CMOS Sensors for Time-Resolved Active Imaging

by

Jihyun Cho

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)
in the University of Michigan
2017

Doctoral Committee:

Professor Euisik Yoon, Chair
Professor Michael P. Flynn
Professor Mary-Ann Mycek
Emeritus Professor Kensall D. Wise

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

CMOS Sensors for Time-Resolved Active Imaging

by

Jihyun Cho

In the past decades, time-resolved imaging such as fluorescence lifetime or time-of-flight depth imaging has been extensively explored in biomedical and industrial fields because of its non-invasive characterization of material properties and remote sensing capability. Many studies have shown its potential and effectiveness in applications such as cancer detection and tissue diagnoses from fluorescence lifetime imaging, and gesture/motion sensing and geometry sensing from time-of-flight imaging. Nonetheless, time-resolved imaging has not been widely adopted due to the high cost of the system and performance limits.

The research presented in this thesis focuses on the implementation of low-cost real-time time-resolved imaging systems. Two image sensing schemes are proposed and implemented to address the major limitations.

First, we propose a single-shot fluorescence lifetime image sensors for high speed and high accuracy imaging. To achieve high accuracy, previous approaches repeat the measurement for multiple sampling, resulting in long measurement time. On the other hand, the proposed method achieves both high speed and accuracy at the same time by employing a pixel-level processor that takes and compresses the multiple samples within a single measurement time. The pixels in the sensor take multiple samples from the fluorescent optical signal in sub-nanosecond resolution and compute the average photon arrival time of the optical signal. Thanks to the multiple sampling of the signal, the measurement is insensitive to the shape or the pulse-width of excitation, providing better accuracy and pixel uniformity than conventional rapid lifetime determination (RLD) methods. The proposed single-shot image sensor also improves the imaging speed by orders of magnitude compared to other conventional center-of-mass methods (CMM).

Second, we propose a 3-D camera with a background light suppression scheme which is adaptable to various lighting conditions. Previous 3-D cameras are not operable in outdoor conditions because they suffer from measurement errors and saturation problems under high background light illumination. We propose a reconfigurable architecture with column-parallel discrete-time background light cancellation circuit. Implementing the processor at the column level allows an order of magnitude reduction in pixel size as compared to existing pixel-level processors. The column-level approach also provides reconfigurable operation modes for optimal performance in all lighting conditions. For example, the sensor can operate at the best frame-rate and resolution without the presence of background light. If the background light saturates the sensor or increases the shot noise, the sensor can adjust the resolution and frame-rate by pixel binning and superresolution techniques. This effectively enhances the well capacity of the pixel to compensate for the increase shot noise, and speeds up the frame processing to handle the excessive background light. A fabricated prototype sensor can suppress the background light more than 100-klx while achieving a very small pixel size of 5.9μm.

Chapter 1

Introduction

Active optical sensing has been an important tool in clinical, scientific, and industrial applications. Such applications use light to measure information such as material properties or micro-environments in a non-invasive and remote way. Its extensive use in scientific fields is well summarized in books and review articles [1]–[4]. Particularly in the biomedical area, a number of studies rely on fluorescent light based active imaging in order to find cancerous cells [5]–[10], hazardous chemicals [3], and cellular environments such as pH, refractive index, and ion or oxygen concentrations [2]. In industry and consumer electronics, optical sensing has been applied for depth measurement based on the time-of-flight (TOF) method [11]. Measuring the depth or geometry can be used for collision avoidance or terrain scan in autonomous vehicles, or gesture recognition in human-computer interaction.

Despite their potentially high applicability, time-resolved images have not yet been widely adopted for various reasons. First of all, time-resolved imaging systems are highly complex and expensive. In order to measure optical phenomena in sub-nanosecond resolution, the system needs multiple devices with extreme timing precision such as a pico- or nanosecond laser source, a special camera, i.e. a time-resolved image sensor, and controlling electronics. Additionally, the devices often only operate with a high voltage source which is potentially hazardous for implantable devices in clinical applications.

Second, the frame-rate or the imaging speed of time-resolved sensor is very low because time-resolved sensors need to capture multiple frames of intensity images at different times to produce a single time image. Some may take less samples, i.e. only one or two frames, to calculate the time image; this, however results in low accuracy due to the under sampling of the optical signal. The trade-off between speed and accuracy needs to be overcome in emerging applications that require both high accuracy and high speed at the same time, such as real-time medical applications.

Lastly, time-resolved imaging does not work well outdoors due to its vulnerability to the ambient light or the background light. Because time-resolved imaging senses photons from its own light emitter, any photons originating from light sources other than the emitter are considered interference. Interference can add measurement error and increase the random shot noise. In the worst case, strong interference can blind the sensor by saturating the sensing node.

In the rest of this chapter, the aforementioned problems are elaborated with the background knowledge of time-resolved imaging. After describing the challenges, we then present our solution to the problems.

1.1 Active Optical Sensing and Time-Resolved Imaging

In conventional camera systems, the optical sensor or the imaging device measures the light from outside sources such as sunlight. This measurement is referred as a passive optical sensing because the measured photons are from outside of the system. In passive optical sensing, the sensors can measure only the brightness of the light or the number of photons for a fixed time period.

On the other hand, active optical sensing employs the measurement of a reflected or fluorescent light that originates from the emitter in the system. The emitter is synchronized with the sensor, giving a system the ability to measure time information from the light.

Timing information can contain valuable information about the sensed object. For example, the depth information, distance between the sensor and the target, can be translated into time delay of the optical signal. This is illustrated in the Fig. 1-1 where a short pulse of light is emitted and then reflected to the sensor. The reflected light is attenuated and delayed from its initial emission when it arrives at the sensor. The returned optical pulse now contains two information, the reflectivity of the target and the distance to the target. The attenuation contains both the reflectivity and the depth information, and the time delay has only one information, the depth. Because the conventional optical sensors measure only the brightness of the light, they cannot measure depth. Measuring the depth (and optionally the reflectivity) from the optical signal requires at least two independent measurements.

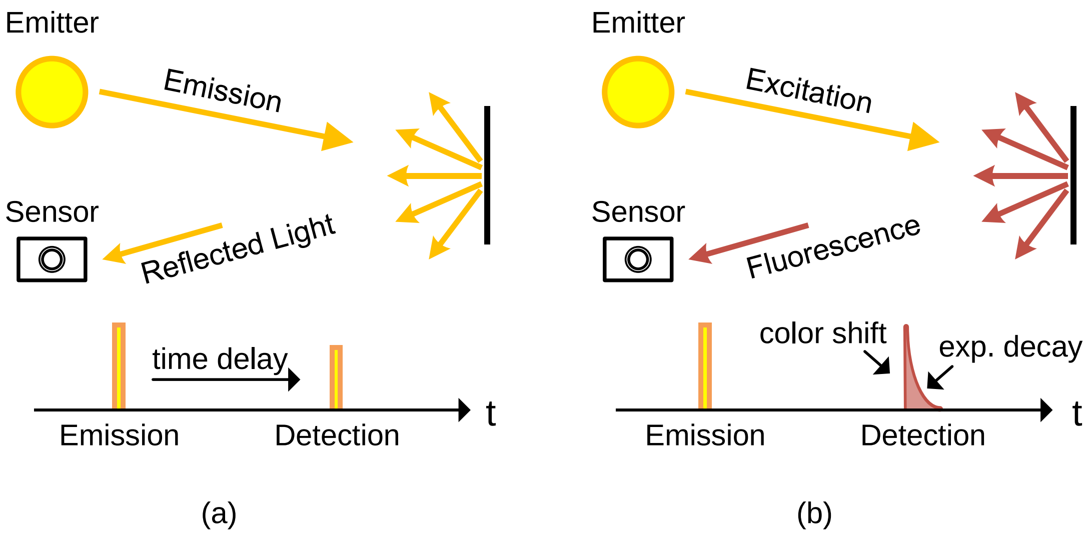


Fig. 1-1. Active optical sensing of (a) depth (b) fluorescence lifetime.

Fig. 1-1(b) shows another example of active sensing where fluorescent targets are measured. Similar to the previous example, the emitter first transmits an optical pulse to the fluorescent target. The emitted light is called an excitation light because the energy transfer from the photons excites the electrons in the target. As shown in the Fig. 1-2, the excitation causes the electron energy to jump from the ground state to an excited state. The excited electrons may lose energy through vibrational relaxation within a few pico-seconds to be in the intermediate states (bottom of S1). They stay there for several nano-seconds before they return to the ground state, releasing photons. Since the stay time is a random variable, the resulting fluorescence from an impulse excitation looks like an exponential decay. It is known that the time constant of the decay or the fluorescence lifetime is sensitive to many important biological features; therefore, the sensor can resolve the lifetime from the decaying optical signals for indirect measure of the biological markers.

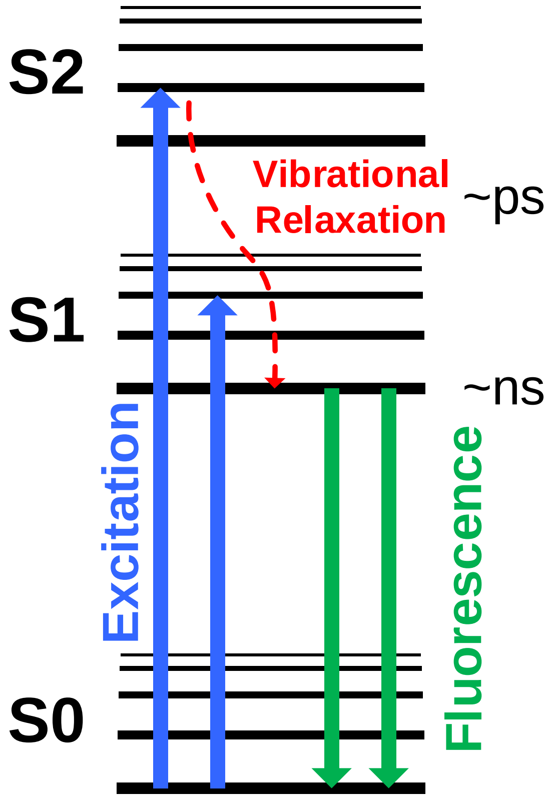


Fig. 1-2. Physics of fluorescence: Jablonski diagram.

As in the previous examples, active optical sensing requires a time measurement. This can be achieved by time-resolved imaging techniques. In general, there are two type of time-resolved imaging, time-correlated single photon counting (TCSPC) and gated-imaging, which are explained in detail in the subsequent sections.

1.1.1 Time-Correlated Single Photon Counting

The time-correlated single photon counting (TCSPC) measures the arrival time of each individual photon as in the example shown in Fig. 1-3(a). In this measurement, it is assumed that the sensor receives one photon at a time. Each time, the detector receives a photon or no photon from the weak fluorescence. When a photon is detected, it generates an electrical pulse and the pulse time is measured by a time-to-digital converter (TDC). After repeating many measurements, the system knows the arrival time of all photons. A histogram is built from the collected information to represent the original optical waveform as shown in the histogram plot of the Fig. 1-3(a). Theoretically, TCSPC measures the arrival time of every single photon without losing any photon, resulting in very high signal-to-noise ratio.

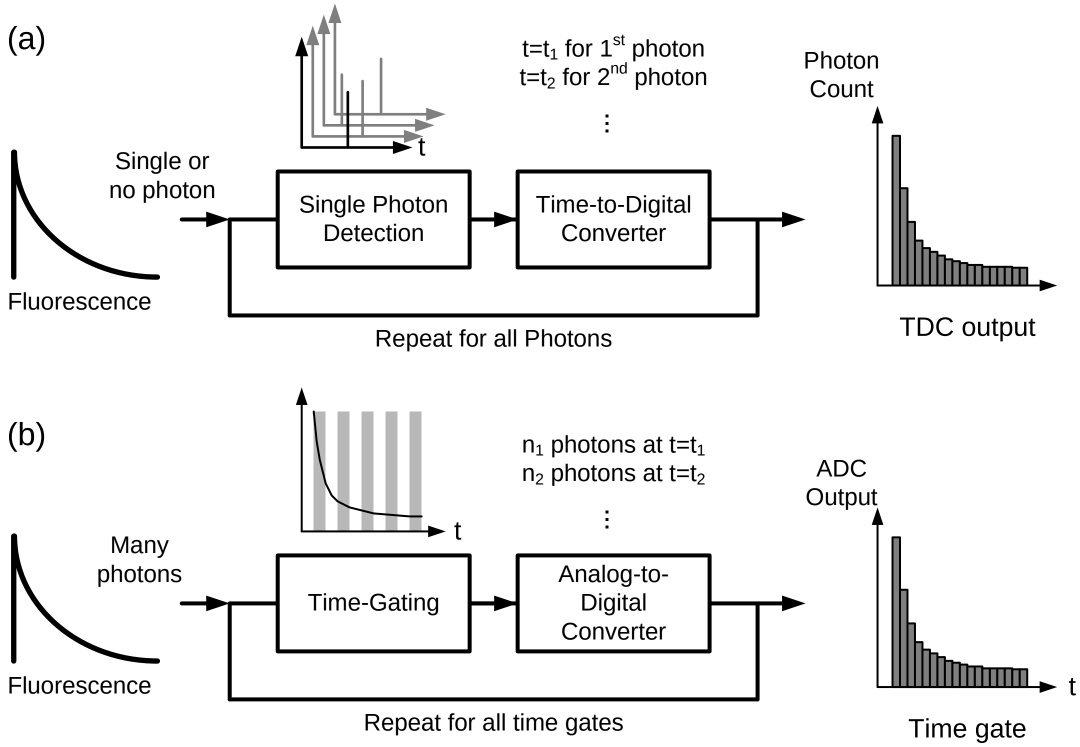


Fig. 1-3. Operation of time-resolve imaging techniques: (a) time-correlated single photon counting (TCSPC) and (b) gated-imaging.

However, TCSPC suffers from slow speed. This is due to the nature of single-photon detection. When a single-photon detector, such as photomultiplier tube (PMT) or single-photon avalanche diode (SPAD), detects a photon, it need to be reset or quenched for the next photon detection. Resetting the detector takes a finite time, and any photons arriving in this period are lost. This is called a photon pile-up and it limits the maximum photon detection rate up to 1% for practical use. Subsequently, TCSPC requires repeated measurement with at least ~100 times redundancy.

1.1.2 Time-Gated Imaging

Unlike TCSPC, gated imaging can receive multiple photons simultaneously. Hence, the measurement can be fundamentally quicker than TCSPC. As shown in Fig. 1-3(b), a time window or a shutter window is defined to selectively receive the photons in each measurement. The width and the number of windows may vary depending on the applications. As an example, the figure shows 5 measurement windows in gray boxes. Each time, only photons within one selected window are integrated while the rest are discarded. The integrated photons are then measured by an analog-to-digital converter (ADC). After repeating this measurement with different time gates, the optical waveform can be reconstructed as shown in the histogram of Fig. 1-3(b).

Due to its fast imaging speed, time-gated imaging is the most suitable method for time-resolved imaging in real-time applications. One drawback of this method is the low SNR especially under low light situation due to the loss of a significant number of photons outside of the time gate. A wider time gate can be defined to have more photons within the time-gate but then the time resolution will be less due to the more coarsely defined time gate.

1.2 Challenges

In many applications, time-resolved imagers are required to have high spatiotemporal resolution (frame-rate and spatial resolution), high accuracy, portability, and background light suppression. For example, many clinical applications require real-time imaging or high frame-rate for blur-free images because the sensors and the subjects are not likely to be fixed or stationary. The sensor should provide accurate results because it may affect the diagnoses. For point-of-care diagnosis or time-of-flight depth imaging, portability is also very important. The portability includes not only the size and weight of the system, but also the robustness of the system to various ambient lighting conditions.

In this section, we present three types of challenges that prevent time-resolved imagers from meeting the requirements.

1.2.1 Spatiotemporal Resolution

For real-time wide-field imaging, we suggest at least 256×256 pixels and 10 frames per second (fps). In some applications, spatial resolution can be easily increased at a reduced frame-rate when mechanical scanning or superresolution technique is involved. Therefore, we define a unified metric, resolution frame-rate product to evaluate the actual performance when used with a mechanical scanner. For example, for real-time wide-field imaging, a line sensor (1-D array) with 256 pixels and 2560 fps is equivalent to a sensor with 256×256 pixels and 10 fps.

The spatial resolution or the number of pixels in an imager is one of the most important factors in most applications. It is related to the field of view and the details of the image. The details can be enhanced by magnifying optics but with the loss of macroscopic information or the field-of-view. Conversely, the details will be lost if the field-of-view is extended.

Many applications require at least 256×256 pixels for both field-of-view and image details. For example, in depth imaging, centimeter range resolution is required for proper interpretation of the geometry. Pixel array of 256×256 can resolve the subject at 3-m distance in 1.35-cm resolution with the viewing angle of 60 degrees. For intra-operative diagnostic imaging, large area of 2.5×2.5 cm^2 can be captured in 100 μm resolution which is enough for tumor removal considering that the accuracy of tumor removal is limited to 500 μm by a surgical hand [8].

