ESSDERC), 230–233 (2012).

[30] F. Acerbi, G. Paternoster, A. Gola, N. Zorzi, and C. Piemonte, "Silicon photomultipliers and single-photon avalanche diodes with enhanced NIR detection efficiency at FBK," Nuclear Inst. And Methods in Physics Research, A (2017).

[31] E. R. Fossum, "Modeling the Performance of Single-Bit and Multi-Bit Quanta Image Sensors," IEEE J. Electron Devices Society 1(9), 166 (2013).

[32] N. A. W. Dutton, T. Al Abbas, I. Gyongy, F. M. D. Rocca, and R. K. Henderson, "High Dynamic Range Imaging at the Quantum Limit with Single Photon Avalanche Diode-Based Image Sensors," Sensors 18(4), 1166 (2018).

[33] I. M. Antolovic, C. Bruschini, and E. Charbon, "Dynamic range extension for photon counting arrays," Opt. Express 26(17), 22234-22248 (2018).

References

[34] S. Kawahito, I. A. Halin, T. Ushinaga, T. Sawada, M. Homma, and Y. Maeda, "A CMOS Time-of-Flight Range Image Sensor With Gates-on-Field-Oxide Structure," IEEE Sensors Journal 7(12), 1578-1586 (2007).

[35] W. van der Tempel, A. Ercan, T. Finateu, K. Fotopoulou, C. Mourad, F. Agavriloaie, S. Resimont, L. Cutrignelli, P. Thury, C. E. Medina, S. Xiao, J.-L. Loheac, J. Perhavc, T. Van de Hauwe, V. Belokonskiy, L. Bossuyt, W. Aerts, M. Pauwels, and D. van Nieuwenhove, "A 320×240 10μm CAPD ToF image sensor with improved performance," Int. Image Sensor Workshop (2017).

[36] C. S. Bamji, S. Mehta, B. Thompson, T. Elkhatib, S. Wurster, O. Akkaya, A. Payne, J. Godbaz, M. Fenton, V. Rajasekaran, L. Prather, S. Nagaraja, V. Mogallapu, D. Snow, R. McCauley, M. Mukadam, I. Agi, S. McCarthy, Z. Xu, T. Perry, W. Qian, V.-H. Chan, P. Adepu, G. Ali, M. Ahmed, A. Mukherjee, S. Nayak, D. Gampell, S. Acharya, L. Kordus, and P. O'Connor, "1Mpixel 65nm BSI 320MHz Demodulated TOF Image Sensor with 3.5μm Global Shutter Pixels and Analog Binning," IEEE Int. Solid-State Circuits Conference (2018).

[37] D. Shin, F. Xu, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "Computational multi-depth single-photon imaging," Opt. Express 24(3), 1873-1888 (2016).

[38] J. Haase, M. Beer, J. Ruskowski, and H. Vogt, "Multi object detection in direct Time-of-Flight measurements with SPADs," 14th Conference on Ph.D. Research in Microelectronics and Electronics (2018).

[39] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded Time of Flight Cameras: Sparse Deconvolution to Address Multipath Interference and Recover Time Profiles," ACM Transactions on Graphics 32(6) (2013).

[40] F. Mochizuki, K. Kagawa, R. Miyagi, M.-W. Seo, B. Zhang, T. Takasawa, K. Yasutomi, and S. Kawahito, "Separation of Multi-path Components in Sweep-less Time-of-flight Depth Imaging with a Temporally-compressive Multi-aperture Image Sensor," ITE Transactions on Media Technology and Applications 6(3), 202-211 (2018).

[41] K. Kitano, T. Okamoto, K. Tanaka, T. Aoto, H. Kubo, T. Funatomi, and Y. Mukaigawa, "Recovering temporal PSF using ToF camera with delayed light emission," IPSJ Trans. Computer Vision and Applications 9(15) (2017).

[42] J. M. Pavia, M. Wolf, and E. Charbon, "Measurement and modeling of microlenses fabricated on single-photon avalanche diode arrays for fill factor recovery," Opt. Express 22(4), 4202-4213 (2014).

[43] I. Gyongy, A. Davies, B. Gallinet, N. A. W. Dutton, R. Duncan, C. Rickman, R. K. Henderson, and P. A. Dalgarno, "Cylindrical microlensing for enhanced collection efficiency of small pixel SPAD arrays in single-molecule localisation microscopy," Opt. Express 26(3), 2280-2291 (2018).

[44] I. M. Antolovic, A.C. Ulku, E. Kizilkan, S. Lindner, F. Zanella, R. Ferrini, M. Schnieper, E. Charbon, and C. Bruschini, "Optical-stack optimization for improved SPAD photon detection efficiency," in Proc. SPIE, Quantum Sensing and Nano Electronics and Photonics XVI, San Francisco, CA, USA (2019).

[45] C. Niclass, C. Favi, T. Kluter, M. Gersbach, and E. Charbon, "A 128 × 128 Single-Photon Image Sensor With Column-Level 10-Bit Time-to-Digital Converter Array," IEEE J. Solid-State Circuits 43(12), 2977-2989 (2008).

[46] S. Burri, Y. Maruyama, X. Michalet, F. Regazzoni, C. Bruschini, and E. Charbon, "Architecture and applications of a high resolution gated SPAD image sensor," Opt. Express 22(14), 17573-17589 (2014).

[47] S. Lindner, C. Zhang, I. M. Antolovic, M. Wolf, and E. Charbon, "A 252 × 144 SPAD pixel FLASH LiDAR with 1728 Dual-clock 48.8 ps TDCs, Integrated Histogramming and 14.9-to-1 Compression in 180nm CMOS Technology," IEEE Int. Symposium on VLSI Circuits (2018).

7 4-dimensional light-in-flight imaging based on megapixel time-gated single-photon camera

Light-in-flight imaging is one of the attractive applications of time-resolved SPAD image sensors. Previous works have demonstrated the light-in-flight observation based on SPAD cameras with up to 1000 pixels. Such a low resolution camera requires complex data processing and spatial interpolation to get acceptable image quality, and is not suitable for precise tracking of spatio-temporal light propagation as well as fast acquisition and rendering of the light-in-flight videos. To address this issue, we employed the megapixel time-gated SPAD image sensor presented in Chapter 6 to achieve extremely high spatio-temporal resolution in this application. This chapter is based on results presented in, K. Morimoto et al. "Superluminal motion-assisted 4-dimensional light-in-flight imaging," Submitted to Light Sci. Appl. (2020).

7.1 Challenges in light-in-flight imaging

Progress in high speed imaging techniques has enabled observation and recording of light propagation dynamics in free space as well as in transparent, translucent and scattering media. Various approaches for capturing the light-in-flight have been demonstrated: holographic techniques [1, 2, 3], photonic mixers [4], time-encoded amplified imaging [5], streak cameras [6, 7] CCDs [8] and silicon CISs [9]. In recent years, 1- and 2-dimensional arrays of SPAD have been adopted for the light-in-flight imaging systems [10, 11, 12, 13, 14, 15]. These sensors boost data acquisition speed by employing pixel-parallel detection of time stamping with picosecond time resolution and single-photon sensitivity. However, the sensors are capable of sampling only 3-dimensional spatio-temporal information (x, y, t) of the target light, and hence the tracking of light-in-flight breaking outside a xy-plane is a challenge.

A method to reconstruct an extra-dimensional position information, z, from measured spatio-temporal information has been recently investigated by researchers [11]. The authors remark that different propagation angle with respect to xy-plane results in the different apparent velocity of light. Thus, comparing the measured spatio-temporal data set with this theory could give an estimation of z-component in the light propagation vector. Yet, the analysis is

limited to a simplified case with single straight light paths within xz-plane passing a fixed point, and complete 4-dimensional reconstruction of light-in-flight in arbitrary paths remains to be verified. In addition, spatial resolution of the detector array is critical for accurate reconstruction of the 4-dimensional trace; the reported SPAD sensor resolutions, e.g. 64×32 [15], 32×32 [10] and 256×1 [13] pixels, are not sufficient, whereas significant improvement of the estimation error for (x, y, z, t) is expected with orders of magnitude larger arrays, typically in the megapixels. CISs routinely achieve megapixel resolutions, e.g. 0.3 Mpixel resolution is reported for high speed imaging [9], however the necessary timing resolution is not reached and it is limited to 10 nanoseconds.

In this chapter we demonstrate the 4-dimensional light-in-flight imaging based on the first time-gated megapixel SPAD camera [17]. In contrast to conventional TCSPC approaches, where power- and area-consuming TDC circuits restrict the scaling of the pixel array [18, 19, 20], our time-gating approach [17] achieves a much more compact pixel circuit, suitable for large-scale arrays. Our camera achieves a megapixel format, while, at the same time ensuring a time resolution comparable to that of a TDC. We introduce a theoretical equation to fit the measured data set in 3-dimensional space. 4-dimensional point cloud is computationally reproduced without prior knowledge, exhibiting a good agreement with the actual light propagation vectors.

7.2   Principle and experimental setup

To illustrate our method, we introduce a Minkowski space that is describing the spatio-temporal propagation of a laser pulse and scattered photons towards fixed observation points. Fig. 7.1(a) shows a schematic illustration of light-in-flight observation in the Minkowski space at $y = 0$. A straight light propagation along x-axis is depicted as (x, y, z) = (ct, 0, 0), where c is the speed of light. In the Minkowski space, this equation corresponds to a single tilted line crossing the origin (purple arrow in Fig. 7.1(a)). A camera is located at the position C_i ($i = $ 1,2,...,5). The observation of the light-in-flight at the camera is mediated by scattered light at each point on the path (colored dots on the purple arrow). The propagation of the scattered light from the points are described as light cones. Here we define t as the time at which the light propagating from a laser source reaches a point (x, y, z), called scattering point on the light path, and t' as a time at which the light scattered by the medium at (x, y, z) reaches the camera. Due to the finite speed of light, a time difference from t to t' is proportional to the spatial distance between the scattering point and the camera location. In Minkowski space, the observation time t' can be visualized as an intersection point of the light cone and a vertical line L passing C_i (dotted line). The relation between the light position x and the corresponding observation time t' is shown in Fig. 7.1(b). The behavior is highly dependent on the camera position with respect to the scattering point, i.e. at certain points it appears to be faster than in others. For example, at C_1 where the light is coming towards the camera, the scattering light from all the points on the light path reaches the camera at the same time.
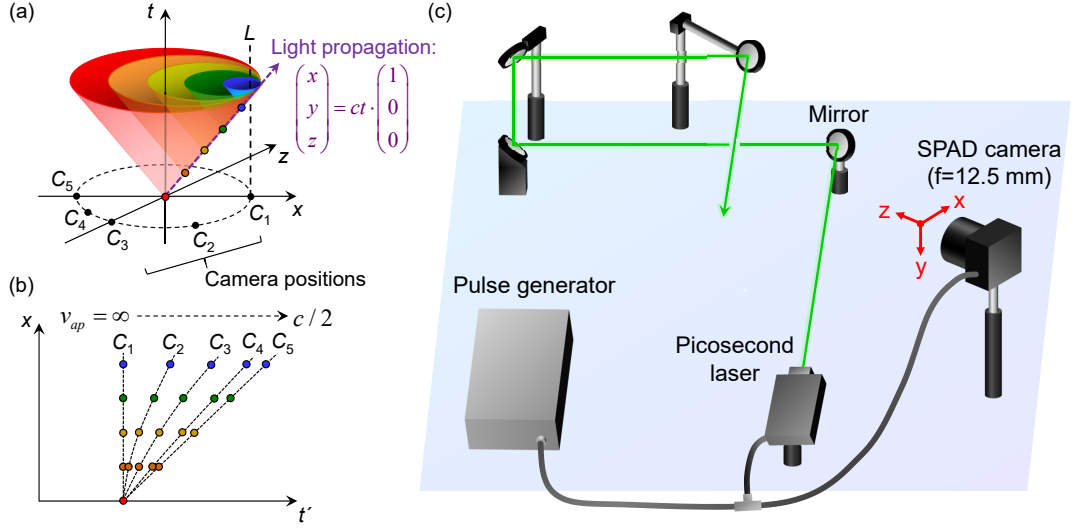
Figure 7.1 – Principe and experimental setup for 4-dimensional light-in-flight imaging. (a) Schematic illustration of light propagation and scattering in Minkowski space. Propagating light is depicted as a purple dashed arrow, and scattered light is shown as a light cone. (b) Schematic plot of t' dependence of x. Slope of the dashed line indicates the apparent velocity. (c) Experimental setup for light-in-flight imaging. Picosecond laser and SPAD camera are both controlled by a pulse generator. When performing the light-in-flight measurements, all the mirrors are confined by a large transparent acrylic box (not shown), and a fog generator forms a small amount of mist in the box to enhance the scattering of the laser in air. The origin of the Cartesian coordinate is set at the optical center of the lens of the SPAD camera.

The corresponding apparent velocity, defined as $v_{ap} = dx/dt'$, is infinite, or t' appears not to change irrespective of the distance of the scattering point (x, y, z) from the camera. At C_5, in contrast, the light propagating away from the camera is observed with the apparent velocity of $c/2$. For other scattering points corresponding to relative camera positions C_2, C_3, and C_4, one observes intermediate apparent velocities, depending on the relative position and angle of the line of sight between scattering point and camera. Hence, analyzing the apparent velocity provides an estimated light propagation vector with extra-dimensional information. The discussion can be generalized for higher dimensions; the detailed analysis of measured (x, y, t) data set and its local apparent velocity would enable a 4-dimensional reconstruction of the light-in-flight history. Note that angle-dependent apparent velocity of light can be seen in nature, for instance in a well-known astronomical phenomenon called superluminal motion. In this case, relativistic jets of matter are emitted from radio galaxies, quasars, and other celestial objects, appearing to travel faster than light to the observer. Light echoes are also known to produce superluminal motion [21, 22].

Fig. 7.1(c) is the experimental setup to verify the 4-dimensional light-in-flight imaging. A 510 nm-laser (average power: 2 mW, frequency: 40 MHz, optical pulse width: 130 ps, PicoQuant GmbH, Berlin, Germany) and the megapixel SPAD camera are synchronized by a pulse generator. The emitted laser pulses are reflected by four mirrors to construct a 3-dimensional

trajectory. The megapixel SPAD camera can be operated in intensity imaging mode to capture the background scene and in time gating mode to capture laser propagation [17].

7.3   Experimental results

7.3.1   3-dimensional light-in-flight imaging

Fig. 7.2(a) describes the principle of time gate scanning method. A 3.8 ns global gate window is synchronized to a pulsed laser source. The pixels detect only photons impinging during the gate window. The gate position can be scanned with a gate shift of 36 ps relative to the laser pulse. At each gate position, 255 binary photon-counting frames are summed to form an 8-bit image. 250 slices of 8-bit images with the scanned time-gating window position over 9 ns range are used to extract the timing information of impinging photons. The gate window is defined by sending a short electrical pulse to the gate switch in the pixels. The data acquisition time for full gate scanning (250 slices) is 15 to 30 seconds, much shorter than in previous works (e.g. hours [6, 7] to 10 minutes [10]). This can be potentially reduced further by introducing a higher power laser.

Fig. 7.2(b) shows the procedure to generate a 3-dimensional light-in-flight video. The captured 250 slices of time-gating images form a photon intensity distribution for each pixel as a function of gate position. A rising edge position τ' and "high" level of photon intensity I in the intensity profile can be extracted for each pixel. Time-of-arrival information t' over the array is obtained by subtracting independently measured pixel-position-dependent timing skew from τ'. The intensity profile is given as a convolution of the arriving photon distribution and the gating profile, and the laser pulse width appears to be elongated by 3.8 ns in the raw intensity profile. Analogously to a previously reported temporal deconvolution technique [10], the laser pulse width can be narrowed to the picosecond scale by replacing the intensity profile with a Gaussian distribution having a mean value of t', standard deviation of 150 ps (corresponding to the combined jitter of laser and detector), and integrated photon counts of I. Fig. 7.2(c) shows the selected frames of a reconstructed 3-dimensional light-in-flight video, where each frame is superimposed with a 2-dimensional intensity image independently captured by the same SPAD camera. A laser pulse starts to be observed around $t' = 0$ ns, is reflected by multiple mirrors, and goes out of sight around $t' = 5.832$ ns. The laser beam at $t' = 5.112$ ns appears to be stretched compared to the beam at $t' = 2.232$ ns. This implies an enhanced apparent velocity of the beam coming towards the camera at $t' = 5.112$ ns. Note that our system requires no mechanical laser scanning, spatial interpolation and dark noise image subtraction owing to its high spatial resolution array and low dark count rate [17].
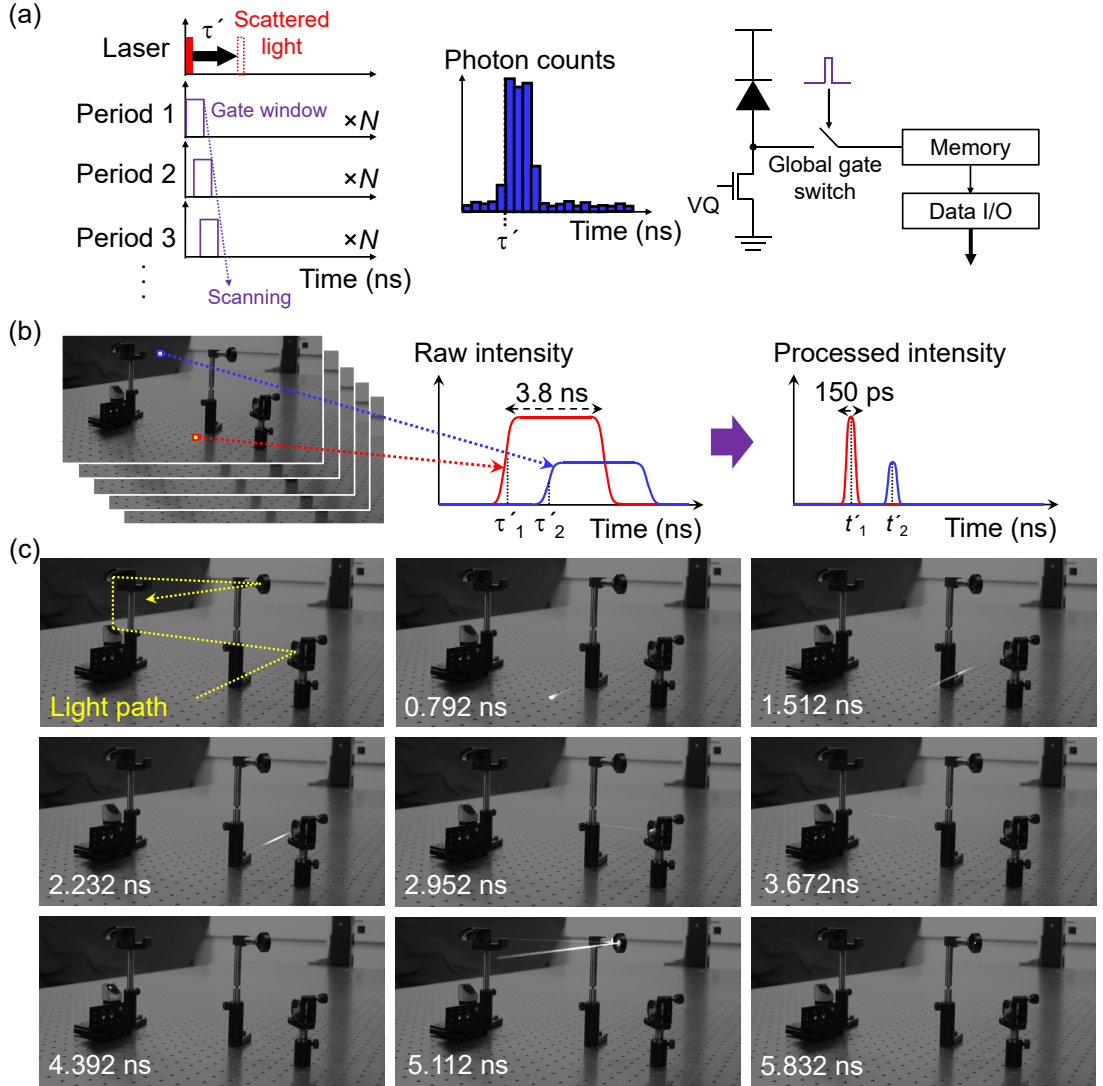
Figure 7.2 – Video frames of 3-dimensional light-in-flight imaging. (a) Principle of time gate scanning method for acquiring time-of-arrival τ' of impinging photons. (b) Data processing flow for generation of light-in-flight video. (c) Laser pulse propagates over time along the dashed line in the top-left image. The light-in-flight video is taken under dark condition, and each frame is superimposed with a 12-bit 2-dimensional intensity image independently taken with the same SPAD camera under room light. The image resolution is 1024×500.