Before we discuss frame-rate requirements, it is important to define related terms correctly to avoid confusion. “Frame-rate” and “time resolution” are confusing terms in time-resolved imaging and they have been distinguished in this thesis for clarity. Time resolution is used to describe the temporal accuracy for representing fast varying optical signal. On the other hand, the term frame-rate is used to describe the rate of the information obtainable by processing the time-varying optical signals. A typical time resolution for time-resolved imaging is sub-nanosecond, while the frame-rate ranges from sub-Hz to 10s Hz.

In order to provide real-time imaging, the frame-rate should be at least 10fps. Particularly, real-time imaging is very critical in depth imaging for human-computer interaction since the information should be processed from the video stream of the depth images. Also, the high frame-rate helps obtaining blur or artifact-free images caused by the motion. For example, *in vivo* imaging through endoscopy can be shaky and hence requires a high frame-rate.

1.2.2 Accuracy

The accuracy of the measurement is fundamentally limited by the photon shot-noise and the electrical noise. The accuracy is also dependent on the used number of intermediate images (intensity images) to get the final image (time image). In other words, the longer the time a sensor spends for more intensity frames, the more accurate results it could provide. This is a trade-off between the speed and the accuracy generally found in many sensing systems.

Most time-resolved image sensors, especially fluorescence lifetime imagers need to take 10s~100s samples to measure the accurate lifetime. This especially deteriorates the imaging speed. Sometimes, it even results in very inaccurate measurement due to the weakened fluorescence (called photo-bleaching) due to the long measurement time.

Therefore, it is critical to develop a new measurement method that provides accurate results without increasing measurement time.

1.2.3 Background Light Suppression

In active imaging, the measurement can be interfered by extraneous lights. This is particularly important for portable applications and fluorescence imaging. Portable devices can be exposed to many different extreme conditions like direct sunlight in a sunny day. In fluorescence imaging, the sensor has to measure the fluorescence while rejecting the excitation light directly from the emitter. Normally, the excitation light is measured around three orders of magnitude higher than the fluorescence.

In most cases, an optical band pass filter is used to filter out the background lights, but implementing a background suppression circuit in the sensor will further extend its usability even in the outdoor conditions.

1.3 Thesis Outline

This thesis presents CMOS time-resolved image sensors for high accuracy, real-time, and outdoor operation. Chapter 2 presents background research, architecture and building blocks of a CMOS image sensor. Chapter 3 discusses the noise sources in image sensors and their representation in time-domain to understand the accuracy limit of the time-resolved imaging. Two image sensors are proposed to address the challenges discussed in the previous sections. Chapter 4 presents a single-shot high speed fluorescence imaging to address the accuracy and imaging speed trade-off in time-gated imaging. Chapter 5 illustrates an adaptable background light suppression scheme for outdoor time-of-flight imaging. Finally, Chapter 6 summarizes this work and discuss about the future works. Appendix introduces a preliminary research for a multi-tap pixel that can provide advantages of both gated-imaging and TCSPC.

Chapter 2

CMOS Image Sensor Basics

2.1 Introduction

Camera system for photography consists of optics and an image sensor as shown in Fig. 2-1 (a). The optics, which are represented as a single lens in the figure, collect the lights from remote locations and focus the lights to form an image on the focal plane inside the camera system. The image sensor, placed at the focal plane, is an array of light-sensitive elements, which measure the brightness of the light at each location of the image.

Traditionally, light-sensitive films were used as image sensor until the advent of the digital camera system where a solid-state device such as charge-coupled devices (CCD) or complementary metal-oxide-semiconductor (CMOS) image sensors measure the image in digital format. At the onset of the digital imaging technology, the CCD sensor was preferred to the CMOS image sensor because of its superior image quality. Nowadays, CMOS sensors are dominant in the market thanks to the improved image quality and the ease of integration with other CMOS circuits to further enhance signal quality and provide additional features.

Fig. 2-1 (b) shows the block diagram and the signal chain of an CMOS image sensor. The unit cell in the sensing array is called a pixel (pictorial element). The photodiode (PD) in each pixel converts an optical signal into an electrical signal. The electrical signal is then conditioned for better digitization. Signal conditioning is typically done through correlated-double sampling (CDS) [12] and amplification. Finally, the signal is converted into the digital domain by the analog-to-digital converters (ADC). The digital image is then stored in a digital memory and transmitted outside of the sensor chip through the interfacing circuits.

In this chapter, we present the CMOS image sensor background knowledge. We first present the device structure of photodiodes and various pixel circuits. We also present image sensor architecture based on the array configuration of ADCs.

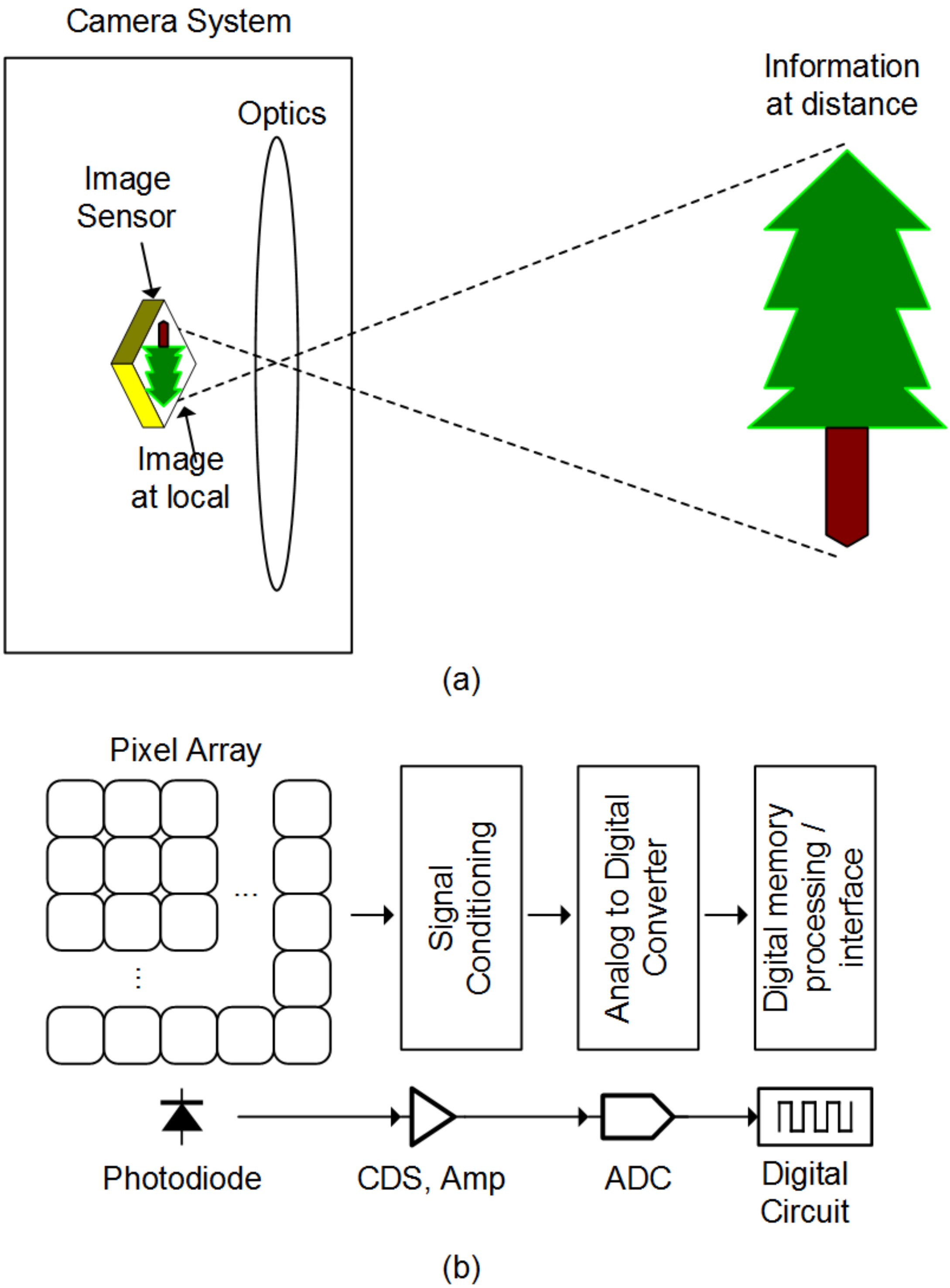


Fig. 2-1. (a) Camera systems for photography (b) signal change in CMOS image sensor.

2.2 Photodiode

A photodiode is used as an optical sensing device in most CMOS pixels. Fig. 2-2 shows the structure of a p-n photodiode. The most common method to create a p-n photodiode in CMOS process is to implant an n-layer on top of the p-type substrate. For photo-sensing, photodiode is usually reversed-biased with grounded anode and floated cathode. The reverse bias forms a depletion region and an electric field around the junction.

When a photon is absorbed by the silicon, an electron-hole pair is generated. If the generation happens within the depletion region, the electron and hole are separated by the electric field. Holes are drained by the ground and electrons are collected in the floating cathode.

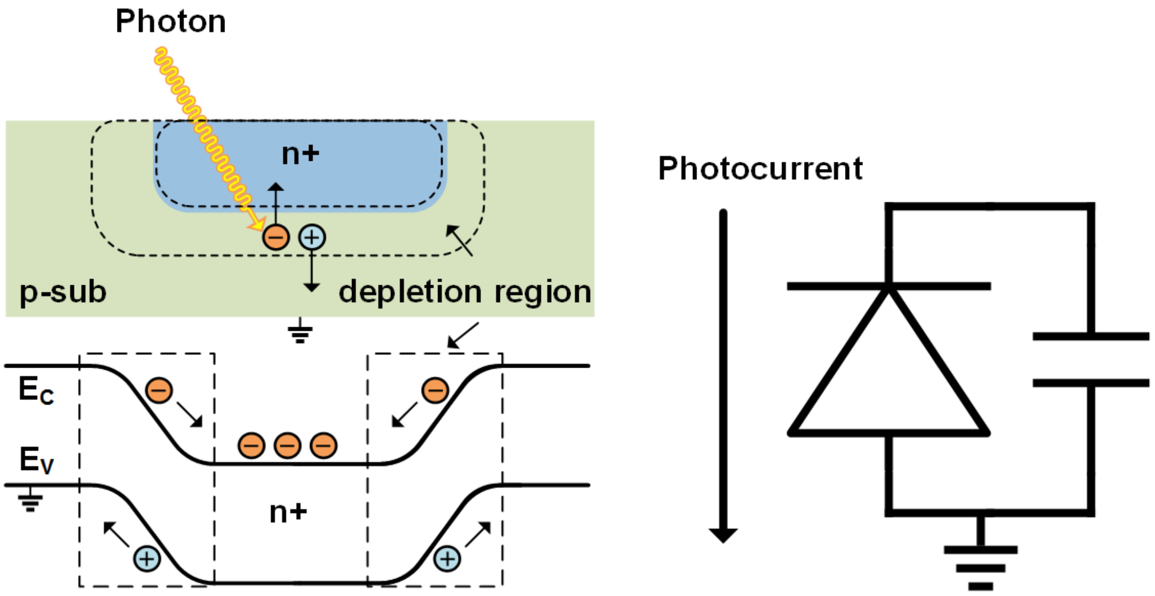


Fig. 2-2. Structure and modelling of p-n photodiode.

The charge separation is modeled as a photocurrent flowing in reverse direction of the diode. The potential well in the electron energy band diagram is the junction capacitance and modeled as a capacitor across the diode as shown in the schematic.

2.3 Operation of CMOS Pixels

Optical signal measurement starts with the photocurrent integration in CMOS pixels. Fig. 2-3 shows the photocurrent integration in a PD. For a proper operation, a reset transistor (RST) is attached to the PD cathode to set the initial condition of the diode voltage V_{PD} before starting the integration. The reset basically acts as a sampling operation adding kT/C noise. This noise is often referred to as reset noise in CMOS image sensors [13]. After the reset, the photodiode generates electrons from the photons and collects them in the potential well, discharging the capacitor during the integration time. After the integration, the difference between the initial voltage and the final voltage of the PD (specified as signal in the figure) measures the light intensity.

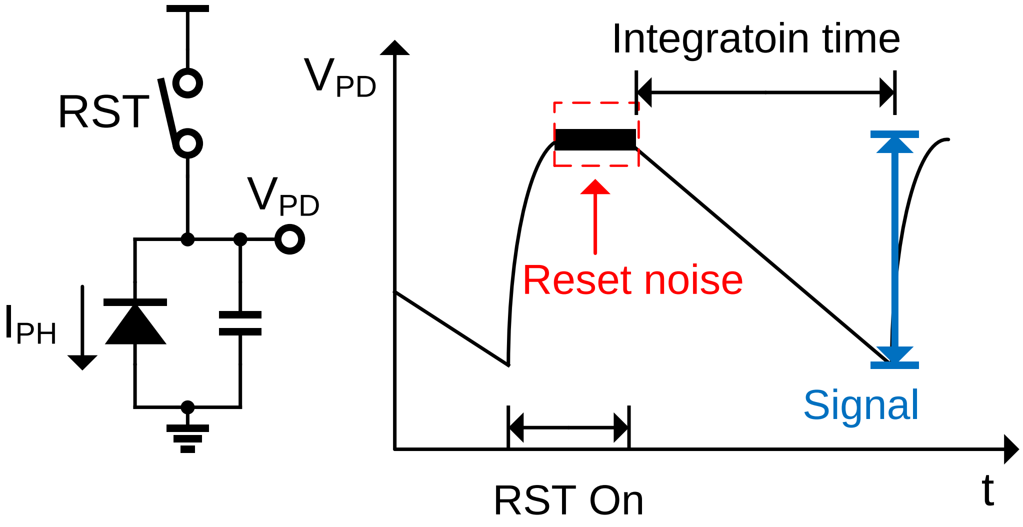


Fig. 2-3. Photocurrent integration in a photodiode.

The charge in the capacitor or the voltage across the diode can be measured by various pixel circuits such as passive pixel sensor (PPS) and active pixel sensor (APS). In the following sections, we present the operation of each pixel and discuss their advantages and disadvantages.

2.3.1 Passive Pixel Sensor

Fig. 2-4 shows a passive pixel sensor circuit. PPS utilizes a charge transfer amplifier to measure the amount of charge in the photodiode capacitor. Often the charge transfer amplifier is implemented in the column-level and shared by many pixels in the same column. The pixel has a switch between the photodiode and the column line. This transistor is used to reset the transistor and to multiplex the pixel signals to the amplifier.

When the switch is on, the photodiode voltage is set by the virtual ground voltage of the amplifier, V_B. After the integration time, the switch is turned on again, and the collected electrons are transferred to the feedback capacitor C_F of the amplifier. Therefore, the amplifier output voltage is given as:

$$V_{OUT} = \frac{\Delta Q}{C_F} = \frac{C_{PD}}{C_F}\Delta V$$

where ΔQ and ΔV represents the charge of photo-generated electrons and the voltage change by the photocurrent.

PPS requires only one transistor per pixel offering high fill factor (portion of the photodiode area in the pixel) and a small pixel design. However, PPS has problem with readout noise and scalability. As the number of pixels increase, the column parasitic capacitance, C_P, increases accordingly, resulting in slower readout speed or increased readout noise [14].

.

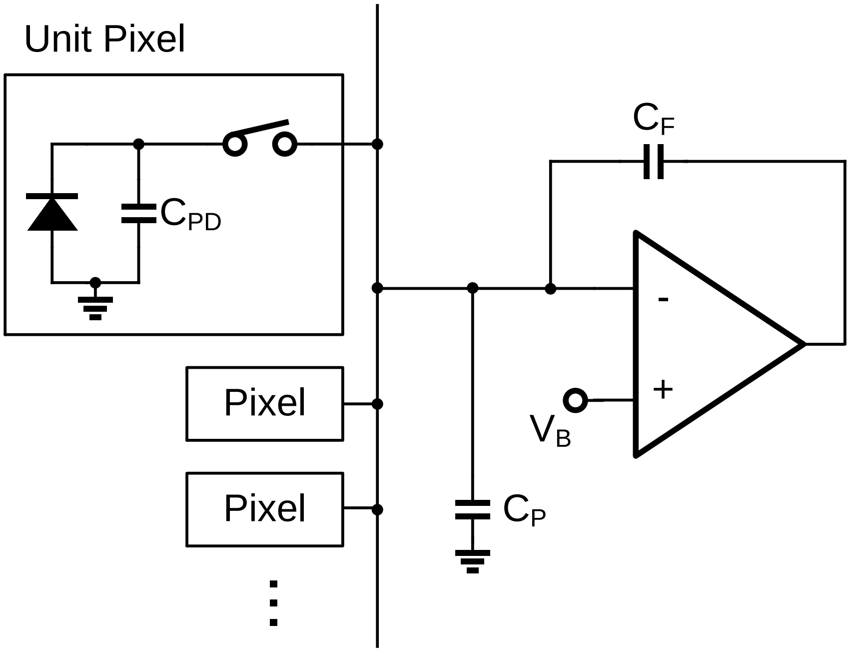


Fig. 2-4. Passive pixel sensor and readout circuit.

2.3.2 3-T Active Pixel Sensor

To overcome the limits of PPS, an amplifier can be inserted after the photodiode to drive the large column capacitance as shown in Fig. 2-5. It is called a 3-T active pixel sensor (APS) because each pixel has three transistors and signal is amplified in the pixel. The load current source of the amplifier is implemented in column-level to save pixel space and prevent static power consumption in inactive pixels. The SEL switch connects only one pixel per column at a time to the load current source. For the amplifier choice, source follower in Fig. 2-5 (a) is preferred over the common source amplifier in Fig. 2-5 (b) for pixel uniformity. Common source amplifier is less commonly used and mostly in special cases such as when high sensitivity is required [15].

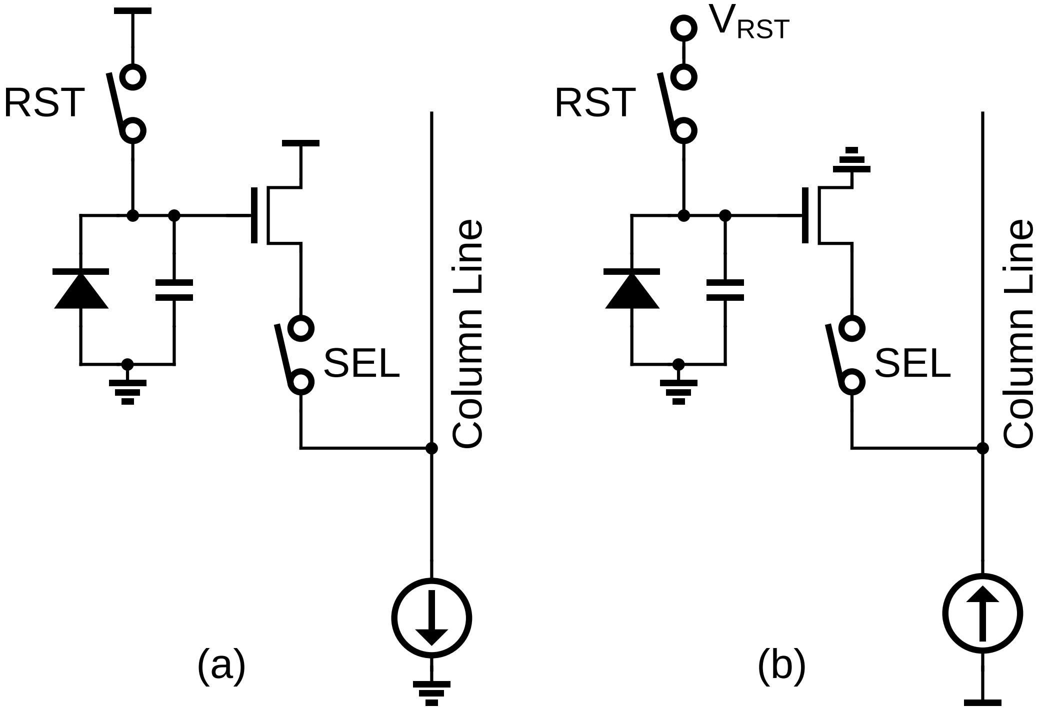


Fig. 2-5. 3-T active pixel sensor with (a) source follower (b) common-source amp.

The reset level of each pixel significantly varies across the pixel array mainly due to threshold voltage mismatch of the reset transistor. This reset level mismatch causes a fixed-patter noise (FPN) in the final image. Therefore, doubling sampling is typically performed for baselining; i.e. both reset level and signal level are measured to get the delta signal.

The double-sampling measurement starts with PD reset. The PD is pre-charged to a reset level to give an initial condition of the PD. The PD voltage is given as $V_{PD0} = V_R + \overline{V_{RN1}}$ where V_R is the reset voltage and $\overline{V_{RN1}}$ is the additional reset noise. After photocurrent integration, the signal voltage is then sampled by the readout circuit. The sampled signal is given as $V_{SIG} = V_R + \overline{V_{RN1}} - \Delta V$ where ΔV is the voltage change due to photocurrent. After that, the PD is reset and sampled for FPN removal, which is given as

$V_{RST} = V_R + \overline{V_{RN2}}$. It is worth noting that the sampling noise in the reset level V_{RST} is different from the sampling noise in the signal level V_{SIG}. Finally, the integrated photocurrent signal, ΔV, can be estimated by subtracting V_{SIG} from V_{RST} as:

$$\Delta V_{est} = V_{RST} - V_{SIG} = \Delta V + \sqrt{\overline{V_{RN1}^2} + \overline{V_{RN2}^2}}$$

Since the two reset noises are not correlated, their noise powers add up. This results in a poor SNR in 3-T APS.

2.3.3 4-T Active Pixel Sensor with Pinned-Photodiode

The reset noise between samples can be suppressed if it is correlated. To achieve this, a pinned-photodiode (pinned-PD) can be utilized as shown in Fig. 2-6 (a). Unlike the p-n photodiode used in 3-T APS, the cathode of the pinned-PD is buried in p-type semiconductor. This is done by implanting p+ layer on top of the n region. The additional p+ layer shifts the integration region (n-type) away from the surface which has defects that cause dark current. Thus, a buried photodiode can achieve extremely low dark current as compared to p-n photodiodes in 3-T APS. The reduced dark current is especially beneficial in low-light condition where the dark current shot noise can be a dominant noise source.

Besides the buried PD, the pixel also has a gate connecting the PD and an n+ region called a floating diffusion (FD). This gate is called transfer gate (TX) because it transfers the collected charges in the PD region to the FD region. The purpose of a pinned-PD is to separate the integrate and readout nodes; the charges are first integrated in the PD region, and then transferred to FD for readout. This separation of integration node and sensing node enables correlated double sampling (CDS) at the FD node, which can eliminate the reset noise of FD capacitance. The electron energy diagram across the line AB inside the photodiode is shown in Fig. 2-6 (a). Note that the potential of the fully depleted n-region at the bottom of the potential well is lower than the n+ floating diffusion (FD). Due to this potential difference, all charges on the PD can be transferred to FD when the transfer gate is turned on. This guarantees complete charge transfer, and the PD is pinned to the voltage called pinning voltage. The full depletion reduces the diode capacitance to zero and enables a noiseless reset of the PD region. The combination of a noiseless PD reset and a FD reset noise suppression through CDS enables virtually noiseless reset.

Fig. 2-6 (b) shows the complete schematic of 4-T APS. The circuit is similar to that of a 3-T APS except for the addition of a transfer gate (TX) at the pinned photodiode instead of a simple p-n photodiode. The operation of 4-T APS with pinned-PD begins with photodiode reset. Photodiode can be reset by simply opening the transfer gate. All charges in the PD region is drained by FD. The PD reset is noiseless thanks to the full depletion of the PD. After PD reset, the photodiode start integrating the photocurrent. The TX is off during this time and the electrons are collected in the potential well in PD region. After the integration, FD is reset and the reset value and reset noise are sampled as $\mathrm{V_{RST}} = \mathrm{V_R} + \overline{\mathrm{V_{RN}}}$. Right after the FD reset, charges are transferred from PD to FD, followed by the sampling of the signal level $\mathrm{V_{SIG}} = \mathrm{V_R} + \overline{\mathrm{V_{RN}}} - \Delta\mathrm{V}$. In 4-T APS, the sampled values $\mathrm{V_{RST}}$ and $\mathrm{V_{SIG}}$ have the same reset noise component $\overline{\mathrm{V_{RN}}}$ unlike in the 3-T APS case. The estimation of the optical signal is then given as: $\Delta\mathrm{V_{est}} = \mathrm{V_{RST}} - \mathrm{V_{SIG}} = \Delta\mathrm{V}$.

Thanks to the correlated noise, CDS completely cancels the FD reset noise and provides the best imaging quality across all types of CMOS pixels.

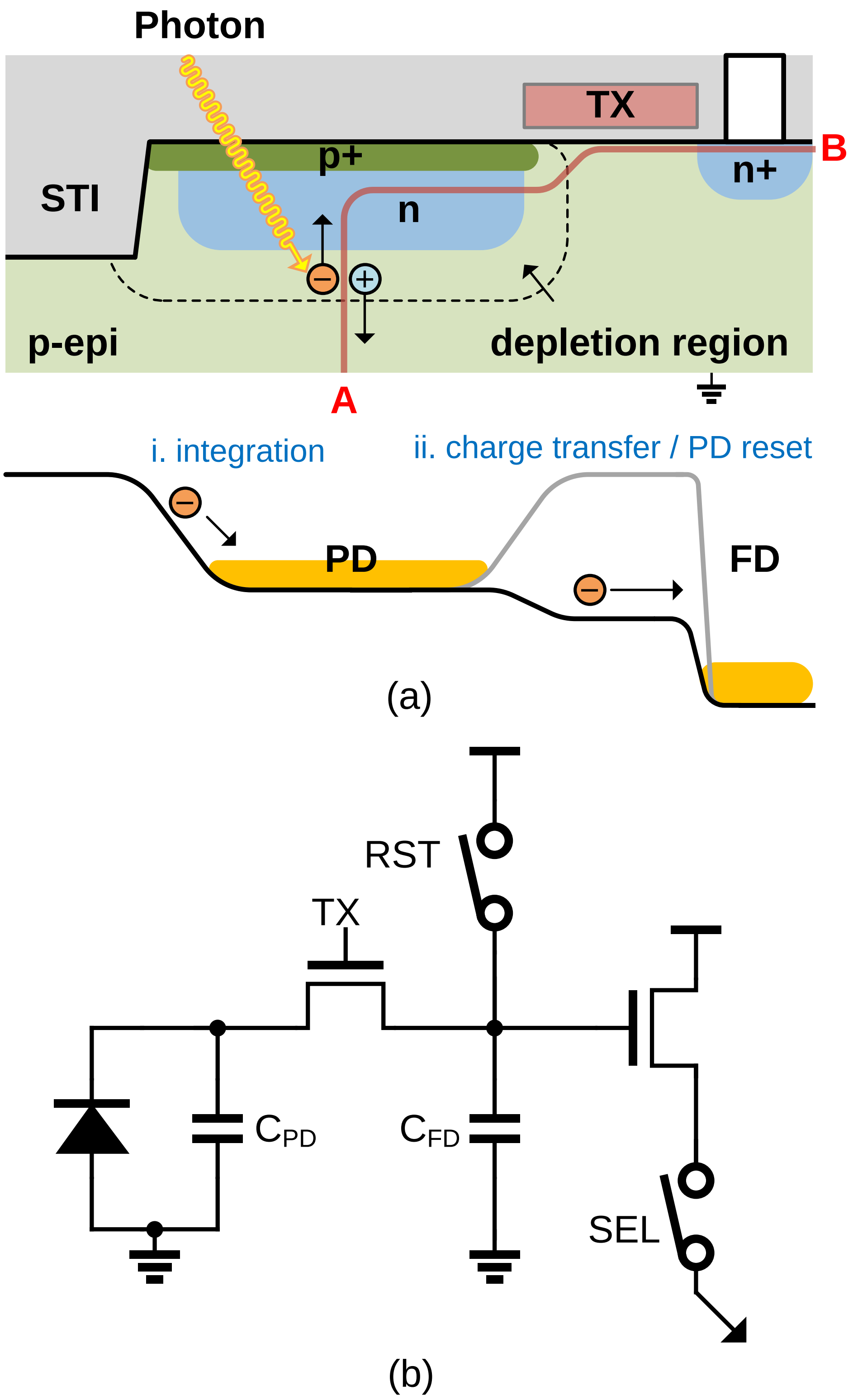


Fig. 2-6. (a) Structure and operation of pinned-photodiode (b) 4-T APS

2.4 CMOS Image Sensor Architecture

A basic CMOS image sensor system, consists of pixels, CDS circuits, and ADCs. The number of ADCs in the image sensor may vary depending on the implementation methods; e.g. one ADC in the chip-level can be shared by the whole pixel array; ADCs in the column-level can be shared by the pixels in the same column; or ADC is implemented per one or a few pixels in the pixel-level. This section introduces the three types of image sensor architecture.

2.4.1 Chip-level ADC

In the chip-level ADC architecture, a single ADC is used to convert the signals from the entire pixel array as shown in Fig. 2-7 (a). In this architecture, the column-level CDS and sample and hold circuits store the pixel values from a selected row. The sampled values are then accessed by the ADC one by one through the column scanner. The ADC then sequentially converts the pixel signals into digital format. After processing the row, the next row is selected for the next A/D conversion. This process is repeated until the whole pixel array is scanned.

In this approach, the speed of ADC determines the overall imaging speed. The ADC sampling-rate should be higher than the product of the resolution and the frame-rate. Therefore, high-speed ADC such as pipeline ADC is preferred. However, as high resolution and high frame-rate are demanded in the high-end digital camera market, the chip-level ADC architecture became less popular in CMOS image sensors industry.

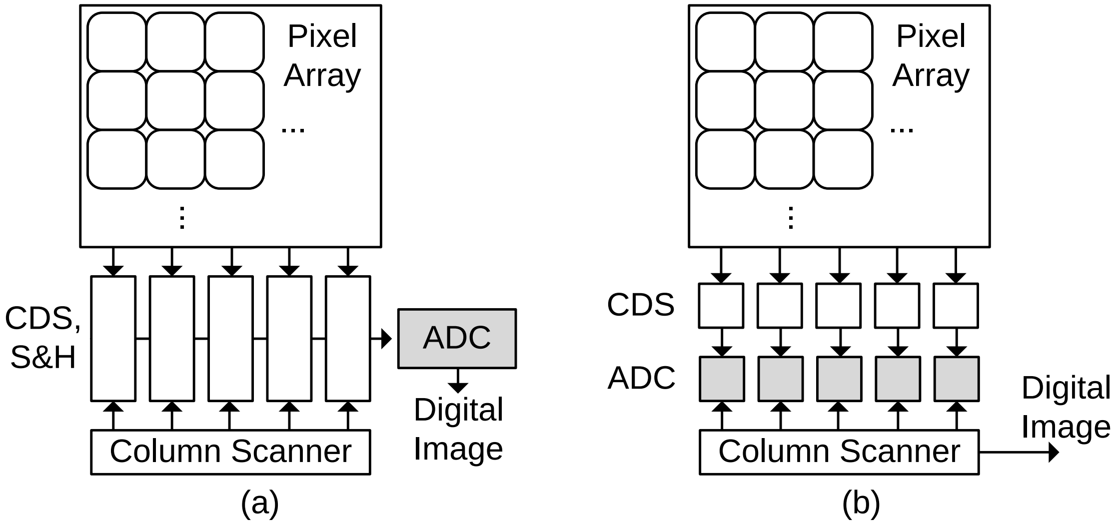


Fig. 2-7. Image sensor architecture (a) serial ADC (b) column-Parallel ADC.

2.4.2 Column-Parallel ADC

Fig. 2-7 (b) shows the column-parallel ADC architecture. ADC is implemented after the CDS circuit in the column-level. Due to the parallel operation of ADCs, the image sensor can achieve higher throughput than a chip-level ADC approach. Simple ADC structures such as single-slope ADC [16]–[18], cyclic ADC [19]–[21] and delta-sigma ADC [22] are preferred for this architecture due to the area restriction. This is currently the most popular architecture for CMOS image sensors because it offers small pixel design, high-resolution array, and high frame-rate.

2.4.3 Pixel-Level ADC and Processing Unit

In this architecture, a complex circuitry is implemented in the pixel. This approach, in general, yields a large pixel size and hence a low-resolution imaging array. Despite of this disadvantage, this approach is particularly interesting and important when extremely high speed imaging or image processing is required.

In [23], a pixel-level ADC, shared by four pixels achieved high dynamic range by multiple capture with different exposure times. The pixel-level ADC was fast enough to convert multiple samples in one frame time. In [24], inter-pixel winner-takes-all circuits are implemented in each pixel, to shrink or expand the pupil area in eye images for eye-tracking application. Thanks to the pixel-level analog processing, it achieved 5000 fps at extremely low power which were never achieved by digital processing approaches. In [25], the pixel level analog processor is implemented for high-speed processing of various computation-intensive operations such as asynchronous propagation and diffusions.