7.3.2   Algorithm

Fig. 7.3(a) shows a color-coded plot of the observation time t' for the propagating light. To reconstruct 4-dimensional light path, the data set needs to be subdivided into straight paths. We adopted the 2D Gaussian mixture model (GMM) fitting for data clustering, which is a common unsupervised machine learning technique. Fig. 7.3(b) shows the data points separated into five clusters. Note that the data clustering is performed based on 2-dimensional spatial information (x and y), whereas the time information t' is not taken into account.

Fig. 7.3(c) depicts the coordinate system for the light-in-flight analysis. Here we derive a theoretical formula to perform the least square regression for estimation of 4-dimensional light trajectory. Time evolution of laser pulse position $\vec{r}(t)$ can be described as $\vec{r}(t) = (x(t), y(t), z(t)) = \vec{r}_0 + ct \cdot \vec{n}$, where $\vec{r}_0 = (x_0, y_0, z_0)$ is the time-independent constant vector, c the speed of light in vacuum, and $\vec{n} = (n_x, n_y, n_z)$ the normalized vector representing the direction of light propagation. Note that t is the time when the laser pulse reaches position $\vec{r}(t)$, and has an offset from
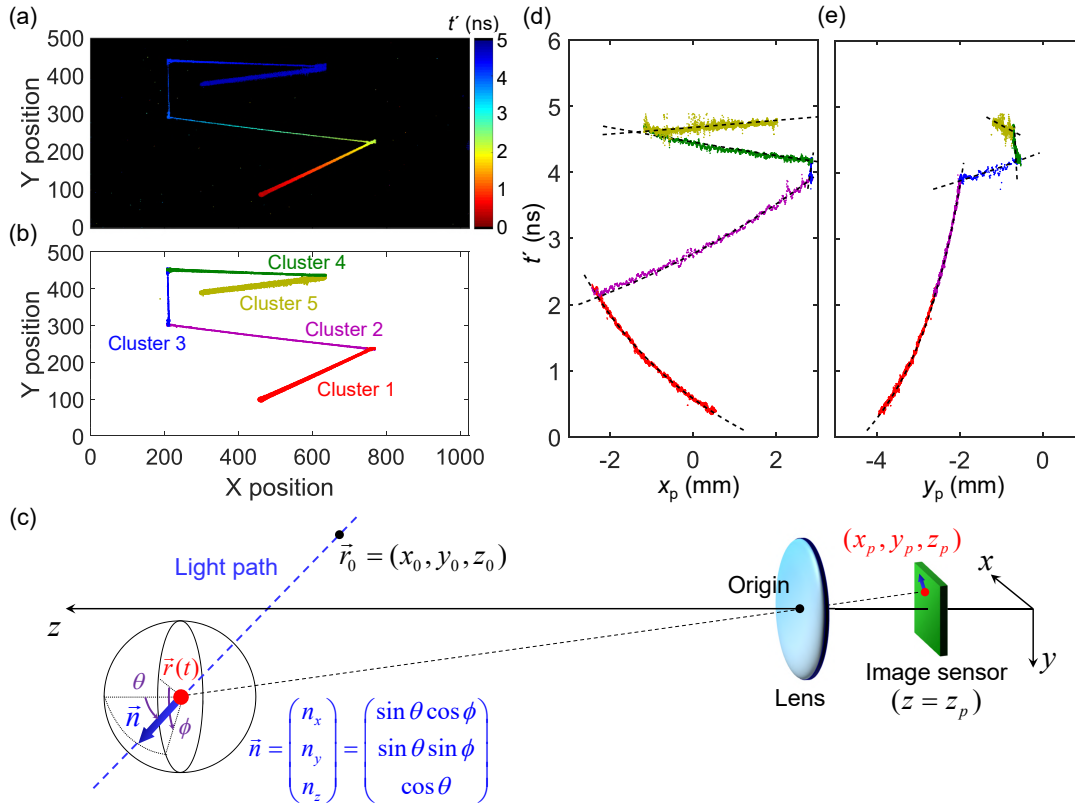
Figure 7.3 – Analysis of extracted spatio-temporal data set. (a) Color-coded plot of observation time for the light path. Red parts are observed earlier, and blue parts are observed later. Dark part has no timing data. (b) The light path data is subdivided into five data clusters of straight paths using 2D GMM fitting. (c) Observation time t' as a function of x_p, the x-position on the focal plane. Each data cluster is individually fitted using a theoretical formula. Dashed lines are the fitting curves. (d) t' as a function of y_p.

t', the time when the laser pulse at $\vec{r}(t)$ is observed at the camera position. The laser pulse position projected to the image sensor plane (focal plane) is $\vec{r}_p(t) = (x_p, y_p, z_p) = \alpha(t) \cdot \vec{r}(t)$, where $\alpha(t)$ is the time-dependent coefficient, and $-z_p$ the focal length (12.5 mm). Given that z_p is time-independent, $\alpha(t)$ can be written as $\alpha(t) = z_p/(z_0 + ct \cdot n_z)$. The movement of the laser pulse projected to the image sensor is described as:

$$x_p(t) = \frac{z_p}{z_0 + ct \cdot n_z} \cdot (x_0 + ct \cdot n_x), \; y_p(t) = \frac{z_p}{z_0 + ct \cdot n_z} \cdot (y_0 + ct \cdot n_y). \tag{7.1}$$

Considering the light propagation time from $\vec{r}(t)$ to the camera, the observation time t' can be written as:

$$t' = t + \frac{|\vec{r}(t)|}{c} = t + \frac{1}{c} \cdot \sqrt{|\vec{r}_0|^2 + 2ct(\vec{r}_0 \cdot \vec{n}) + c^2 t^2}. \tag{7.2}$$

Solving this equation will give the following:

$$t = f(t') = \frac{1}{2} \cdot \frac{c^2 t'^2 - |\vec{r}_0|^2}{c^2 t' + c(\vec{r}_0 \cdot \vec{n})}. \tag{7.3}$$

By substituting Eq. (7.3) to Eq. (7.1), the projected laser pulse position as a function of the observation time t' is expressed as:

$$x_p(t') = \frac{z_p}{z_0 + c \cdot f(t') \cdot n_z} \cdot (x_0 + c \cdot f(t') \cdot n_x), \; y_p(t') = \frac{z_p}{z_0 + c \cdot f(t') \cdot n_z} \cdot (y_0 + c \cdot f(t') \cdot n_y). \tag{7.4}$$

From the time-resolved measurement, N sets of 3-dimensional data points (x_p^i, y_p^i, t'^i) are obtained ($i = 1, 2, \ldots, N$). To reconstruct 4-dimensional light-in-flight history, the six parameters $x_0, y_0, z_0, n_x, n_y, n_z$ have to be determined by solving the following optimization problem:

$$(\vec{r}_0, \vec{n}) = \arg\min_{\vec{r}_0, \vec{n}} [\sum_i^N \{x_p^i - \frac{z_p}{z_0 + c \cdot f(t'^i) \cdot n_z} \cdot (x_0 + c \cdot f(t'^i) \cdot n_x)\}$$

$$+ \sum_i^N \{y_p^i - \frac{z_p}{z_0 + c \cdot f(t'^i) \cdot n_z} \cdot (y_0 + c \cdot f(t'^i) \cdot n_y)\}], \tag{7.5}$$