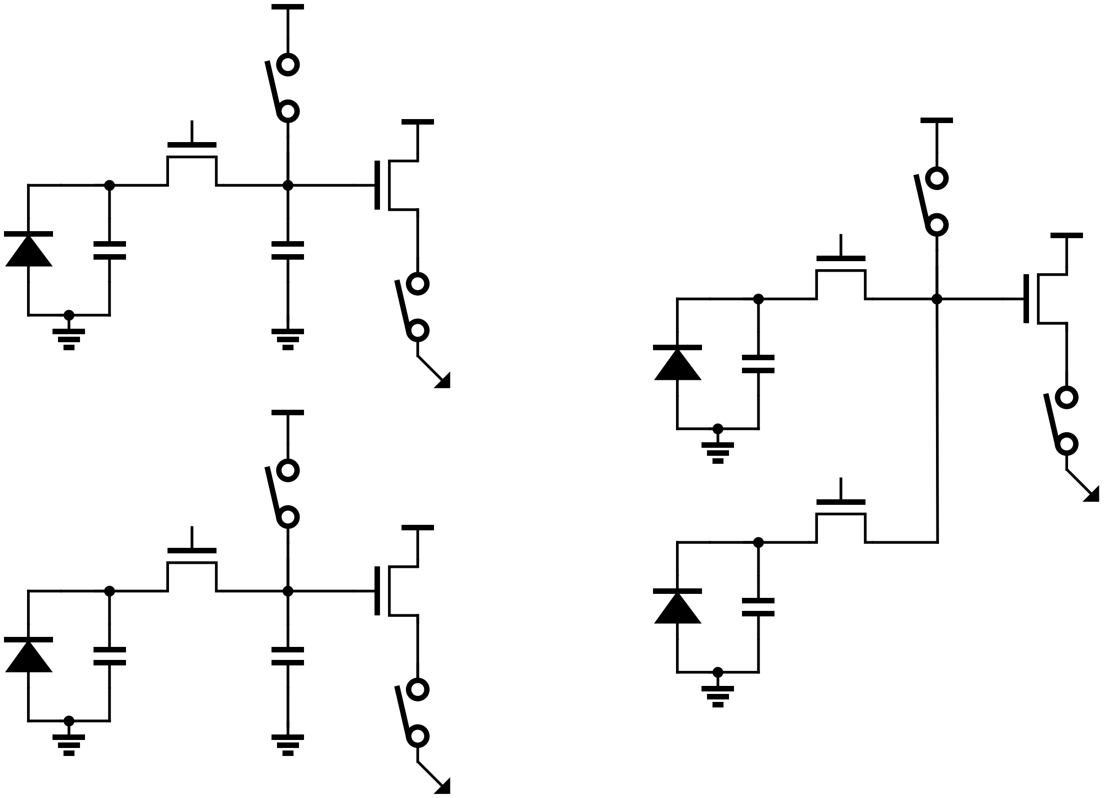
2.5 Pixel Performance

This section describes the key parameters of pixels that are useful in design and evaluation of image sensors.

2.5.1 Fill Factor

In CMOS image sensors, a pixel is composed of a photodiode and a few transistors such as reset transistor, select transistor, and a source follower. The circuit area is covered by metal to protect the circuit part from the photo-generation, whereas the photodiode area is exposed for photon reception.

Fill factor is defined as the ratio of the exposed photodiode region to the whole pixel area. In 4-T APS, larger fill factor means higher sensitivity because the pixel receives more photons. To increase the fill factor in 4-T APS, the pixel transistors except for transfer gates can be shared by multiple pixels. Fig. 2-8 shows an example of this shared architecture where two 4-T APSs can equivalently be implemented in 5-T shared structure. Besides, advanced image sensor fabrication processes offer backside illumination (BSI) technology to further increase the fill-factor. In this technology, the lights are coming from the bottom-side of the substrate, maximizing the fill factor regardless the metal routing in the front-side.



(a) (b)

Fig. 2-8 (a) Two 4-T pixels (b) equivalent 5-T pixels in shared architecture.

2.5.2 Dark Current

Under completely dark conditions, the photodiode still experiences a leakage current called a dark current. The dark current causes fixed pattern noise and adds shot noise deteriorating the SNR especially in low-light conditions. The dark current mechanisms are categorized into two: the reverse-bias leakage current, and the surface generation current. Because the reverse-bias leakage current tends to increase with temperature, cooling mechanisms are sometimes employed in imaging systems to suppress the dark current. The second mechanism is well suppressed by a buried-type diode such as pinned photodiode, as explained in 2.3.3.

2.5.3 Full-Well Capacity

The full well capacity is defined as the electron capacity of the pixel. When more electrons than the capacity are introduced, the pixels are saturated and cannot accept the excessive electrons. In 4-T pixels, the full-well capacity is limited by the FD capacitance and the voltage swing of FD, while in 3-T pixels, it is determined by the PD capacitance and the voltage swing of PD. The amount of charge or the number electrons the capacitor can hold is given as:

$$\mathrm{N_{SAT}} = \frac{\mathrm{C_{PD;FD}V_{SWING}}}{\mathrm{q}}\ [\mathrm{electrons}]$$

where, q is the charge of a single electron (1.6e-19 C).

Larger full-well capacity enables the measurement of higher illumination, providing a higher dynamic range. Full well capacity is however limited by the technology and the pixel size. However, the full-well capacity can effectively be increased by various sensing techniques such as a multiple capturing method [23], [26].

2.5.4 Sensitivity

The sensitivity is defined as the ratio of the output voltage to the input illumination level. The unit of sensitivity is $[\mathrm{V/lx \cdot s}]$. The sensitivity in 4-T APS is dependent on various factors such as the quantum efficiency, conversion gain, and the amplifier gain. The quantum efficiency is the ratio of collected electrons to the incident photons. One can use a smaller FD capacitance to increase the conversion gain; this, however, decreases the full well capacity, requiring the designer to accommodate for the trade-off. The pixel amplifiers are typically source followers with a gain of less than unity, but common source amplifiers are also used in CMOS image sensors at the expense of increased fixed-pattern noise [15].

2.5.5 Dynamic Range

The dynamic range is defined as the ratio of maximum measurable optical power to the minimum resolvable optical power, or the noise floor of the image sensor. High dynamic range is very important in low-light imaging as well as in outdoor imaging. In low-light conditions, the ADC noise and dark current noise are dominant. Therefore, developing low-noise readout circuit is very important. Recent literature reported sub-electron noise performance for low-light imaging [27], [28]. For outdoor imaging, nonlinear photo-response is also often employed [29] to extend the maximum measurable light intensity. Other dynamic range extension techniques include dual- capture [26], multiple-capture [23], pixel-wise integration time control [30], [31], and time-domain measurement [32], [33].

2.6 Summary

In this chapter, the basic knowledge of CMOS image sensor has been described. First, the image sensor is introduced as a part of the camera system with the signal chain in the optical sensing system. The operation and the characterization of various pixel circuits are explored; first, PPS and 3-T APS are discussed with the introduction of p-n photodiode; then, pinned-photodiode, 4-T APS and CDS operation are explained in detail. This chapter also introduced three types of image sensor architecture depending on the ADC location in the system. Finally, pixel performance metrics are presented for evaluation and design of the image sensors.

Chapter 3

Noise in Time-Resolved Imaging

3.1 Introduction

In this chapter, we analyze the noise in CMOS time-resolved image sensors to provide a theoretical limit for designing and evaluating the time-resolved image sensors especially with two-tap pixels [34]–[42]. Although the noise in time-resolved image sensors originates from the same sources as in a CMOS image sensors for digital photography, the noise analysis for CMOS image sensors such as in [43] cannot be directly applied to time-resolved imaging due to the completely different noise characteristics in the final time images. The difference comes from the fact that the time-resolved imaging has an additional post-processing step that calculates the time image from multiple intensity images. This process of time-estimation translates the noise from the intensity domain into the time domain by non-linear equations.

In this chapter, we analyze the noise of two-tap measurement for the time-of-flight depth imaging and the fluorescence lifetime imaging. The noise sources are first identified in the intensity domain and then they are translated into the time domain. Finally, we verify the derived equation with a Monte Carlo simulation.

3.2 Noise in CMOS Image Sensors

The noise sources in CMOS image sensors are well identified in literature [43]. Although they originate from different random processes, they can be categorized into two types based on their physical behavior.

The first type is signal independent; i.e. the noise power is constant in all lighting conditions. Electrical noises such as pixel reset noise, amplifier noise, and A/D converter noise belong in this category. The noises are, for convenience, often referred to the sensing node, and then summed up as total noise. The total noise is measured in the root-mean-squared number of electron.

In contrast, the second type of noise, called a shot noise, is signal dependent; the noise power is proportional to the light intensity (i.e. number of photons). Shot noise is introduced during the photon-to-electron conversion. This photoelectron conversion follows Poisson process, so the variance and mean are the same. That is, when a photodetector measures, on average, N photoelectrons during exposure, the uncertainty of the measurement is characterized as a variance of N, or standard deviation of $\sqrt{N}$.

In this chapter, we will use the terms readout noise and shot noise to represent the two noise types, respectively. Also, we will use the terms noise power and noise to represent the variance and standard deviation of the signals, respectively.

Fig. 3-1 shows the signal and noise components with respect to the input light intensity. Units for x- and y-axes are the number of electrons. It is obvious that shot noise is smaller than the readout noise when the electron count is small (i.e. low-light condition). In contrast, in bright condition, the shot noise is much larger than the readout noise. Despite of the increased shot noise in bright condition, the SNR is large because the signal power increases faster than the shot noise as the light intensity increases.

The maximum SNR (40dB in this example) is determined by the well-capacity (N_{FW}) or the maximum number of electrons that the sensor can measure.

$$\mathrm{SNR_{MAX}[dB]} = 10 \cdot \log\left(\frac{\text{Signal Power}}{\text{Noise Power}}\right) = 10 \cdot \log\left(\frac{N_{FW}^2}{N_{FW}}\right) = 10 \cdot \log N_{FW}$$

Therefore, the design goal for high-quality intensity imaging is straightforward. To improve the low-light imaging quality, low-noise readout circuits must be used. To improve the maximum SNR in the brightest lighting condition, the well-capacity must be increased.

The dynamic range is defined as the ratio of maximum detectable light without sensor saturation to the minimum resolved light, which is read as 4e- in the figure. Therefore, the dynamic range of this example is given as $10^4/4$ or 64dB.

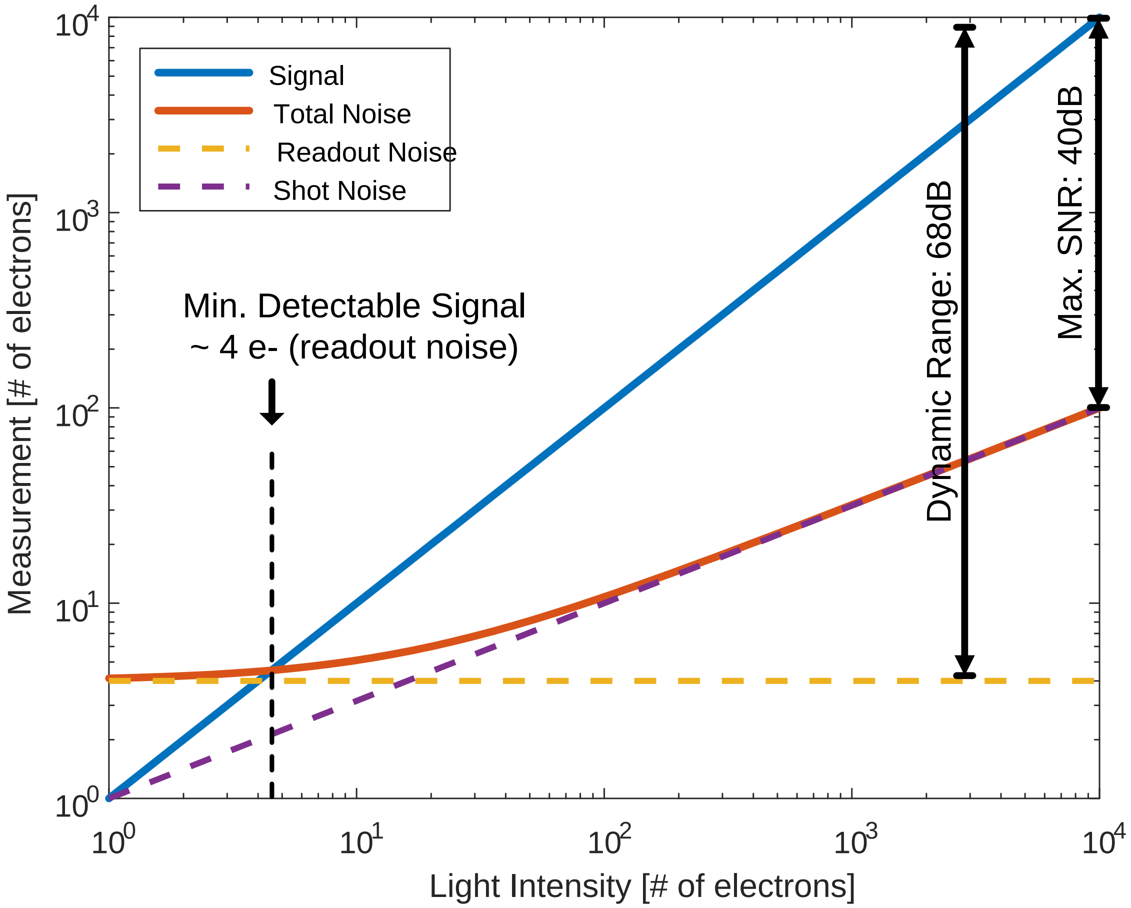


Fig. 3-1 Signal and noise vs. light intensity with full well-capacity of 10^4 and readout noise of 4 electrons.

3.3 Noise of Two-Tap Time-of-Flight Depth Imaging

In this section, we analyze the theoretical limit of the two-tap time-of-flight depth imaging. Fig. 3-2(a) shows a model of the two-tap measurement including the post-processing step. The optical signal $f(t, \tau, A)$ (system input) contains the time-of-flight (τ), the amplitude of the optical signal (A), and the background light component (B). For simplicity, we assume that the background light component is either suppressed or non-existing. The two-tap pixel integrates the input signal using two time gates and provides two outputs (X and Y in the figure), containing the variables, τ and A. To measure the time information, τ, we first remove the amplitude A by normalizing the signals. The normalized value Z is only function of τ, not necessarily in a linear fashion. If Z is a nonlinear function of τ, a linearization step is required at the end of the process.

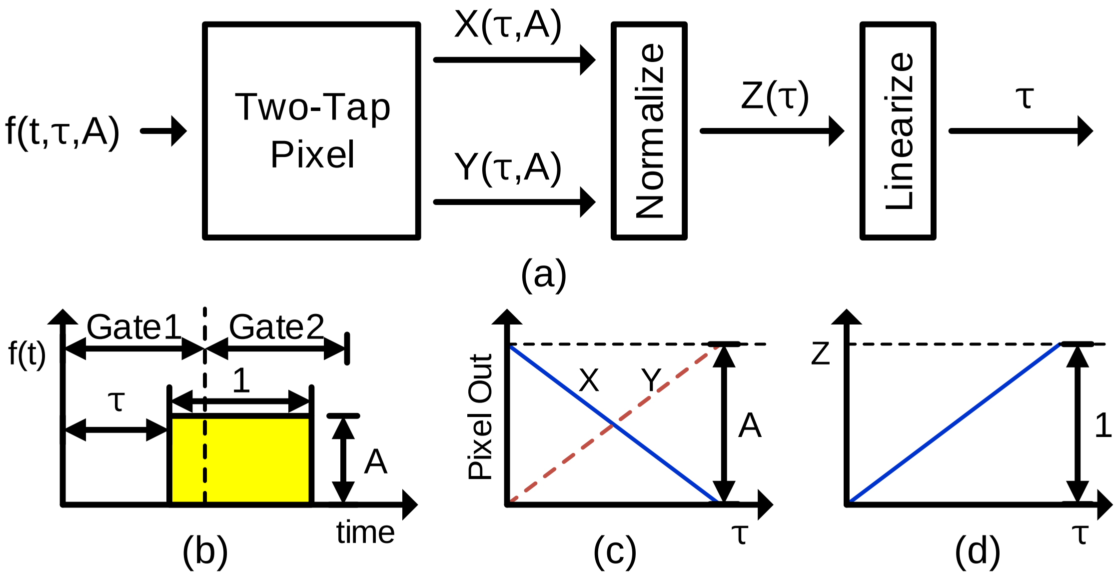


Fig. 3-2 (a) Block diagram of two-tap measurement; (b-d) signal plots in time-of-flight measurement.

For time-of-flight depth imaging, the signal light is often chosen to be a rectangular pulse. For convenience, we assume a perfectly rectangular light pulse as in Fig. 3-2(b). The mathematical form of the returned optical signal is given as:

$$f(t, \tau, A) = A \cdot \mathrm{rect}(t - 0.5 - \tau)$$

where A is the amplitude of the pulse and τ is the time-of-flight to be measured. In this analysis, the pulse width is normalized to 1, so the range of τ is from 0 to 1. Later, the result can be accordingly scaled for actual pulse widths.

Two-tap pixel integrates the returned optical signal $f(t, \tau, A)$ with two time-gates. Here, we define two complementary time-gates of [0,1] and [1,2]. The output values of the pixel, X and Y, plotted in Fig. 3-2(c), are then given as:

$$X(\tau, A) = \int_0^1 f(t)dt = A(1 - \tau) + \overline{\sigma_X}$$

$$Y(\tau, A) = \int_1^2 f(t)dt = A \cdot \tau + \overline{\sigma_Y}$$

where $\overline{\sigma_X}$ and $\overline{\sigma_Y}$ represent the noise in the intensity imaging. Then, to solve for τ, the amplitude A is eliminated by normalizing the signals as in the following equation:

$$y = \frac{Y}{X + Y} = \tau$$

Plotted in Fig. 3-2(d), y is equal to τ in this example, i.e. time-of-flight measurement with perfect rectangular pulse.

The normalization equation above translates the noises of X and Y (σ_X^2 and σ_Y^2, respectively) into the noise of y (σ_y^2) as in the following equation:

$$\sigma_y^2 = \sigma_X^2 \cdot \left(\frac{dy}{dX}\right)^2 + \sigma_Y^2 \cdot \left(\frac{dy}{dY}\right)^2 = \frac{Y^2 \sigma_X^2}{(X + Y)^4} + \frac{X^2 \sigma_Y^2}{(X + Y)^4}$$

In the following section, we examine the translation of each noise component (i.e. readout and shot noises discussed in the previous section) from intensity images to time domain.

3.3.1 Readout Noise and Shot Noise

To evaluate the readout noise of the CMOS image sensor, we substitute both σ_X^2 and σ_Y^2 with a single term, σ_{RD}^2 because the readout noise powers for both X and Y are the same. The readout noise of y is then expressed in terms of X, Y and σ_{RD}^2 as:

$$\sigma_{y;RD}^2 = \sigma_{RD}^2 \cdot \frac{X^2 + Y^2}{(X+Y)^4}.$$

To quantify the effect of noise translation by the post-processing, noise figure (NF) is defined as the increased noise power by the process.

$$NF = \frac{\sigma_{y;RD}^2}{\sigma_{RD}^2} = \frac{X^2 + Y^2}{(X+Y)^4}.$$

The shot noise of X and Y are not the same unlike the readout noise because they are signal dependent. The noise powers, $\sigma_{X;SH}^2$ and $\sigma_{Y;SH}^2$, are given as the electron counts of the signals (N_X and N_Y.) Therefore substituting (X, Y, σ_X^2, σ_Y^2) with (N_X, N_Y, N_X, N_Y) gives the shot noise power of y as:

$$\sigma_{y;SH}^2 = \frac{X \cdot Y}{(X+Y)^3} \cdot \frac{1}{N_X + N_Y} = \frac{X \cdot Y}{(X+Y)^3} \cdot \frac{1}{N_{PH}}$$

where N_{PH} is the total electron count.

For shot noise, a common figure of merit [44], the photon economy or the F-value, is often used to evaluate and compare the measurement method. It is defined as the ratio between the relative noise ($\sigma_{y;SH}/y$) in time domain and the relative noise in intensity domain. The relative noise or the SNR by the shot noise in intensity domain is given as

$\sqrt{\mathrm{N}_{\mathrm{PH}}}$. The F-value, similar to the noise figure, measures the noise increase due to the intensity-to-time domain translation. For simplicity, the square of F number, F^2 is defined as:

$$\mathrm{F}^2 = \sigma_{y;\mathrm{SH}}^2 \cdot \mathrm{N}_{\mathrm{PH}} = \frac{\mathrm{X} \cdot \mathrm{Y}}{(\mathrm{X} + \mathrm{Y})^3}$$

The variables NF and F^2, can be represented in terms of the estimated time-of-flight value $\mathrm{y} = \mathrm{Y}/(\mathrm{X} + \mathrm{Y})$, by substituting X and Y with $(1 - \mathrm{y})$ and y, respectively, as in the following equations:

$$\mathrm{NF} = 2\mathrm{y}^2 - 2\mathrm{y} + 1, \qquad \mathrm{F}^2 = \mathrm{y} \cdot (1 - \mathrm{y})$$

In summary, the noise figure and F^2 is defined as the increased noise power by post-processing for readout noise and shot noise, respectively. The total noise around the measurement point y is then given as:

$$\mathrm{N}_{\mathrm{TOT}} = \mathrm{NF}(\tau) \cdot \sigma_{\mathrm{RD}}^2 + \mathrm{F}^2(\tau)/\mathrm{N}_{\mathrm{PH}}.$$

NF and F^2 are both plotted in Fig. 3-3. As expected, they differ from the noise characteristics of intensity noise plotted in Fig. 3-1. The readout noise reaches minimum when X is equal to Y and maximum when either X or Y equals 0. Shot noise has the opposite behavior; the minimum F^2 is found when X or Y equals 0 and maximum occurs when X is equal to Y.

In good lighting conditions, the accuracy of the time-of-flight imaging is limited by the readout noise. However, in intensity imaging, if the signal component, N_{PH}, is large enough, the readout noise is drowned out by the signal, making the SNR limited solely by the shot noise. However, in time imaging, the N_{PH} does not represent the signal and the readout noise cannot be ignored. Instead, when N_{PH} is large and the shot noise component becomes negligible due to small F^2, the accuracy becomes limited by the readout noise.

Under low-light conditions, the accuracy of the time-of-flight imaging is limited by shot noise unless the difference between X or Y is very large where the F-number becomes zero, yielding shot-noise free measurement.

To verify the theoretical results, we performed a Monte Carlo simulation with 5,000 repetitions, shown in the scatter plot in Fig. 3-3. The simulation results prove the validity of the noise analysis above. The same results are reported in other literatures such as [41].

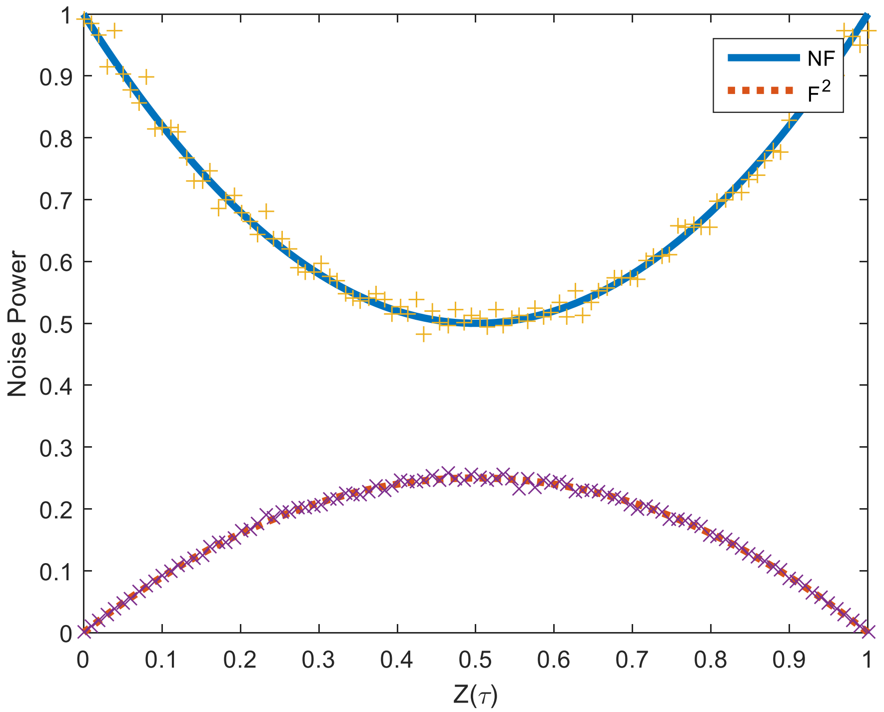


Fig. 3-3 Noise power in two-tap time-of-flight measurement; theoretical results are plotted in solid lines and Monte Carlo simulation results in markers.

3.3.2 Depth Noise

Finally, the derived equations for time noise can be represented in the depth domain. In time-of-flight depth imaging, the depth is calculated from the time as following:

$$\mathrm{D} = \frac{\mathrm{c} \cdot \tau}{2} = 15 \cdot \tau \cdot \mathrm{T_P}\ [\mathrm{cm}]$$

where c is the speed of light, and T_P is the pulse width of the light. Since the maximum τ is 1, the maximum measurable depth is $15 \cdot \mathrm{T_P}\ [\mathrm{cm}]$. Therefore, the depth noise is

$$\sigma_{\mathrm{D}} = 15 \cdot \sigma_\tau \cdot \mathrm{T_P}.$$

It is seen from the above equation that small depth error can be achieved by short pulse width (short T_P), but the maximum measurable depth or the dynamic range is reduced. For this reason, the pulse width should be chosen considering the applications. For example, [45] used impulse-like laser pulse in order to achieve high depth-resolution of 0.3mm. In other cases such as gesture sensing, the pulse width from 10ns to 50ns is adequate for a cm-resolution with the dynamic range of a few meters [37], [39].

In depth imaging, we also have to consider that the number of measured photoelectrons varies with depth. At large depth, the sensor receives less photons and has less electrons due to the beam divergence. Therefore, N_{PH} is a function of the depth D:

$$\mathrm{N_{PH}}(\mathrm{D}) = \mathrm{N_{FW}} \cdot \left(\frac{\mathrm{D_{MIN}}}{\mathrm{D}}\right)^2.$$

where N_{FW} is the full well-capacity, or the maximum measurable N_{PH}, and D_{MIN} is the minimum measurable depth when the electron count N_{PH} is the maximum (N_{FW}). Best accuracy is achieved at the depth of D_{MIN}, but as the depth becomes greater the accuracy decreases due to the reduced electron count.

3.4 Noise of Two-Tap Fluorescence Lifetime Imaging

In fluorescence lifetime imaging, the input light signal is an exponential decay as shown in Fig. 3-4(a). It can be written in a mathematical form as:

$$f(t) = A \cdot \exp(-t/\tau)$$

where τ is the lifetime to be resolved and A is the amplitude of the signal. The background light component is ignored similar to the case of time-of-flight imaging for simplicity. The two-tap pixel integrates the light from time 0 to h for the first value, and h to infinity for the second value, producing two outputs X and Y.

$$X = \int_0^h f(t) \cdot dt, \qquad Y = \int_h^\infty f(t) \cdot dt$$

We define the lifetime relative to h, by setting h to be 1. The real τ value can be obtained later by simply scaling the resulting lifetime values with the used h value in the measurement. Then the pixel outputs are given as:

$$X = \int_0^1 f(t) \cdot dt = A \cdot \tau\left(1 - e^{-1/\tau}\right)$$

$$Y = \int_1^\infty f(t) \cdot dt = A \cdot \tau\left(e^{-1/\tau}\right).$$

Fig. 3-4(b) shows the plot of pixel outputs, X and Y with respect to τ. To remove the amplitude dependency from the measurement, y is calculated as

$$y = \frac{Y}{X + Y} = e^{-1/\tau}$$

So far, the processing is similar to time-of-flight. However, y is now a nonlinear function of τ as shown in Fig. 3-4(c), translating the noise characteristic into a more complicated form.

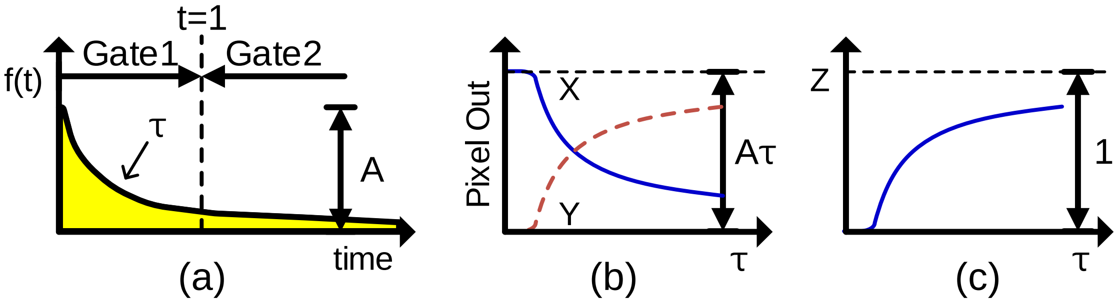


Fig. 3-4 Signal plots in fluorescence lifetime measurement (a) input light (b) pixel outputs (c) normalized pixel output.

3.4.1 Readout Noise and Shot Noise

The noise components of two-tap time-of-flight imaging were already characterized in the previous section as:

$$\mathrm{NF_y} = 2y^2 - 2y + 1, \qquad \mathrm{F_y^2} = y \cdot (1 - y).$$

However, the normalized pixel output has now a different function $y = e^{-1/\tau}$. Since y is a nonlinear function of τ, getting τ from Z requires an additional linearization step such as $\tau = -1/\ln(y)$. The noise translation by this linearization is then

$$\left(\frac{d\tau}{dy}\right)^2 = \frac{1}{y^2 \ln^4(y)} = \tau^4 \cdot e^{2/\tau}$$

Therefore, NF and F^2 are then given as:

$$\mathrm{NF} = \frac{\mathrm{NF_y}}{\tau^2} \cdot \left(\frac{d\tau}{dy}\right)^2 = \tau^2\left(e^{2/\tau} - 2e^{1/\tau} + 2\right)$$

$$F^2 = \frac{F_y^2}{\tau^2} \cdot \left(\frac{d\tau}{dy}\right)^2 = \tau^2 \cdot \left(e^{1/\tau} - 1\right)$$

The derived noise equations for NF and F^2 are plotted in Fig. 3-5. Monte Carlo simulation result is also overlaid in the same figure to verify the derived equations.

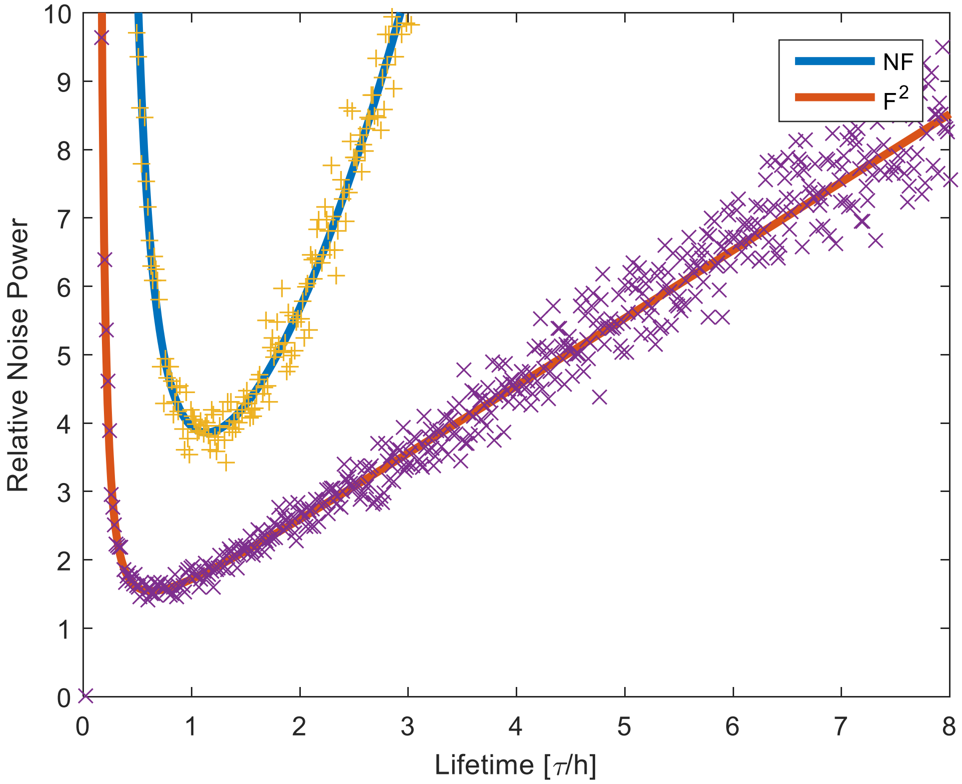


Fig. 3-5 Noise components in two-tap fluorescence lifetime imaging, readout noise (NF) and shot noise (F^2); theoretical results are in solid line and Monte Carlo simulation results in markers.

Unlike the time-of-flight computation, the noise in the lifetime measurement is not bounded; i.e. it can diverge to infinity for some τ values. From the derived noise characteristics, the NF is the main factor that limits the dynamic range of the sensor because the relative noise power of NF increases much faster than that of F^2. In other words, fluorescence lifetime imaging with two-tap integration method provides good accuracy only for the lifetime of between 0.5x and 2.5x of the time-gate width. The only way to achieve a good SNR in two-tap fluorescence lifetime imaging is to find the optimal h for a given lifetime which has been stated in previous literature [46].