Now the following equations for optimization can be derived by converting Eq. (7.5) and Eq. (7.3), respectively, to a polar coordinate system, where $\vec{n} = (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta)$.

$$(\vec{r}_0, \theta, \phi) = \arg\min_{\vec{r}_0, \theta, \phi} [\sum_i^N \{x_p^i - \frac{z_p}{z_0 + c \cdot f(t'^i) \cdot \cos\theta} \cdot (x_0 + c \cdot f(t'^i) \cdot \sin\theta \cos\phi)\}$$

$$+ \sum_i^N \{y_p^i - \frac{z_p}{z_0 + c \cdot f(t'^i) \cdot \cos\theta} \cdot (y_0 + c \cdot f(t'^i) \cdot \sin\theta \sin\phi)\}], \tag{7.6}$$

$$f(t') = \frac{1}{2} \cdot \frac{c^2 t'^2 - (x_0^2 + y_0^2 + z_0^2)}{c^2 t' + c(x_0 \sin\theta \cos\phi + y_0 \sin\theta \sin\phi + z_0 \cos\theta)}. \tag{7.7}$$

A single straight light trajectory can be expressed as $\vec{r} = \vec{r_0} + ct \cdot \vec{n}$, defined by five fitting parameters, x_0, y_0, z_0, θ, ϕ. Those parameters are estimated by solving the above optimization problem for each data cluster.

Note that data cluster comprising of smaller number of data points could have a convergence problem during the least square regression. This can be avoided by taking a continuity assumption for the light path into consideration; starting point of the target data cluster should coincide with the end point of the previous data cluster. In practice, this assumption can be implemented by adding an extra cost function $\lambda \cdot \{(x_c - x_0 - ct_c \cdot n_x)^2 + (y_c - y_0 - ct_c \cdot n_y)^2 + (zc - z_0 - ct_c \cdot n_z)^2\}$ to Eq. (7.5), where (x_c, y_c, z_c, t_c) is the 4-dimensional coordinate for the end point of the previous data cluster, and λ the positive coefficient.

Fig. 7.3(d) and (e) show the measured observation time t' as a function of x_p and y_p, where (x_p, y_p) is the position of the pixel on the focal plane. Different colors of the data points correspond to the different data clusters. The result of fitting for each data cluster is shown as a dashed line, exhibiting a good agreement with the measured points for both x_p and y_p. Note that the fitting involves five independent parameters and is sensitive to the variation of the data. Hence, the spatio-temporal resolution of the camera is a critical factor determining the accuracy of the 4-dimensional reconstruction.

7.3.3   4-dimensional reconstruction of light-in-flight

As shown in Fig. 7.4(a), the 4-dimensional point cloud is reconstructed without prior knowledge. The time evolution t of the beam is shown as colors of points, having a certain offset from the observation time t' due to the finite traveling time from the scattering point to the camera. The fully-recovered 4-dimensional information for each point enables the visualization of the point cloud from arbitrary viewpoint apart from the actual location of the SPAD camera. In contrast to the 3-dimensional light-in-flight video where the apparent velocity of light changes as a function of propagation angle, the 4-dimensional point cloud indicates the uniform speed of light irrespective of the beam position and angle. Relatively larger variation of the data points is observed for cluster 4 and 5. This originates from the enhanced apparent velocity of light coming towards the camera, leading to the reduced fitting accuracy and increased deviation.

The accuracy of the reconstruction is evaluated in Fig. 7.4(b) and (c). The figures show the fitted light propagation angles θ and ϕ as a function of the actual geometrical angles. Five dots correspond to the fitting results of the five data clusters. All the dots are along with the ideal trend (dashed line), indicating that the fitting results are in good agreement with the actual light propagation vectors.
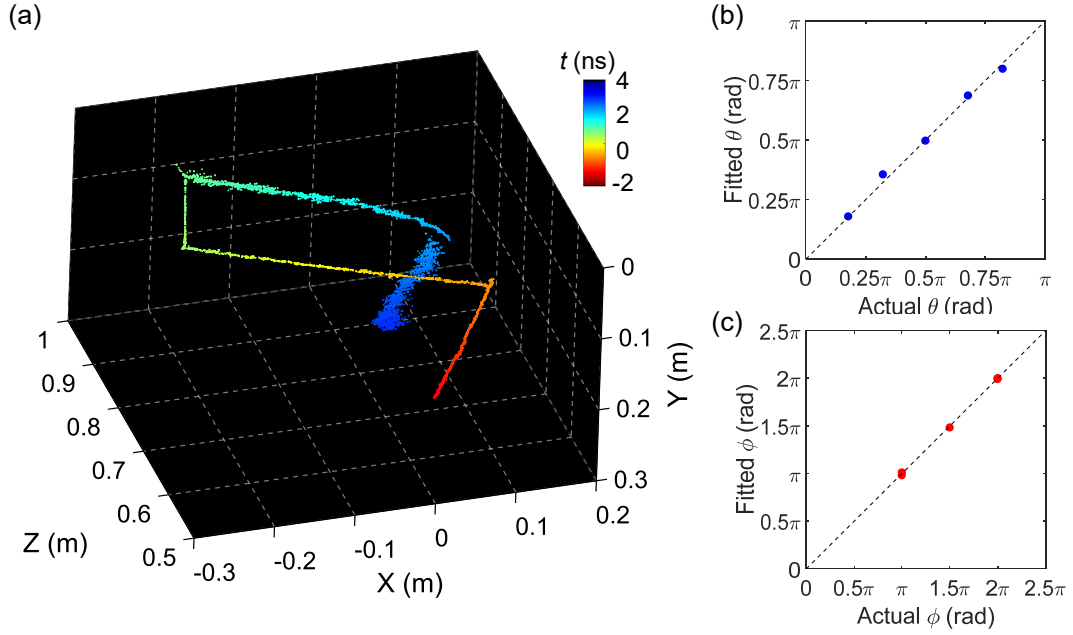
Figure 7.4 – Reconstructed 4-dimensional light-in-flight observation. (a) 4-dimensional point cloud reproduced from the measured 3-dimensional spatio-temporal data. The origin of the Cartesian coordinate system is defined at the optical center of the lens for the SPAD camera. (x, y, z, t) of each data point is reproduced using obtained fitting parameters for the corresponding data cluster. (b) (c) Comparison of fitted light propagation angles θ and ϕ, respectively, with the actual angles for five data clusters, showing a very good agreement with each other.

7.3.4 Monte Carlo simulation

For more detailed performance analysis of our approach, Monte Carlo simulations were performed. Fig. 7.5(a) shows the virtual setup for the Monte Carlo simulation. For simplicity, the light path is assumed to belong to xz-plane at $y = 0$ and the beam is assumed to be crossing point $(x, z) = (0, L)$ at $t = 0$, regardless of propagation angle θ. The horizontal sensor resolution is defined as N_H, whereas the focal length $-z_p$ is set at 25 mm to evaluate the reconstruction performance for a wide range of θ. The sensor size in horizontal direction is fixed at 9.6256 mm, equivalent to the actual dimension of the developed SPAD sensor, and L is fixed at 1 m, so as to reproduce a comparable time scale to the experiment. Fig. 7.5(b) shows examples of simulated spatio-temporal data set for $N_H = 32$, and $\theta = n\pi/8$ (n=1,2,3,...,7). The observation time t' is calculated for each pixel position x_p based on a simple geometrical calculation, where the temporal variation of t' is added by generating random variables following a normal distribution with the standard deviation $\sigma = 30$ ps (rms), thereby emulating the error of measurement. Fig. 7.5(c) shows the simulated spatio-temporal data set for $N_H = 1024$, corresponding to the megapixel resolution.