3.5 Summary

In this chapter, we identified and reviewed two types of noise in CMOS image sensors, the signal independent readout noise and the signal dependent shot noise. Then we provided a mathematic model for time-resolved imaging with two-tap pixels. The model explains the noise translation from intensity domain into time domain. After that, the noise equations are derived for time-of-flight imaging and the fluorescence lifetime imaging. Theoretical noise equations are plotted along with the Monte Carlo simulation results for verification.

Chapter 4

Single-Shot High-Speed Fluorescence Lifetime Imaging

4.1 Introduction

In the last decade, the use of fluorescence lifetime imaging microscopy (FLIM) had been expanded from lab bench applications towards real-world applications such as clinical use. Traditionally, fluorescence lifetime imaging served as a powerful tool in life sciences and biomedical engineering but was limited only to bench testing setup [1]. Recent studies have demonstrated that this technique can be very effective in intra-operative uses [47], [48]. In these demonstrations, the fluorescence lifetime of cells or tissue is imaged to distinguish cancerous cells from normal cells, helping the surgeons determine the boundary of cancer region.

Despite of a wide range of potential applications, FLIM system is not widely available due to high cost, large size, and slow imaging speed. The system typically requires a bulky and costly short-pulse laser system and an expensive high speed photo-sensor. In addition, most fluorescence lifetime imaging systems requires a very long photon acquisition time as well as a long post-processing for improved accuracy. For example, time-correlated single photon counting (TCSPC), which is the most accurate method, suffers from very slow measurement speed. Faster time-gated imaging systems are in general not as accurate and require multiple-sampling for improved accuracy at slower speeds.

Solid-state devices have been developed to replace the traditional expensive and bulky hardware with low cost and portable systems [49]–[56]. Solid-state laser such as vertical cavity surface-emitting laser (VCSEL) is much more compact and as fast as the traditional Ti:Sapphire lasers, which is bulky, expensive, and difficult to maintain and operate. On the sensor side, solid-state devices again surpass the performance of traditional devices. The single-photon avalanche diodes (SPAD) and CMOS active pixel sensors (APS) are faster, cheaper and more compact than their traditional counterparts, the photo-multiplier tube (PMT), and the intensified charge coupled device (ICCD), respectively. Therefore, for low-cost and small size implementation, we focus on solid-state devices for both emitters and sensors.

For real-time and wide-field applications, time-gated imaging is the only viable option, as TCSPC has fundamentally slow imaging speed and low number of pixels. Although the imaging speed of TCSPC using semiconductor devices has been improved by multi-channel integration and advanced semiconductor technologies [49], it is still far too low for real-time operation or high resolution imaging.

By tackling the challenges mentioned above, this work focuses on implementing and improving the solid-state time-gated imaging systems. We propose a single-shot center-of-mass method, which is a new time-gated imaging scheme that improves the accuracy without losing the imaging speed or increasing the hardware complexity. The proposed method uses a distinctive integration method that acquires and processes multiple frames of the optical signal within one time frame to greatly improve the accuracy, especially when the system response time is much slower than the fluorescence lifetime of the samples.

In the following sections, we will review the limitations of conventional time-gated imaging for fluorescence lifetime imaging and compare it with time-of-flight imaging. Then we will propose a new time-gated imaging methodology that achieves both fast imaging speed and high accuracy. Finally, we will show the experimental results with two different fluorescent samples, fluorescein and green fluorescence protein (GFP) cells. The measurement of proposed method is compared with conventional time-gated imaging methods.

4.2 Fluorescence for Lifetime Measurement

4.2.1 Sampling Theory for Single Exponential Fluorescence

Fluorescence is known to exponentially decay if the excitation light is an infinitesimally short pulse. In most cases, when the pulse-width of the excitation light is sufficiently shorter than the fluorescence lifetime, the pulse-width no longer affects the response of the fluorescence and the short pulse can be considered as impulse excitation. The fluorescence waveform after impulse-excitation is given by

$$f(t) = \frac{A}{\tau} \exp\left(-\frac{t}{\tau}\right) \tag{1}$$

where A is the optical energy of fluorescence and $\tau$is the fluorescence lifetime.

To measure the lifetime, τ, we need to acquire samples from the waveform. First, we can assume a very densely sampled optical pulse. The measured data points can then be curve-fitted to equation (1) in order to find the lifetime. This method appears straightforward and accurate, but in theory, we only need two samples because there are only two variables in the equation (1).

Fig. 4-1 (a) shows how fluorescence lifetime is calculated from two samples. Two finite time-gate integrations are defined and the fluorescence light is integrated over each time-gate period. A two-tap pixel, such as in [34]–[42], can be used to measure the two integration values. The ratio of the integrated values, X and Y, varies depending on the lifetime as shown in Fig. 4-1 (b). Then, the analytic solution for the lifetime is calculated from X and Y as in [57]:

$$\tau = \frac{\Delta t}{\ln X/Y}. \tag{2}$$

This method is well-known as rapid lifetime determination (RLD) method. The time-gates can be overlapped and have different widths, but the analytic solution for τ can be determined [57].

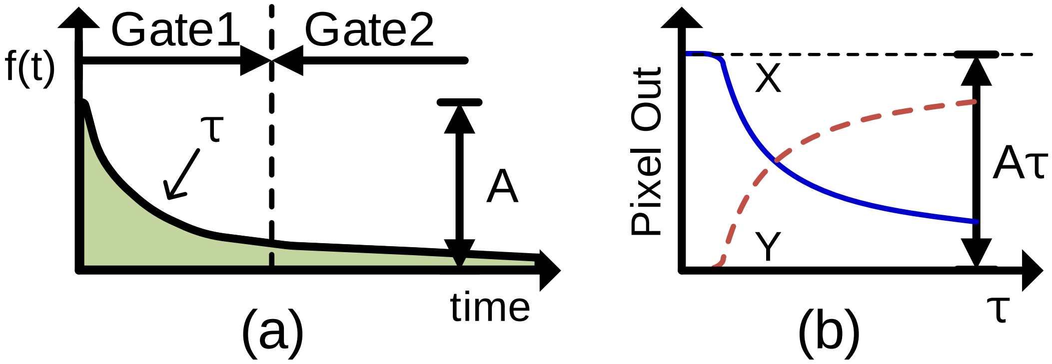


Fig. 4-1. Two-sample measurement of fluorescence for rapid lifetime determination.

This two-sample measurement or RLD has been demonstrated with solid-state devices with ultra-short pulse laser used as an excitation source [55]. The lifetime from a two-sample measurement showed very accurate results for green quantum dots with a long lifetime (10s of nanoseconds); however, it showed very inaccurate results for short lifetimes (a few nanoseconds) as shown in Fig. 4-2.

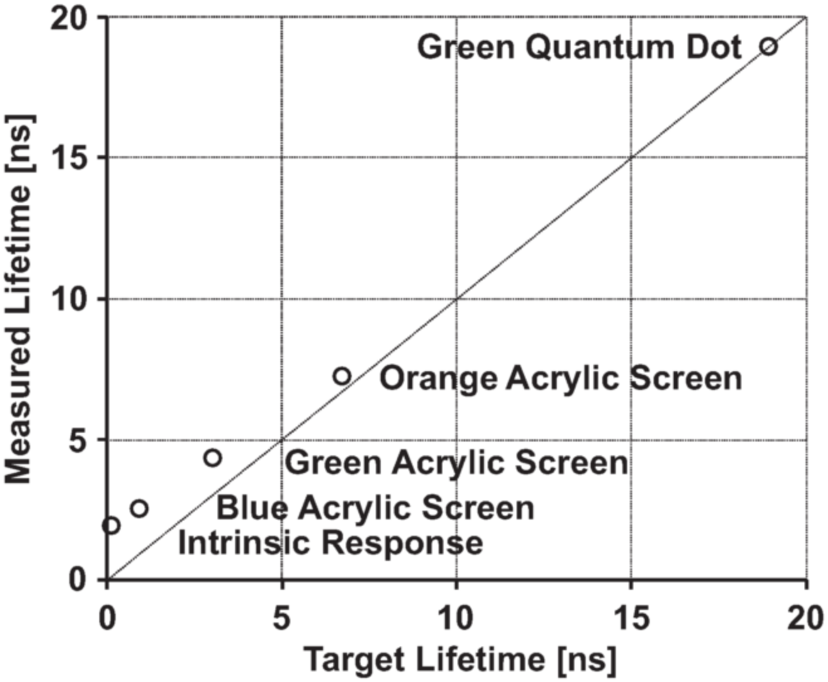


Fig. 4-2. Measured lifetime vs. ideal lifetime (image from [55]).

The cause of these poor results is due to an imperfect impulse excitation source, which leads to an imperfect exponential decay necessary for two-sample measurements. Even with expensive laser systems, the instrument response function (IRF), or the waveform of the excitation light measured by the sensor, has a finite pulse-width as shown in Fig. 4-3. If this width is comparable to the lifetime of an object, the optical waveform is no longer like an exponential decay. In this figure, the response to the fluorescence looks like appears as a low-pass-filtered exponential decay. Thus, for short lifetimes or a sensing system with a long tail of IRF, the two-sample measurement is not sufficient to obtain accurate lifetime information; i.e. it is equivalent to under-sampling of the fluorescent image.

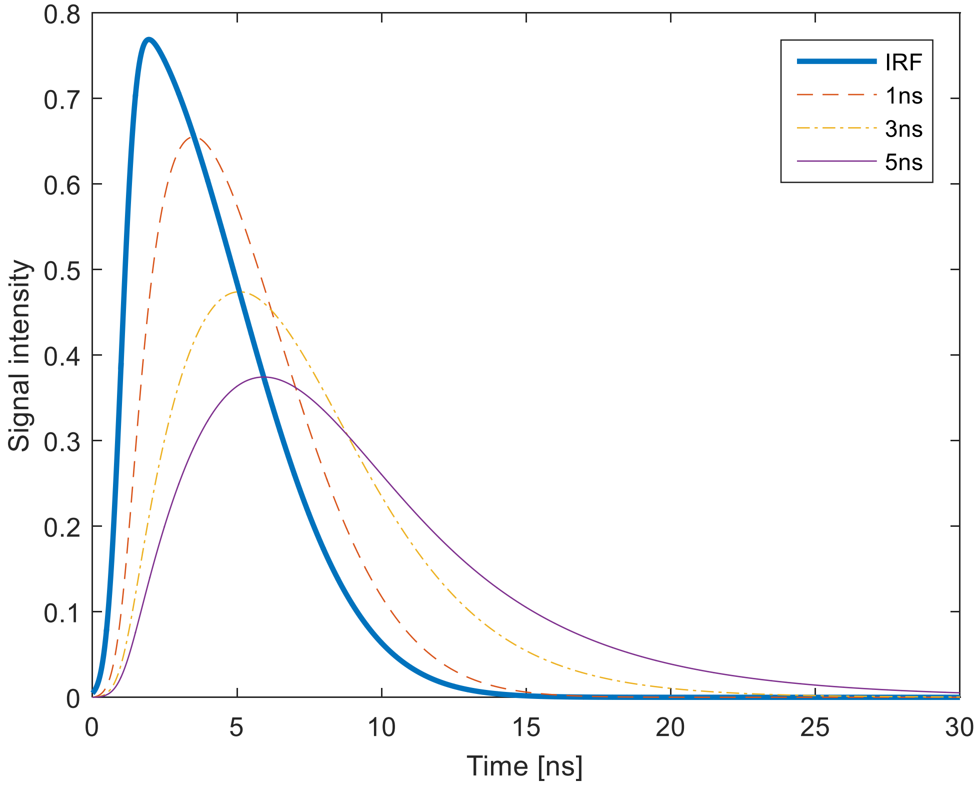


Fig. 4-3. System response to (i) the excitation light with a finite width (IRF), and (ii) fluorescence with three different lifetimes.

To address this under-sampling problem, full sampling of the signal is inevitable. In [52], lifetime was measured using both methods: a two-sample measurement (RLD method) and a full waveform sampling. It was shown that full sampling of the signal could provide a more accurate lifetime assessment.

From among many calculation methods for lifetime estimation using fully-sampled data, in the next section we will discuss a method called center-of-mass method (CMM). CMM provides an accurate estimation of lifetime. Then we will present a new measurement methods for CMM that does not require multiple-sampling for fast and accurate fluorescence lifetime measurement.

4.3 Center-of-Mass Method (CMM) for Lifetime Estimation

From among many algorithms that can deconvolve the effects of IRF from the fully sampled optical signals, center-of-mass method (CMM) is the best in terms of processing time [46]. Provided that the number of samples is sufficient, CMM can provide the same accuracy as the maximum-likelihood method, which provides the best accuracy but requires a very long iterative computation, and henceforth it is not preferred in real-time applications.

The CMM equations can be generally expressed as an average time-delay of photon arrival from the fluorescence signal, or $\tau_{\mathrm{CT}} = \mathrm{E}[\mathrm{t} \cdot \mathrm{f}(\mathrm{t})]/\mathrm{E}[\mathrm{f}(\mathrm{t})]$, where f(t) is the function that describes the optical waveform [46], [58], [59]. As the optical signals are sampled, a discrete time version can also be used as $\tau_{\mathrm{DT}} = \Delta\mathrm{t} \cdot \mathrm{E}[\mathrm{n} \cdot \mathrm{f}[\mathrm{n}]]/\mathrm{E}\big[\mathrm{f}[\mathrm{n}]\big]$, where $\Delta\mathrm{t}$ is the sampling period. The CMM equations for both continuous-time and discrete-time domains can be rewritten as

$$\tau_{\mathrm{CT}} = \frac{\int \mathrm{t} \cdot \mathrm{f}(\mathrm{t})\mathrm{dt}}{\int \mathrm{f}(\mathrm{t})\mathrm{dt}} \text{ and } \tau_{\mathrm{DT}} = \Delta\mathrm{t} \cdot \frac{\Sigma\, \mathrm{n} \cdot \mathrm{f}[\mathrm{n}]}{\Sigma\, \mathrm{f}[\mathrm{n}]}. \tag{3}$$

The CMM processing at an early stage of the signal chain greatly reduces the post-processing time and the data transmission time because CMM generates only one output from multiple inputs. It has been implemented as a post-processing part in [60]. In this case, the sensing devices first produce the full optical signal waveform, i.e. the sampled optical signal, f[n]. Then, a digital processor performs the discrete-time center-of-mass processing to estimate the lifetime information τ_{DT}. This processing naturally reduces the data size from N to 1, and as a result, the speed and power requirements for I/O devices are greatly reduced.

In another work, CMM, referred to as analog-mean-delay (AMD) method, has also been used to overcome the bandwidth limitation of the measurement equipment [58], [59], [61]. Here, the measured photon signals are intentionally elongated in the time domain by incorporating a 10^{th}-order Gaussian low-pass filter (GLPF) in the signal chain to reduce the bandwidth of photon signals. In this situation, the IRF is much longer than the lifetime and CMM is shown to be effective in deconvolving the IRF. Thanks to the elongated signal, a low frequency sampler was used to acquire the full waveform of the photon signal. After digitizing the sampled signals, lifetime was then calculated in digital domain by subtracting the mean-delay of IRF from the mean-delay of the elongated fluorescence signal using the discrete-time domain CMM equation.

Despite the reduced hardware requirements during and after the post-processing stage, the implemented CMM still does not improve the overall imaging speed. Instead, the system suffers from the increased signal acquisition time compared to the RLD method because it requires a full waveform measurement. For example, the sensor in [60] still relies on TCSPC, the slowest optical signal measurement method. In [58], [59], [61], 64 samples were required to calculate one sample of lifetime. The fluorescence lifetime camera based on this method will be 64 times slower than the conventional cameras using the same electronics.

In addition, full-waveform measurement with time-gated imaging is not recommended in many fluorescence applications where the fluorescence is very weak and suffers from photo-bleaching. Fig. 4-4 shows the intensity image of green fluorescence protein (GFP) cells measured by one of our custom-built cameras with an integration time of 253ms. The integration time is set relatively long to get a sufficient signal intensity level (maximum value of 157 out of 256 in the brightest pixel.) Measuring 10~100 frames of image for the full waveform measurement will limit the frame-rate by 0.04Hz~0.4Hz. Moreover, a measurement of multiple frames eventually deteriorates in signal quality because of a phenomenon called photo-bleaching; the intensity of the fluorescence is getting weaker as we measure multiple consecutive frames as shown in the Fig. 4-5.