A cycle of Monte Carlo-based data generation and numerical fitting is repeated for 500 times to estimate the accuracy and precision of the fitting for various combinations of parameters.
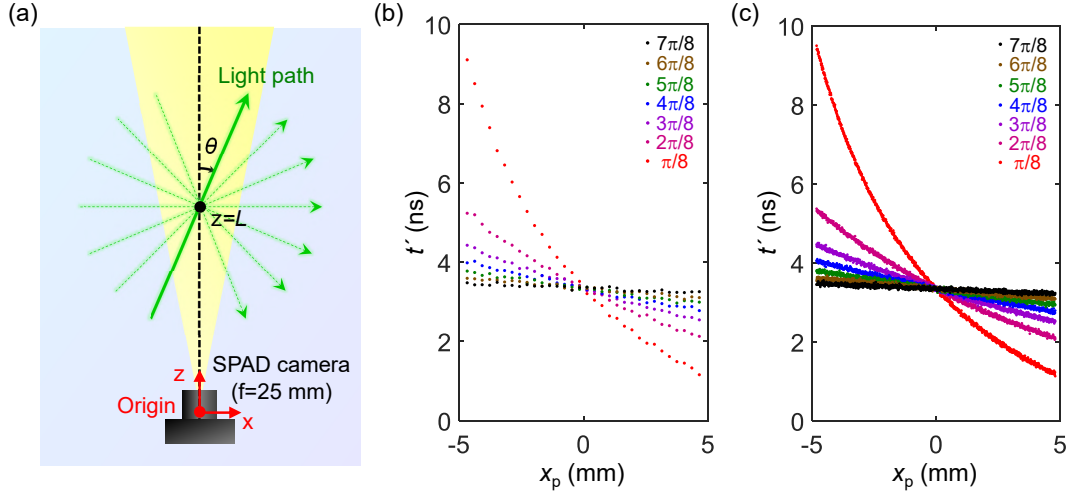
Figure 7.5 – Monte Carlo simulation setting. (a) Schematic illustration of virtual setup for Monte Carlo simulation of 4-dimensional light-in-flight imaging. The laser beam is assumed to pass a fixed point $(0, L)$ at $t = 0$. (b) Simulated x_p dependence of t' for $N_H = 32$, and $\theta = n\pi/8$ (n=1,2,3,...,7). Temporal variation is assumed to be $\sigma = 30$ ps (rms) for all the simulated points. (c) Simulated x_p dependence of t' for $N_H = 1024$, and $\theta = n\pi/8$ (n=1,2,3,...,7).
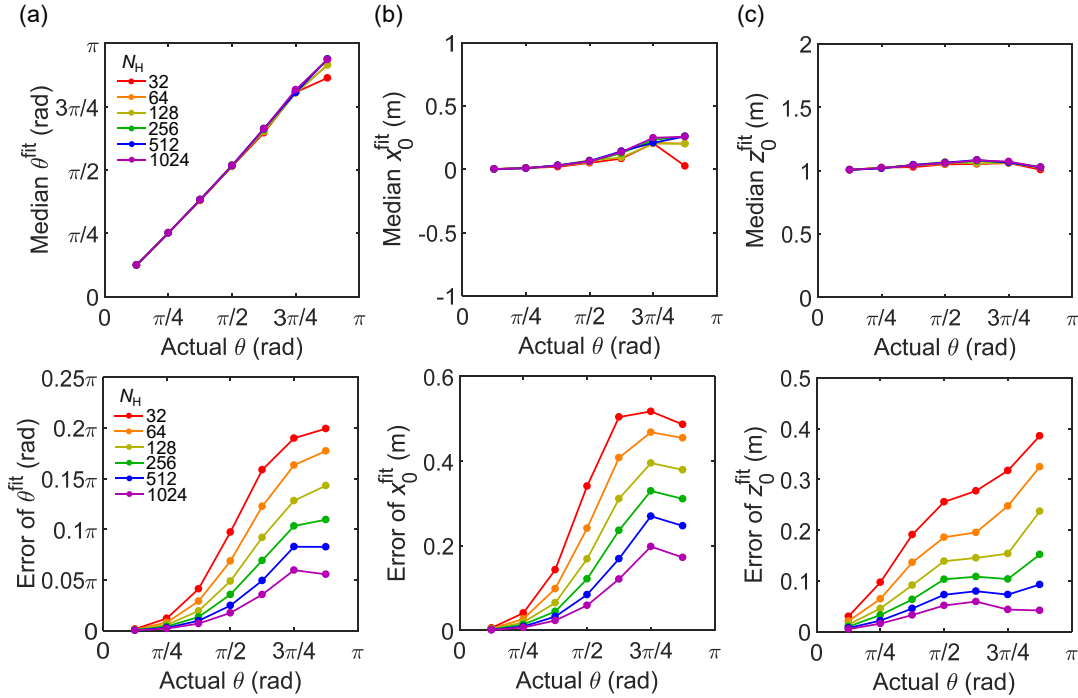
Figure 7.6 – Monte Carlo simulation results. (a) Propagation angle dependence of the median of fitted angle θ (top) and its estimation error (bottom) for $N_H = 32$, 64, 128, 256, 512, 1024, based on 500 cycles of data generation and fitting. (b) Propagation angle dependence of the median of fitted x_0 (top) and its estimation error (bottom). (c) Propagation angle dependence of the median of fitted z_0 (top) and the corresponding estimation error (bottom).

Fig. 7.6(a) shows the propagation angle dependence of the median of fitted angle θ (top) and its estimation error (bottom) for N_H = 32, 64, 128, 256, 512, 1024. The fitting is performed based on the equations shown in the main text, which are down-converted to 2-dimensional expressions in xz-plane. Three fitting parameters, θ, x_0, and z_0, are estimated based on the least-square regression. To avoid instability of the fitting, the fitting parameter θ is constrained between 0.06π and 0.94π. The median of the fitted angle shows a good agreement with the actual angle except for slight deviation at $\theta = 7\pi/8$. This is due to the potential instability of the algorithm for the angle close to $\theta = \pi$, giving a singular point where the apparent velocity becomes infinite. For all N_H in the bottom figure, the angle estimation error shows positive correlation to the actual angle θ. This stems from the fact that the light coming towards the camera leads to higher apparent velocity, and hence the effectively smaller change of the observation time makes the fitting process more noise-sensitive. For smaller N_H, the estimation error is saturated towards the maximum possible value of 0.25π corresponding to the standard deviation of the uniform distribution, where the fitting gives a totally random estimation of the parameter. For N_H = 1024, the estimation error is improved by factor of 6 with respect to N_H = 32 for unsaturated regions ($\theta \leq 4\pi/8$). For N_H = 1024, the estimation error of the angle below $\pm 0.06\pi$, or $\pm 11°$ is guaranteed for a whole range of θ ($\pi/8 \leq \theta \leq 7\pi/8$), whereas for N_H = 32, the equivalent level of error is only guaranteed for a range $\pi/8 \leq \theta \leq 3\pi/8$. This result suggests that the previously reported SPAD cameras with 32×32 resolution [10, 16] could suffer from the large angle estimation error, while justifying the necessity of the megapixel resolution for high precision 4-dimensional light-in-flight imaging. At larger values of θ, a smaller camera will be inherently inaccurate, thus requiring the use of multiple exposure and complex scanning systems, whereas a larger format camera is more reliable and can ensure a larger field of view in a reduced number of required shots.

Fig. 7.6(b) shows the propagation angle dependence of the median of fitted x_0 (top) and its estimation error (bottom) for N_H = 32, 64, 128, 256, 512, 1024. Similar to the fitted angle in Fig. 7.6(a), the median of fitted x_0 shows a slight deviation from the actual value $x_0 = 0$ (m) at larger angle θ. The corresponding estimation error of x_0 shows consistent trend with the estimation error of θ, the estimation error for N_H = 1024 is improved by factor of 6 with respect to N_H = 32 for unsaturated region ($\theta \leq 4\pi/8$). Fig. 7.6(c) shows the propagation angle dependence of the median of fitted z_0 (top) and the corresponding estimation error (bottom). In the top figure, the fitted z_0 shows a good agreement with the actual value $z_0 = 1$ (m). The estimation error for N_H = 1024 shows improvement by factor of 9 at $\theta = 7\pi/8$ with respect to the error for N_H = 32.

The above simulation focuses on the light path confined in xz-plane, while the discussion can be naturally extended to the full three-dimensional space. Note that the actual imaging system for 4-dimensional light-in-flight can be affected by various factors which are not taken into consideration in the Monte Carlo simulation. The estimation error could potentially be influenced by number of impinging photons, finite width and length of observed laser beam, timing non-uniformity over the SPAD array, dark counts, ambient light, hot pixels, and aberration of objective lens. The development of a more comprehensive simulation model could help understanding the performance limitation of the proposed approach, which

is essential for further improvement of the SPAD camera system and the reconstruction algorithm.

The above results obtained from Monte Carlo simulations suggest that a further increase of N_H could have a non-negligible impact on the stability and accuracy of light-in-flight measurements, along with the potential of further expansion of the measurement in distance and field-of-view.

7.4 Conclusion

Our approach of reconstructing extra-dimensional light-in-flight information with high spatio-temporal resolution single-photon camera can be applied to a wide range of high speed imaging techniques. One of the promising applications is the NLOS imaging [23, 24, 25, 26, 27]; by extending our approach for concentrated linear beam towards a diffused light, hidden objects can potentially be imaged with more simplified setup. The technique could also be used for estimating additional information of the hidden objects based on the currently used setup. Another potential application is the combination with optical tomography technique [28, 29, 30]; generalization of our theory can be useful for non-invasive 3-dimensional monitoring inside a target object structure as well as non-destructive measurement of 3-dimensional distribution of physical parameters such as refractive index and transmittance.

References

[1] N. Abramson, "Light-in-flight recording by holography," Opt. Lett. 3, 121–123 (1978).

[2] N. Abramson, "Light-in-flight recording: high-speed holographic motion pictures of ultrafast phenomena," Appl. Opt. 22, 215–232 (1983).

[3] N. H. Abramson, and K. G. Spears, "Single pulse light-in-flight recording by holography," Appl. Opt. 28, 1834–1841 (1989).

[4] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich, "Low-budget transient imaging using photonic mixer devices," ACM Trans. Graph 32, 45:1–45:10 (2013).

[5] K. Goda, K. Tsia, and R. Jalali, "Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena," Nature 458, 1145–1149 (2009).

[6] A. Velten, E. Lawson, A. Bardagjy, M. Bawendi, and R. Raskar, "Slow art with a trillion frames per second camera," Proc. SIGGRAPH 44 (2011).

[7] A. Velten, D. Wu, A. Jarabo, B. Masia, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar, "Femto-photography: capturing and visualizing the propagation of light. ACM Trans," Graph 32, 44:1–44:8 (2013).

[8] D. Faccio, and A. Velten, "A trillion frames per second: the techniques and applications of light-in-flight photography," Rep. Prog. Phys. 81, 105901 (2018).

[9] G. Etoh, T. Okinaka, Y. Takano, K. Takehara, H. Nakano, K. Shimonoura, T. Ando, N. Ngo, Y. Kamakura, V. T. S. Dao, A. Q. Nguyen, E. Charbon, C. Zhang, P. D. Moor, P. Goetschalckx, and L. Haspeslagh, "Light-In-Flight Imaging by a Silicon Image Sensor: Toward the Theoretical Highest Frame Rate," Sensors 19, 2247 (2019).

[10] G. Gariepy, N. Krstajic, R. Henderson, C. Li, R. R. Thomson, G. S. Buller, B. Heshmat, R. Raskar, J. Leach, and D. Faccio, "Single-photon sensitive light-in-flight imaging," Nat. Commun. 6(6021) (2015).

[11] M. Laurenzis, J. Klein, E. Bacher, and N. Metzger, "Multiple-return single-photon counting of light in flight and sensing of non-line-of-sight objects at shortwave infrared wavelengths," Opt. Lett. 40, 20 (2015).

References

[12] R. Warburton, C. Aniculaesei, M. Clerici, Y. Altmann, G. Gariepy, R. McCracken, D. Reid, S. McLaughlin, M. Petrovich, J. Hayes, R. Henderson, D. Faccio, and J. Leach, "Observation of laser pulse propagation in optical fibers with a SPAD camera," Sci. Rep. 7, 43302 (2017).

[13] M. O'Toole, F. Heide, D. B. Lindell, K. Zang, S. Diamond, and G. Wetzstein, "Reconstructing Transient Images from Single-Photon Sensors," IEEE Proc. CVPR (2017).

[14] K. Wilson, B. Little, G. Gariepy, R. Henderson, J. Howell, and D. Faccio, "Slow light in flight imaging," Phys. Rev. A 95, 023830 (2017).

[15] Q. Sun, X. Dun, Y. Peng, and W. Heidrich, "Depth and Transient Imaging with Compressive SPAD Array Cameras," IEEE Proc. CVPR (2018)

[16] M. Laurenzis, J. Klein, and E. Bacher, "Relativistic effects in imaging of light in flight with arbitrary paths," Opt. Lett. 41, 9 (2016).

[17] K. Morimoto, A. Ardelean, M.-L. Wu, A. C. Ulku, I. M. Antolovic, C. Bruschini, and E. Charbon, "Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," Optica 7, 346–354 (2020).

[18] C. Veerappan, J. Richardson, R. Walker, D.-U. Li, M. W. Fishburn, Y, Maruyama, D. Stoppa, F. Borghetti, M. Gersbach, R. K. Henderson, and E. Charbon, "A 160×128 single-photon image sensor with on-pixel 55ps 10b time-to-digital converter," IEEE Int. Solid-State Circuits Conference (2011).

[19] A. Carimatto, S. Mandai, E. Venialgo, T. Gong, G. Borghi, D. R. Schaart, and E. Charbon, "A 67,392-SPAD PVTB-compensated multi-channel digital SiPM with 432 column-parallel 48ps 17b TDCs for endoscopic time-of-flight PET," IEEE Int. Solid-State Circuits Conference (2015).

[20] R. K. Henderson, N. Johnston, F. M. D. Rocca, H. Chen, D. D.-U. Li, G. Hungerford, R. Hirsch, D. Mcloskey, P. Yip, and D. J. S. Birch, "A 192×128 Time Correlated SPAD Image Sensor in 40-nm CMOS Technology," IEEE J. Solid-State Circuits 54, 1907-1916 (2019).

[21] H. E. Bond, A. Henden, Z. G. Levay, N. Panagia, W. B. Sparks, S. Starrfield, R. M. Wagner, R. L. M. Corradi, and U. Munari, "An energetic stellar outburst accompanied by circumstellar light echoes," Nature 422, 405-408 (2003).

[22] K. P. Mooley, A. T. Deller, O. Gottlieb, E. Nakar, G. Hallinan, S. Bourke, D. A. Frail, A. Horesh, A. Corsi, and K. Hotokezaka, "Superluminal motion of a relativistic jet in the neutron-star merger GW170817," Nature 561, 355–359 (2018).

[23] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar, "Looking around the corner using ultrafast transient imaging," Int. J. Comput. Vis. 95, 13–28 (2011).

[24] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," Nat. Commun. 3, 745 (2012).

[25] O. Katz, E. Small, and Y. Silberberg, "Looking around corners and through thin turbid layers in real time with scattered incoherent light," Nat. Photonics 6, 549–553 (2012).

[26] M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," Nature 555, 338–341 (2018).

[27] X. Liu, I. Guillén, M. L. Manna, J. H. Nam, S. A. Reza, T. H. Le, A. Jarabo, D. Gutierrez, and A. Velten, "Non-line-of-sight imaging using phasor-field virtual wave optics," Nature 572, 620-623 (2019).

[28] D. Huang, "Optical coherence tomography," Science 254, 1178-1181 (1991).

[29] T. Durduran, R. Choe, W. B. Baker, and A. G. Yodh, "Diffuse optics for tissue monitoring and tomography," Rep. Prog. Phys. 73, 076701 (2010).

[30] E. B. de Haller, "Time-resolved transillumination and optical tomography," J. Biomed. Opt. 1, 1-17 (1996).

8 Conclusion

Since 2003, researchers have pioneered new application possibilities of SPAD detectors along with successful scaling of the SPAD arrays. To reach the important milestone of 1 megapixel SPAD, we have investigated various possibilities of further miniaturization and scaling of SPAD image sensors towards time-resolved applications. In this concluding chapter, the main findings and scientific contributions presented in this thesis are summarized. In addition, a technical perspective of the research and development for the future-generation SPAD image sensors is provided.

8.1 Summary

Three objectives of this thesis were defined in Chapter 1:

1. formulate the scaling laws of SPAD pixels and exemplify fundamental tradeoffs in the miniaturized SPAD arrays below 10 μm pitch;

2. propose and implement novel concepts to overcome the tradeoffs in the miniaturized SPADs based on insights from semiconductor device physics, fabrication process technology, and IC design methodology;

3. employ the large-scale SPAD arrays towards various time-resolved applications.

The first objective is addressed in Chapter 2, where the scaling laws of miniaturized SPAD arrays for fill factor, PDP, PDE, DCR, afterpulsing, crosstalk, power consumption, and timing jitter are investigated based on the theoretical analysis. The second objective is fulfilled in Chapter 3, 4, 5 and 6; we proposed and implemented device-level concepts including guard-ring sharing technique and charge focusing SPAD, circuit-level concepts including readout circuit sharing, system-level concepts including SAR-ToF and coded time gating, and process-level concepts including 3D-stacked configurations, all of which are of great importance to overcome the fundamental tradeoffs in the miniaturization. The third objective is achieved

in Chapter 6 and 7, where the megapixel time-gated SPAD cameras demonstrated versatility towards 2D, 3D and even 4D imaging applications.

8.2   Scientific contributions

Towards the megapixel SPAD implementation for time-resolved applications, we have investigated the novel technologies from diversified standpoints: theoretical analysis, conceptual proposals, implementation to ICs, performance characterization, and demonstration towards applications. Major scientific contributions of this thesis are listed in the following:

- Fundamental tradeoff relations in the SPAD pixel scaling are clarified based on theoretical analysis.

- A novel guard-ring-sharing technique is proposed to push the limit of SPAD pixel scaling, and SPAD operation with 2.2 μm-pitch, the world's smallest SPAD pitch is experimentally demonstrated in 4×4 test SPAD arrays.

- A novel charge focusing SPAD is proposed to overcome a fundamental tradeoff between PDE and DCR, and the proof-of-concept image sensor achieved the world's lowest DCR density, while on-chip color filter is implemented in the SPAD image sensor for the first time.

- Novel pixel-wise/region-wise time-gating approaches are proposed, providing high speed and scalable solutions for ToF ranging and FLIM applications.

- A 1 megapixel SPAD image sensor is demonstrated for the first time.

- A novel algorithm for 4-dimensional light-in-flight reconstruction is developed, and the 4-dimensional light-in-flight imaging is experimentally demonstrated for the first time, based on the 1 megapixel SPAD image sensor.

In Chapter 2, theoretical analysis of scaling laws in SPAD pixel was conducted by Kazuhiro Morimoto (KM), and directed by Prof. Edoardo Charbon (EC). In Chapter 3, the concept of guard-ring sharing was conceived by KM. The chip design and experiment were conducted by KM, and directed by EC. In Chapter 4, the concept of charge focusing SPAD was conceived by KM and Mahito Shinohara (MS). The chip design and experiment were conducted by KM and the colleagues in Canon Inc. The device and process optimization was conducted by Junji Iwata (JI), MS, KM, and the colleagues in Canon Inc. The project was directed by Shunsuke Inoue (SI) and Takeshi Ichikawa (TI). In Chapter 5, the concepts of SAR-ToF and coded time gating were conceived by KM. The design of the 3D-stacked chips was conducted by KM, Myung-Jae Lee (MJ), and Andrei Ardelean (AA), and directed by EC. In Chapter 6, the pixel architectures were conceived by KM and EC. The chip design was conducted by KM. The design of PCB and FPGA firmware was conducted by KM, AA, and Arin Can Ulku (AU). The

experiment and data analysis were conducted by KM, Ming-Lo Wu (MW), and AA, and directed by EC. In Chapter 7, the concept of 4-dimensional light-in-flight reconstruction was conceived by KM and MW. The theoretical formulation was conducted by KM and AA. The experiment was conducted by KM, MW, and AA, and directed by EC.