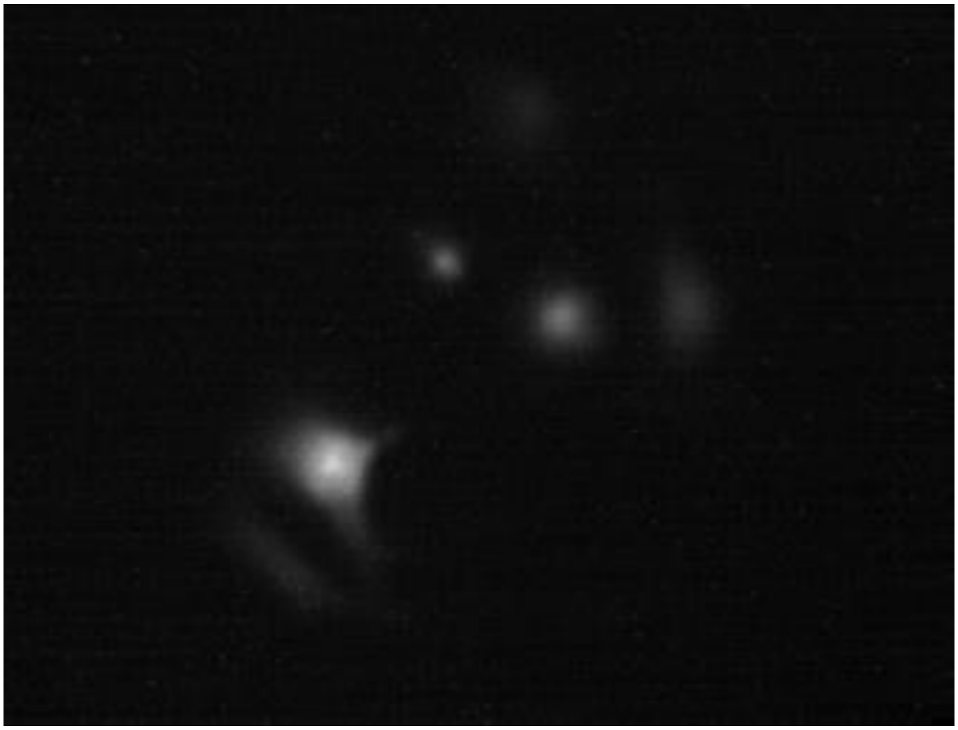


Fig. 4-4. GFP cell image taken by a custom-built camera with an integration time of 253ms.

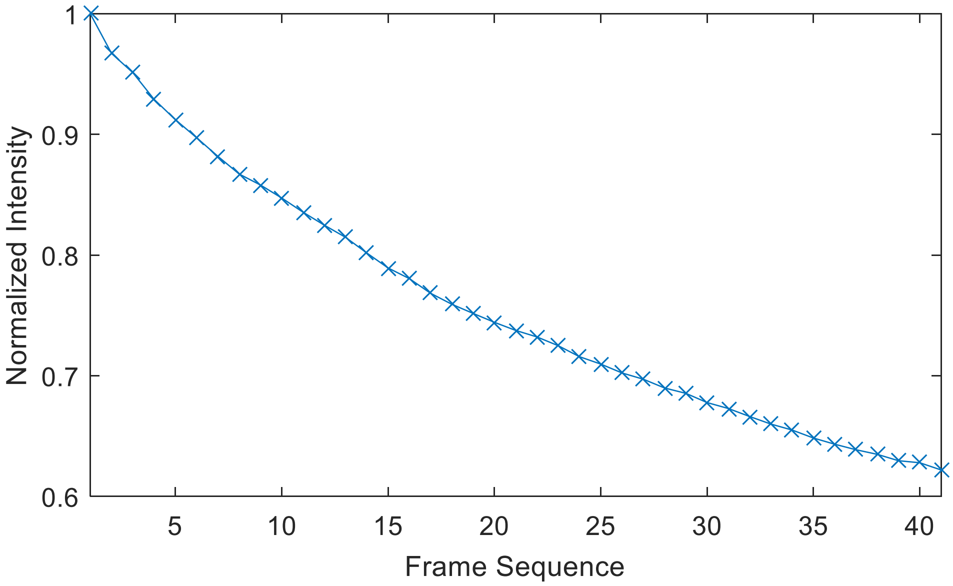


Fig. 4-5. Photo-bleaching of fluorescein. Signal intensity decreases as we continue to capture the same image multiple times.

To solve this problem for time-gated imaging, the CMM processing should be implemented in the analog domain before analog-to-digital conversion takes place. However, implementing the CMM in analog domain is very challenging. In the next section, we discuss three possible architectures for implementing CMM in analog domain. Finally, we will propose a new CMM architecture that provides a high accuracy without increasing hardware complexity.

4.3.1 Hardware Implementation of Center-of-Mass Method

In this section, we discuss the hardware implement of the equation (3), which is rewritten below again as a reminder:

$$\tau_{CT} = \frac{\int t \cdot f(t)dt}{\int f(t)dt}, \qquad \tau_{DT} = \Delta t \cdot \frac{\Sigma\, n \cdot f[n]}{\Sigma\, f[n]}.$$

First, the location of analog-to-digital converters (ADC) must be determined. Depending on the ADC location, the CMM calculation is implemented in digital domain, analog domain, or mixed-signal domain as shown in the Fig. 4-6 (a), (b), and (c), respectively.

Two parameters need to be analyzed before implementation. First, the sampling rate requirement of ADC should be minimized for high speed. If the number of samples is too large, the overall imaging speed will be limited by the ADC speed. Second, the use of analog circuits should be minimized for high accuracy. Analog computation usually suffers from PVT variations as well as systematic errors. We will evaluate each architecture in Fig. 4-6 and find the best architecture that can meet our goals.

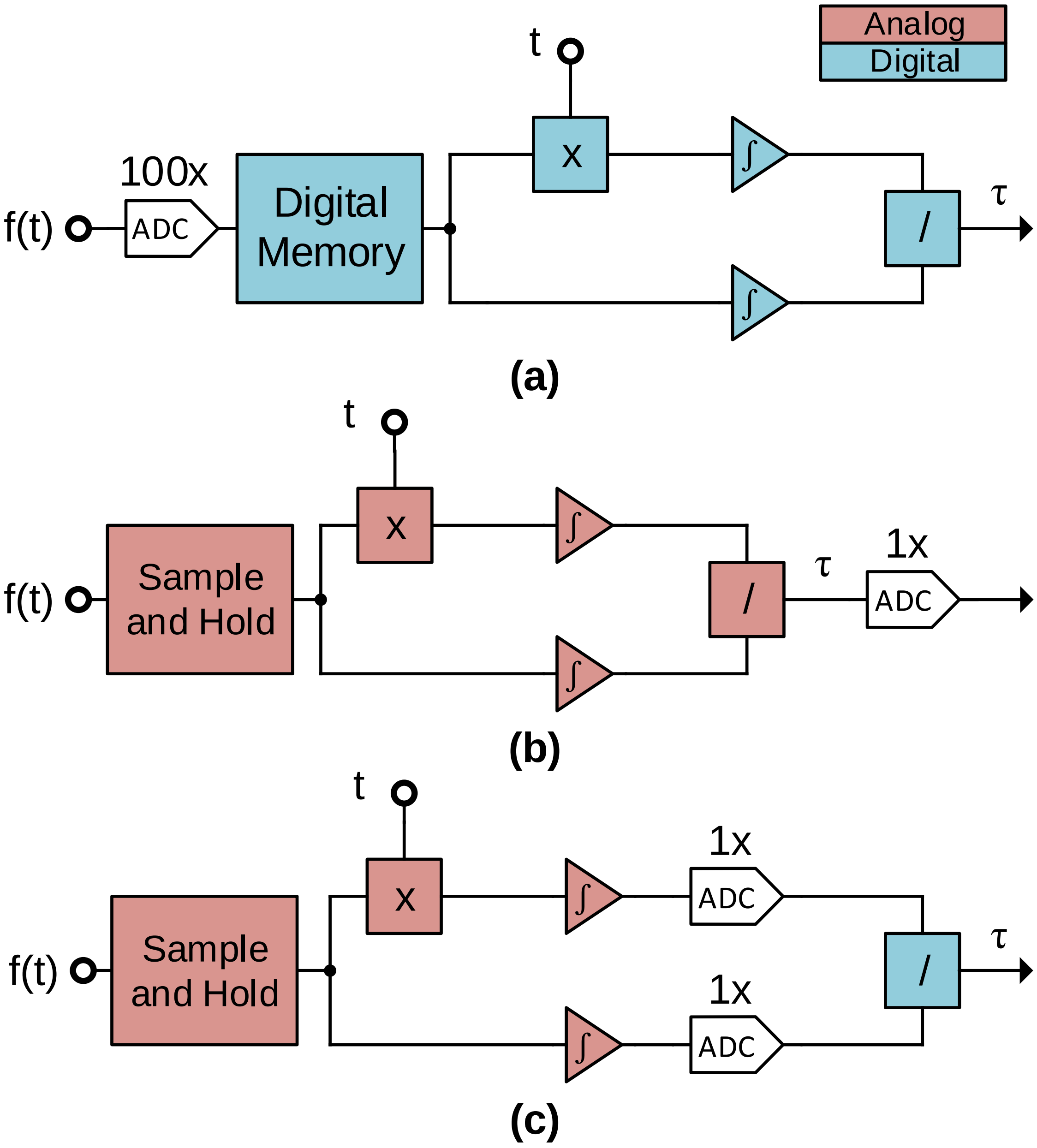


Fig. 4-6. Hardware implementation options of CMM equation; (a) pure digital implementation, (b) pure analog implmentation, and (c) mixed-mode implmenetation.

The direct implementation of the discrete-time CMM is shown in Fig. 4-6 (a). Since the building blocks are in digital domain, the implementation is robust and relatively easy to design. However, it requires the ADC with a sampling rate of ~100 times faster than the frame-rate to acquire all optical samples within one frame time [46]. This means that the fluorescnece lifetime imager in this architecture will be slower by two orders of magnitude than the conventional image senosrs.

A pure analog implementation is shown in Fig. 4-6 (b). The advantage of this approach are the relaxed ADC requirements. However, each pixel or column circuit requires analog computational blocks such as multipliers and dividers. A high precision multiplier is available in switched capacitor circuits but implementing an accurate divider in analog domain is very challenging. The analog divider will suffer from PVT variations as well as systematic errors.

Finally, we show the mixed-mode implementation of the CMM equation in Fig. 4-6 (c), a hybrid approach where the divider is implemented in digital domain. This implementation provides a good compromise between the two extreme implementations above. The requirement for ADC speed is still by orders of magnitude less demanding than the pure digital approach and the design can be very robust thanks to the digital implementation of the divider.

Note that all cases require a frame buffer to save 100 frames before the integration block, where the data rate is reduced by 100 times. Such a requirement for large memory space is not practical. To solve this problem, we propose a practical implementation of the mixed-mode approach that combines a multiplier and an integrator to eliminate the need for a frame buffer.

4.3.2 Proposed Pixel Hardware for Center-of-Mass Method

We propose a new pixel hardware that does not require a frame buffer and a multiplier. This implementation substitutes the multiplying operation by multiple integrations, combining the multiplication and the integration in one operation. Consequently, there is no need for switched capacitor multipliers or frame-buffers, but only two integration nodes will be sufficient, which already exist in the conventional two-tap pixels [34]–[42].

To explain the implementation, first, we express CMM equation in infinite series as given in the following equation.

$$\int_0^T tf(t)\,dt \approx \int_0^\Delta f(t)\cdot\Delta dt + \int_\Delta^{2\Delta} f(t)\cdot 2\Delta dt + \cdots \tag{4}$$

In this equation, each term has an integration width of Δ and the optical signal f(t) is multiplied by a corresponding weight, i.e. $n \cdot \Delta$ for n-th term. This calculation method is illustrated in Fig. 4-7 (a).

This multiplication can be achieved by multiple addition if we change the equation in another form as:

$$\frac{1}{\Delta}\int_0^T tf(t)\,dt \approx \int_0^T f(t)dt + \int_\Delta^T f(t)dt + \cdots + \int_{T-\Delta}^T f(t)dt. \tag{5}$$

In the equation, f(t) is added N times for the time $[0, \Delta]$, and 1 time for the time $[T - \Delta, T]$. This is illustrated in Fig. 4-7 (b). For easiness of implementation, we assume that T is sufficiently large and the signal content is negligible after the time T. The equation then becomes

$$\frac{1}{\Delta}\int_0^T tf(t)\,dt \approx \int_0^T f(t)dt + \int_\Delta^{\Delta+T} f(t)dt + \cdots + \int_{T-\Delta}^{2T-\Delta} f(t)dt. \tag{6}$$

Each term on the right-hand side represents a time-gated integration with a time-gate width of T. Therefore, any time-gated imagers that can integrate the optical signals for a defined period can perform the proposed CMM calculation, provided that the location of the integration widow can be shifted or delayed by Δ.

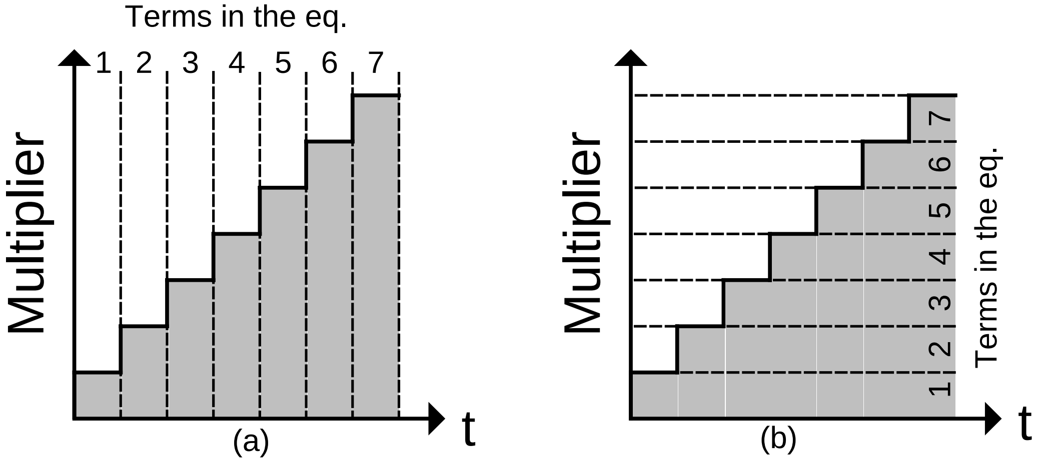


Fig. 4-7. The multiplying operation of (a) traditional CMM and (b) proposed CMM.

Fig. 4-8 shows the timing diagram of a proposed CMM calculation using a two-tap pixel. The time windows TX1 and TX2 are running at a period of 2W, whereas the excitation light has a period of 2W-Δ. This results in the shift of the time window by Δ relative to the excitation light.

The first turn-on period of the TX2 integrates the fluorescence from t=Δ to T+ Δ, representing the second term in the equation (6). In the next turn-on time, TX2 then integrates the fluorescence from t=2 Δ to T+2 Δ, representing the third term. This integration is continued until the equation (6) is completed. As a result, the floating diffusion for TX2 will contain the result of the CMM equation. The fluorescence signals outside of the TX2 window are integrated in the floating diffusion for TX1. This signal provides the intensity information and the lifetime can be finally calculated from both floating diffusion signals, X and Y as:

$$\tau_{\mathrm{CMM}} = \mathrm{T} \cdot \frac{\mathrm{X}}{\mathrm{X} + \mathrm{Y}} \tag{7}$$

where X and Y are the measured fluorescence using the time gates TX1 and TX2, respectively, and T is the total delay of the time-gates.

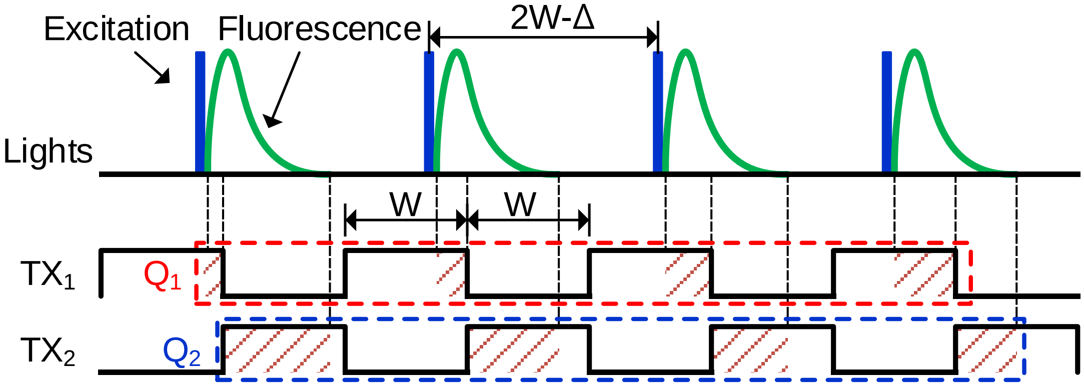


Fig. 4-8. Timing diagram for multiplying-integration in the proposed CMM pixel.