8.3 Future perspective

The aforementioned contributions are made in individual sub-projects, but the underlying key technical concepts are complementary to each other. Further miniaturization and scaling of the SPAD image sensors can be realized by combining these concepts into a single chip together with established techniques, potentially achieving multi-megapixel time-resolved SPAD cameras to open up a new dimension of applications. Finally, we give some perspectives in technology directions of SPAD imagers, hopefully to provide inspiration for research and development of future-generation SPAD cameras.

8.3.1 10 megapixel SPAD arrays

We demonstrated a megapixel SPAD, which has been considered one of the most important milestones for over 10 years. Exponential scaling trends of SPAD array size in recent years gives a prediction towards 10 megapixel SPAD arrays achieved within the next 10 years. Further scaling will confront several technical challenges. First, SPAD-originated power consumption needs to be managed properly. Assuming passive quenching architecture with the quenching resistance $R_q = 200$ kOhm, the excess bias $V_{ex} = 3.3$ V, and the breakdown voltage $V_B = 22$ V, the worst-case power consumption per pixel is calculated as:

$$P = V \cdot I = (V_{ex} + V_B) \cdot \frac{V_{ex}}{R_q} = 0.42 \text{ (mW)}. \tag{8.1}$$

This indicates that a 10 megapixel SPAD sensor will theoretically consume more than 4 kW under strong illumination, which is not practical due to heat dissipation, voltage drop at power lines, etc. A possible countermeasure is to introduce active quenching and/or recharging. For example, the megapixel SPAD array in Chapter 6 operates in active recharging mode. From the measurement result, the worst-case SPAD-originated power consumption with a megapixel array (pixel B) is 41 mW, whereas passive quenching operation with the same array would consume 420 W. In this specific case, the active recharging can suppress the power consumption by a factor of 10^4. Note that the worst-case power consumption in the active recharging operation is proportional to a maximum allowable photon count rate, indicating a tradeoff with dynamic range.

Further reduction of the power consumption can be realized by reducing V_{ex}, V_B, and parasitic capacitance at the SPAD output node. For high level integration towards potential applications, care must be taken to the underlying tradeoffs between power consumption and PDE, DCR,

dynamic range, pixel functionality, readout speed, chip size, cost, robust manufacturability, etc. In addition, power consumption from pixel circuits such as quenching and recharging, digital counters, signal processing logic circuits, and TDCs should also be taken into consideration for scaling of the SPAD array size.

Second, non-uniformity is a critical issue towards 10 megapixel SPADs. In addition to the non-uniformity of basic characteristics such as breakdown voltage, PDE and DCR, the non-uniformity of timing performance must be strictly controlled for time-resolved applications. As the array size increases, longer signal lines and larger parasitics hinder uniform distribution of clock signals over the array. For the time-gating approach, balanced binary trees are proven to be feasible for clock distribution in 1 megapixel SPADs, whereas more advanced approaches need to be investigated for 10 megapixels to ensure the comparable timing skew to the megapixel SPAD sensors.

Third, innovation is required for boosting readout speed in 10 megapixel SPAD image sensors. To achieve a comparable frame rate to the current megapixel SPAD imagers, designers have to deal with an amount of output data 10 times larger. Developing new algorithms for on-chip data compression would help efficiently reducing the bit rate. 3D-stacked approaches with advanced CMOS technology node for circuit wafer are a promising direction to implement compact and highly functional processing circuits.

8.3.2   1 μm-pitch SPAD pixels

Despite the demonstration of the world's smallest 2.2 μm-pitch SPAD arrays in this thesis, the pixel size of SPAD imagers still falls behind that of CISs, where 1 μm or smaller pixels are already employed for commercial products. Further innovations and continuous efforts on the miniaturization could potentially enable 1 μm-pitch SPAD arrays in the future. To target this, the 3D-stacked approach combined with advanced technology node is indispensable. In addition, the charge focusing approach and guard-ring sharing technique could play key roles to ensure the high PDE and low DCR. In such a super-miniaturized SPAD device, electric field distribution must be carefully designed to avoid premature edge breakdown. In addition, optical and electrical crosstalk components could become major limiting factors for the aggressive miniaturization. For both aspects, further breakthrough in the device, process, and optics technology is a key requirement to push the physical limit of current SPAD devices.
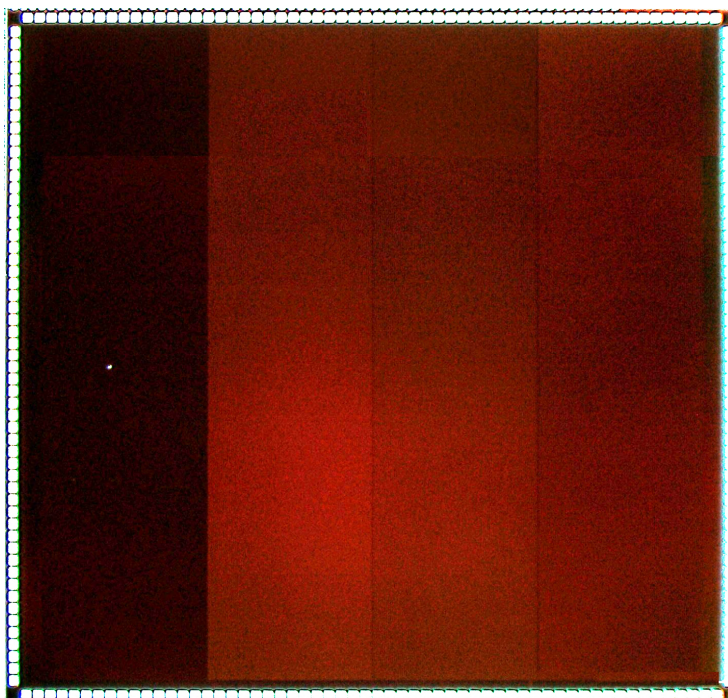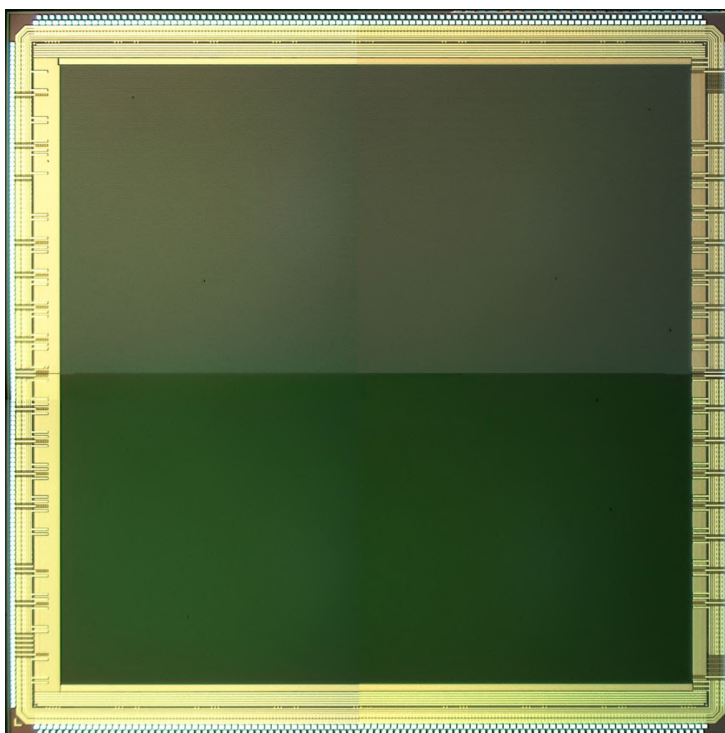
Chip Gallery
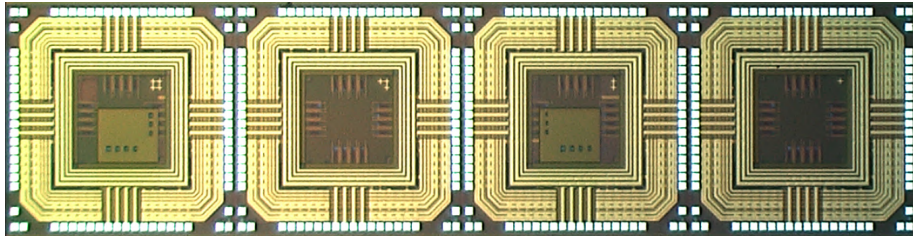
Guard-ring-shared 4×4 SPAD arrays, Chapter 3
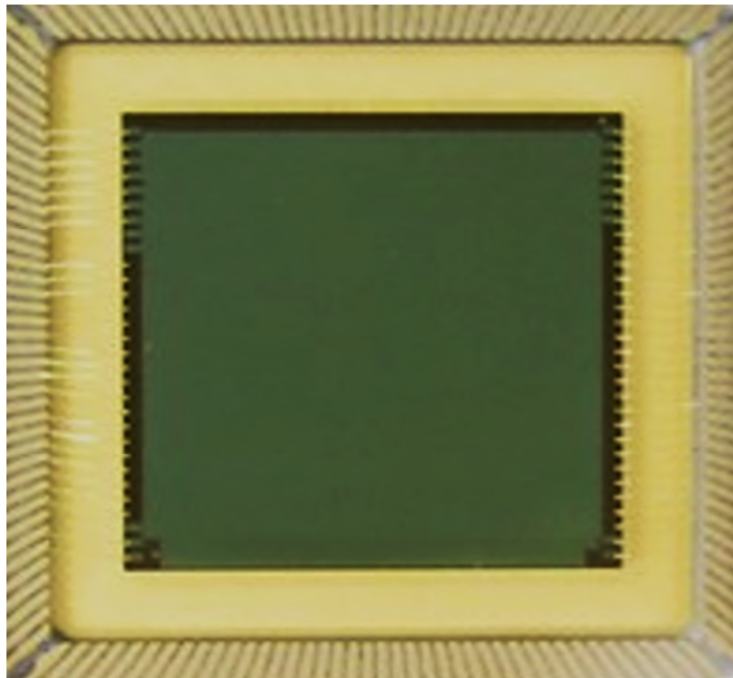
128×128 charge focusing SPAD imager, Chapter 4

BSI 3D-stacked megapixel SPAD imager, Chapter 5

FSI megapixel time-gated SPAD imager, Chapter 6

List of Publications

International peer-reviewed journal articles

K. Morimoto, A. Ardelean, M.-L. Wu, A. C. Ulku, I. M. Antolovic, C. Bruschini, and E. Charbon, "Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," Optica 7(4), 346–354 (2020).

K. Morimoto, and E. Charbon, "High fill-factor miniaturized SPAD arrays with a guard-ring-sharing technique," Opt. Express 28(9), 13068-13080 (2020).

K. Morimoto, M.-L. Wu, A. Ardelean, and E. Charbon, "Superluminal motion-assisted 4-dimensional light-in-flight imaging," Submitted to Light Sci. Appl. (2020).

V. Zickus, M.-L. Wu, **K. Morimoto**, V. Kapitany, A. Fatima, A. Turpin, R. Insall, J. Whitelaw, L. Machesky, C. Bruschini, D. Faccio, and E. Charbon, "Wide-field fluorescence lifetime imaging microscopy with a high-speed mega-pixel SPAD Camera," Submitted to Light Sci. Appl. (2020).

Preprints

K. Morimoto, A. Ardelean, M.-L. Wu, A. C. Ulku, I. M. Antolovic, C. Bruschini, and E. Charbon, "A megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," arXiv:1912.12910 (2019).

Conference presentations and demonstrations

K. Morimoto, A. C. Ulku, I. M. Antolovic, C. Bruschini, and E. Charbon, "A single-photon camera with 97 kfps time-gated 24 Gphotons/s 512 × 512 SPAD pixels for computational imaging and time-of-flight vision," Technical demonstration in ICCP2019 (2019).

E. Charbon, **K. Morimoto**, and C. Bruschini, "Large-format SPAD cameras," Image Sensor

Europe (2020).

K. Morimoto, A. Ardelean, M.-L. Wu, A. C. Ulku, I. M. Antolovic, C. Bruschini, and E. Charbon, "A megapixel time-gated SPAD image sensor for 2D and 3D imaging applications," Workshop on Information Sensing Technologies (2020).

J. Iwata, **K. Morimoto**, Y. Kuroda, T. Goden, T. Sasago, Y. Ota, K. Ota, M. Shinohara, Y. Fukuchi, A. Chiba, M. Furuta, H. Takahashi, and T. Ichikawa, "A 0.16e-rms Temporal Noise 0.015Hz/μm^2 Dark Count Rate Charge Focusing SPAD Image Sensor for Low Light Imaging," Workshop on Information Sensing Technologies (2020).

K. Morimoto, "Charge-Focusing SPAD Image Sensors for Low Light Imaging Applications," Int. SPAD Sensor workshop (2020).

Patents

K. Morimoto, H. Ikeda, J. Iwata, inventors; Canon Inc. PHOTODETECTION DEVICE AND PHOTODETECTION SYSTEM. U. S. Patent, Application. 20180033896 (2017).

K. Morimoto, M. Kobayashi, K. Kawabata, T. Kato, H. Sekine, inventors; Canon Inc. IMAGING DEVICE, IMAGING SYSTEM, AND MOVING OBJECT. U. S. Patent, Application. 20180035068 (2017).

K. Morimoto, M. Shinohara, inventors; Canon Inc. PHOTO-DETECTION APPARATUS AND PHOTO-DETECTION SYSTEM. U. S. Patent, Application. 20180108800 (2017).

K. Morimoto, H. Ikeda, J. Iwata, inventors; Canon Inc. PHOTODETECTION DEVICE AND SYSTEM HAVING AVALANCHE AMPLIFICATION. U. S. Patent, Application. 20180151758 (2017).

K. Morimoto, M. Kobayashi, K. Kawabata, T. Kato, H. Sekine, inventors; Canon Inc. IMAGING DEVICE, IMAGING SYSTEM, AND MOVING OBJECT. U. S. Patent, Grant. 10142574 (2017).

K. Morimoto, H. Ikeda, J. Iwata, inventors; Canon Inc. PHOTODETECTION DEVICE AND SYSTEM HAVING AVALANCHE AMPLIFICATION. U. S. Patent, Grant. 10283651 (2019).

K. Morimoto, H. Ikeda, J. Iwata, inventors; Canon Inc. PHOTODETECTION DEVICE AND PHOTODETECTION SYSTEM. U. S. Patent, Grant. 10497818 (2019).

K. Morimoto, M. Shinohara, inventors; Canon Inc. PHOTO-DETECTION APPARATUS AND PHOTO-DETECTION SYSTEM. U. S. Patent, Grant. 10497822 (2019).

K. Morimoto, M. Shinohara, inventors; Canon Inc. PHOTO-DETECTION APPARATUS AND PHOTO-DETECTION SYSTEM. U. S. Patent, Application. 20200052143 (2019).

K. Morimoto, H. Ikeda, J. Iwata, inventors; Canon Inc. PHOTODETECTION DEVICE AND PHOTODETECTION SYSTEM. U. S. Patent, Application. 20200058808 (2019).

K. Morimoto, inventors; Canon Inc. To be published (2019).

K. Morimoto, inventors; Canon Inc. To be published (2020).

K. Morimoto, inventors; Canon Inc. To be published (2020).

D. Shirahige, **K. Morimoto**, and H. Sekine, inventors; Canon Inc. To be published (2020).

H. Sekine, and **K. Morimoto**, inventors; Canon Inc. To be published (2020).

K. Morimoto, inventors; Canon Inc. To be published (2020).

K. Morimoto, inventors; Canon Inc. To be published (2020).

*NOTE: Some of the patents may be related to the works presented in this thesis, but the detailed relations are not explicitly stated for the purpose of IP protection.

Press releases

The Optical Society: https://www.osa.org/en-us/about_osa/newsroom/news_releases/2020/new_photon-counting_camera_captures_3d_images_with/

The Optical Society: https://www.osa-opn.org/home/newsroom/2020/april/single-photon_detectors_build_a_super-fast_camera/

EPFL: https://actu.epfl.ch/news/megax-the-first-camera-to-capture-the-smallest-par/

Phys.org: https://phys.org/news/2020-04-photon-counting-camera-captures-d-images.html

ScienceDaily: https://www.sciencedaily.com/releases/2020/04/200416114537.htm

Canon: https://global.canon/en/news/2020/20200624.html

Nikkei: https://www.nikkei.com/article/DGXLRSP536439_U0A620C2000000/

Curriculum Vitae

Kazuhiro Morimoto

1988   Born in Tokyo, Japan

Professional Experience

2013-   **Canon Inc.**
Image Sensor Device, Optics & Circuit Design Engineer

Education

2017-2020   **École polytechnique fédérale de Lausanne (EPFL)**
Advanced Quantum Architecture Laboratory (AQUA)
The Doctoral Program in Microsystems and Microelectronics (EDMI)
Thesis: "Megapixel SPAD Cameras for Time-Resolved Applications"

2011-2013   **The University of Tokyo**
Master of Engineering, Department of Applied Physics,
Graduate School of Engineering
Thesis: "Detection of Single Photon-Single Electron Spin Conversion
in g-factor Engineered Quantum Dots"

2007-2011   **The University of Tokyo**
Bachelor of Engineering, Department of Applied Physics,
Faculty of Engineering
Thesis: "Electric Field Modulated Rashba Effect in Liquid Gated Transistors"

2001-2007   **Azabu Junior and Senior High school**

Awards

2020	**Nominated for EPFL Doctorate Awards**	
	École polytechnique fédérale de Lausanne	
2020	**Swiss Nanotechnology PhD Award**	
	swiss mnt network	
2020	**Patent Award C**	
	Device Tech. Development HQ., Canon Inc.	
2017	**Best Patent Award**	
	Device Tech. Development HQ., Canon Inc.	
2013	**Excellent Master's Thesis Award**	
	Department of Applied Physics, The University of Tokyo	
2011	**Excellent Bachelor's Thesis Award**	
	Department of Applied Physics, The University of Tokyo	