4.4 Simulation of Lifetime Estimation Methods

In this section, we show the simulation results comparing RLD and CMM implementations to prove the effectiveness of the proposed CMM method. To be realistic, we assume that IRF is a Gaussian pulse with a full width at half maximum (FWHM) of 3~5ns. This pulse width can be achieved if the pulse driver can operate at >100MHz. For simplicity, other non-idealities such as pulse skews are excluded in this simulation. The simulation model of IRF is then

$$\mathrm{irf}(t) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[\frac{-x^2}{2}\right], x = \frac{t}{\mathrm{FWHM} \cdot 2.355}. \tag{8}$$

Then, the fluorescence response is modeled as a first order LTI system with single pole. An impulse response function of the system can be defined as:

$$h(t) = u(t) \cdot \frac{1}{\tau} \exp[-t/\tau]. \tag{9}$$

The measured fluorescence by the sensor is then acquired by convoluting (8) and (9).

$$f(t) = \mathrm{irf}(t) * h(t) = \int_0^t \mathrm{irf}(\tau) \cdot h(t-\tau)\, d\tau. \tag{10}$$

Fig. 4-9 shows the simulated response curves for the excitation light with a pulse width of 5 ns (IRF), plotted along with the various fluorescence responses of different lifetimes: 1ns, 3ns and 5ns, respectively. As expected, the waveform does not look like a single exponential decay because of the wide width of the excitation light. We can observe that the peak locations of response curves are different; i.e. the peak position is delayed further as the fluorescence lifetime increases. Therefore, the estimated lifetime using RLD will not be consistent and will be sensitive to the location and the width of the time gates.

To simulate the RLD method, two time gates are defined. Since the IRF peak location is found at 7.5ns in this simulation model, we defined two time gates away from the peak location for best accuracy. The time gates used in this simulation are [10-12.5ns] and [12.5-15ns]. Each response is integrated using the defined time gates, and lifetime is estimated using equation (2). In this case, Δt of 2.5ns is used because it is the width of the time window.

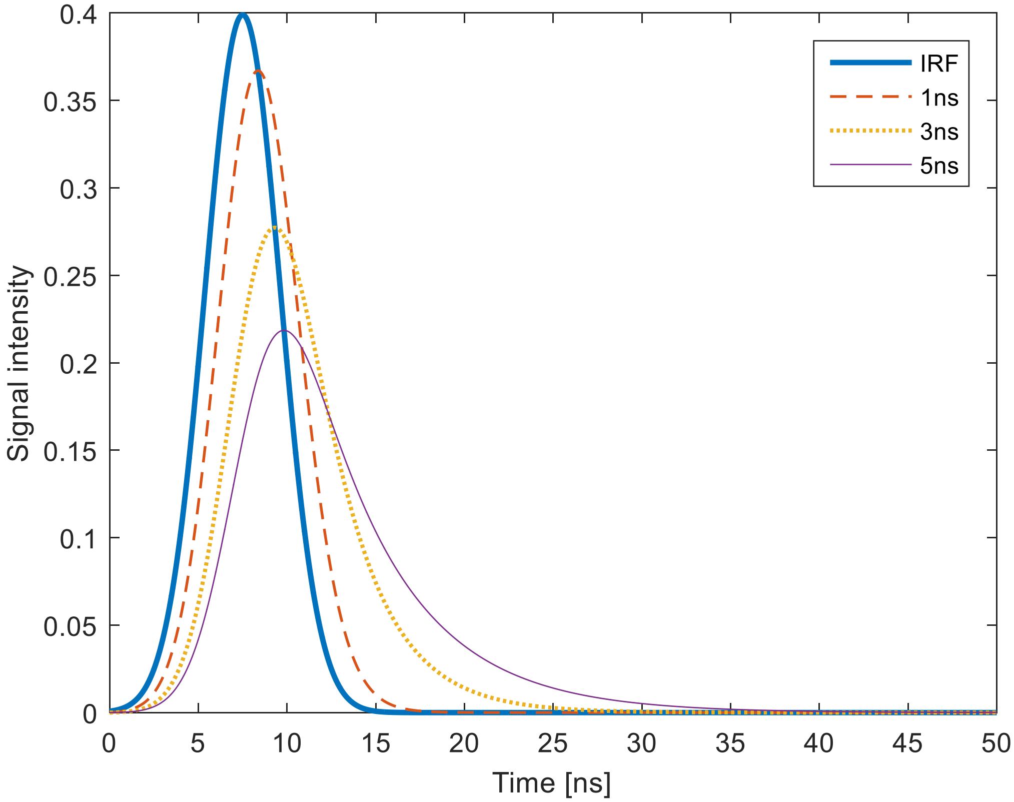


Fig. 4-9. Sensor system responses to an excitation light of a 5ns pulse-width and various fluorescence responses with different lifetimes of 1ns, 3ns and 5ns, respectively.

The proposed CMM method is simulated using the equation (6). At the beginning of the integration time, the center locations of the two time-gates are located at t=0. As shown in Fig. 4-8, the time-gates are shifted by 0.5ns for 80 steps throughout the integration time, ending at t=40ns.

The simulation results for both RLD and CMM are plotted in Fig. 4-10. The CMM measurement results are very consistent regardless of the IRF pulse width and target lifetime. On the other hand, the lifetime estimation using RLD fails at small target lifetimes as expected and as it is already reported in [52]. One additional observation is that for the IRF widths of 4ns and 5ns, the estimated lifetime value using RLD is completely off from the target values, as the peak location of the fluorescence is approaching the time gates. This limited capability of the RLD method is well known in literature [46]. The quality of the RLD method highly depends on the time-gate configuration, and hence the dynamic range of the RLD method is very narrow; i.e. it cannot measure the short lifetime and long lifetime at the same time because time-gate cannot be optimized for wide range of lifetimes.

To quantify the quality of both methods, the accuracy is measured and summarized in Table 4-1. As discussed above, the accuracy of the CMM is independent of the IRF's pulse width unlike the RLD method. In general, the CMM method always provides accurate results regardless of the shape of IRF as it effectively measures the full waveform and calculates the average delay of photons. The error increases for short target lifetimes due to the shift resolution of the time-gates. Therefore, the scanning resolution (0.5ns in this simulation) of the time-gates during the integration determines the dynamic range of lifetime measurement in the low-end. On the other hand, the error increases for long target lifetimes. This is due to the limited shift coverage of the time-gates, which is 40ns in this simulation. For large target lifetimes, the tail of the exponential decay will last longer than this range, resulting in errors. Therefore, the total shift amount of the time-gates determines the dynamic range in the high end.

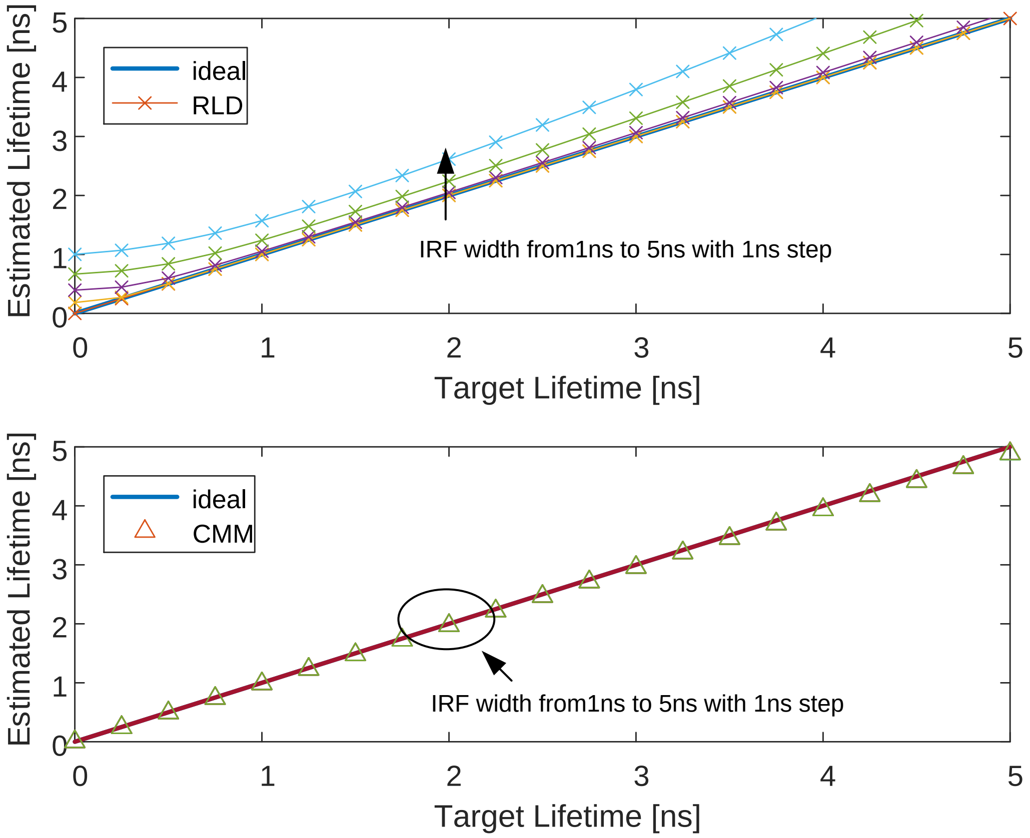
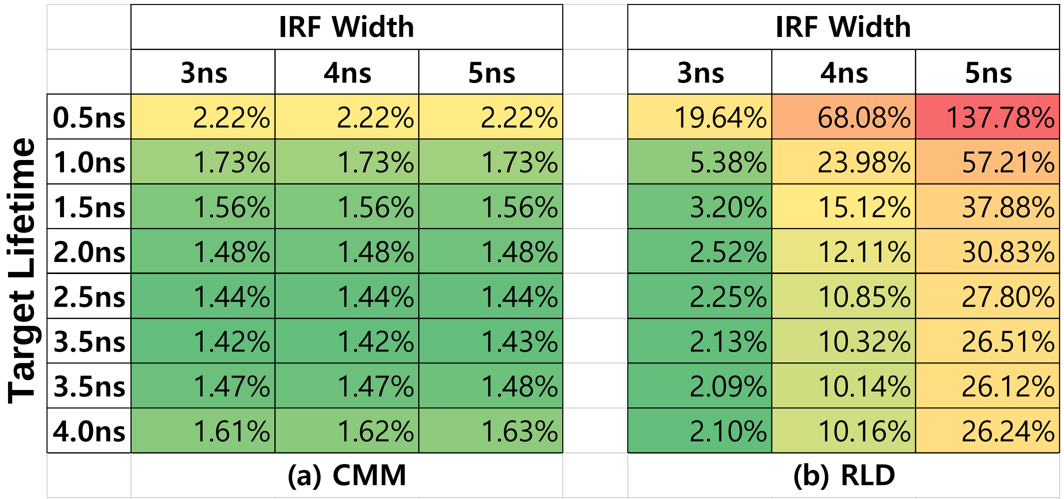


Fig. 4-10. Estimated lifetime using RLD and CMM with various IRF pulse-widths.

Table 4-1. Accuracy comparison of lifetime estimation using (a) CMM and (b) RLD.



	IRF Width			IRF Width		
Target Lifetime	**3ns**	**4ns**	**5ns**	**3ns**	**4ns**	**5ns**
0.5ns	2.22%	2.22%	2.22%	19.64%	68.08%	137.78%
1.0ns	1.73%	1.73%	1.73%	5.38%	23.98%	57.21%
1.5ns	1.56%	1.56%	1.56%	3.20%	15.12%	37.88%
2.0ns	1.48%	1.48%	1.48%	2.52%	12.11%	30.83%
2.5ns	1.44%	1.44%	1.44%	2.25%	10.85%	27.80%
3.5ns	1.42%	1.42%	1.43%	2.13%	10.32%	26.51%
3.5ns	1.47%	1.47%	1.48%	2.09%	10.14%	26.12%
4.0ns	1.61%	1.62%	1.63%	2.10%	10.16%	26.24%
	(a) CMM			**(b) RLD**		

4.5 Measurement Setup

Fig. 4-11 shows the block diagram of our custom FLIM system used for measurement. The top and left blocks in the diagram are electronics parts which are composed of a custom-built two-tap pixel sensor [38], [39], a delay IC on the board, a laser diode driver (ICHaus), and a laser diode (445nm). The image sensor board has an FPGA that synchronously controls the image sensor and the excitation light, respectively.

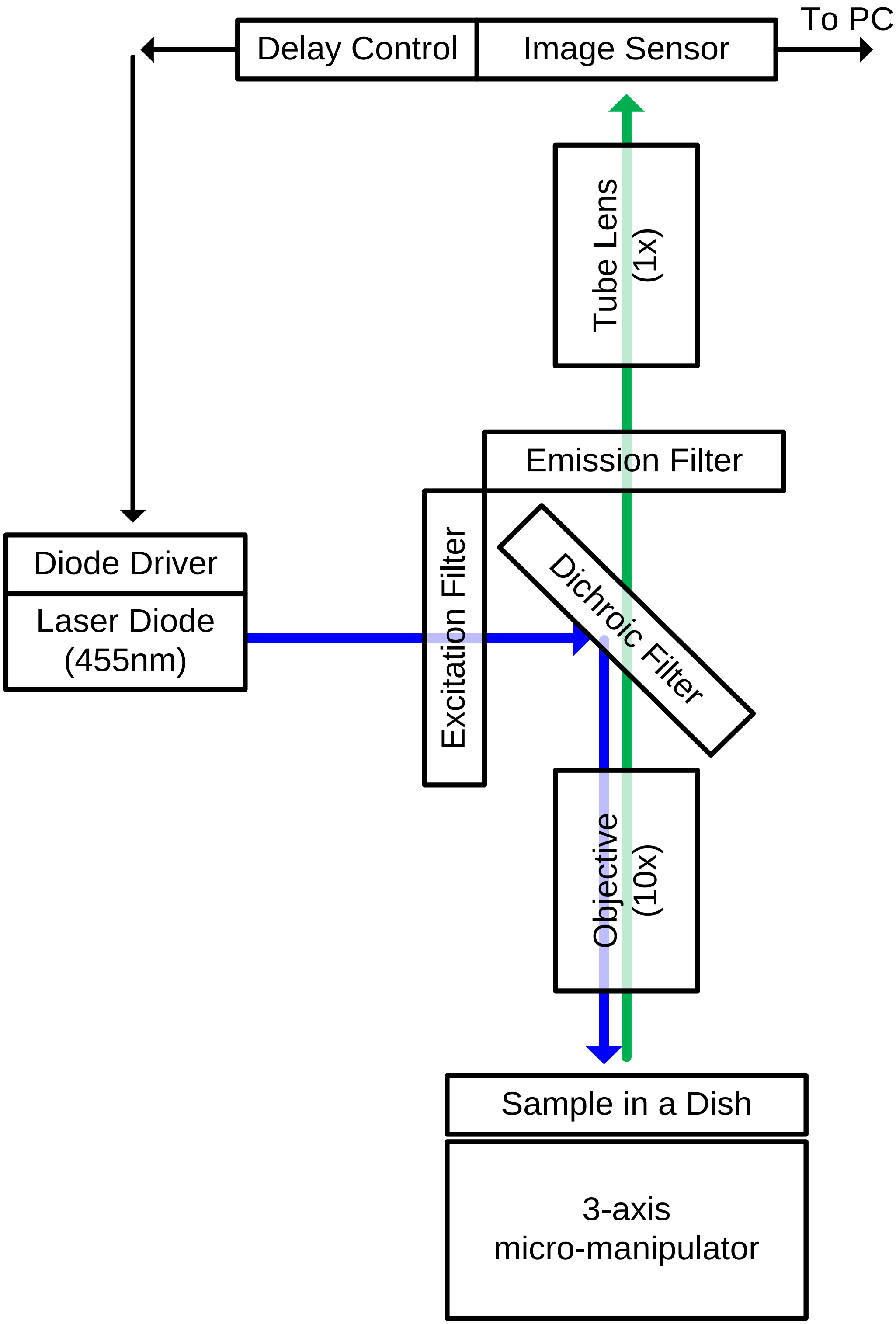


Fig. 4-11. Block diagram of FLIM measurement setup.

The excitation optical pulse from the laser diode passes through the excitation filter. This excitation filter blocks all the light except for the excitation wavelength. The filtered light is then reflected at the dichroic mirror and is illuminated on the sample through an objective lens. The fluorescence response from the sample, which is composed of longer wavelengths than the excitation light, goes through the same objective, the emission filter, and the tube lens, is eventually detected by the image sensor. The infinity-corrected objective lens and the tube lens are aligned to make a parallel optical path between them. In this way, the distance between the two lenses does not change the focus, making it much easier to put the filter cube between the lenses. Fluorescence is then finally detected by the image sensor. The sensor can be configured to perform a 2-tap pixel operation for image capture either in the RLD mode or in the full-sampling CMM mode. The sensor also supports the proposed CMM integration method. The delay is shifted during the integration time to enable a single-shot IRF independent measurement of fluorescence lifetime. The control board can provide delays from 0ns up to 94.72ns with an 0.37ns interval.

In addition, the system has two mechanical components. A 3-axis micromanipulator is used to move the sample in three different axes. Especially, the z-axis location is very important as the depth of field of the objective is very shallow. Although it is not shown in the figure, a mirror is placed between the laser and the filter cube to manipulate the direction of the laser.

Fig. 4-12 shows the picture of the custom-built FLIM system on an optical table. All measurement results in the next section is acquired using this setup.

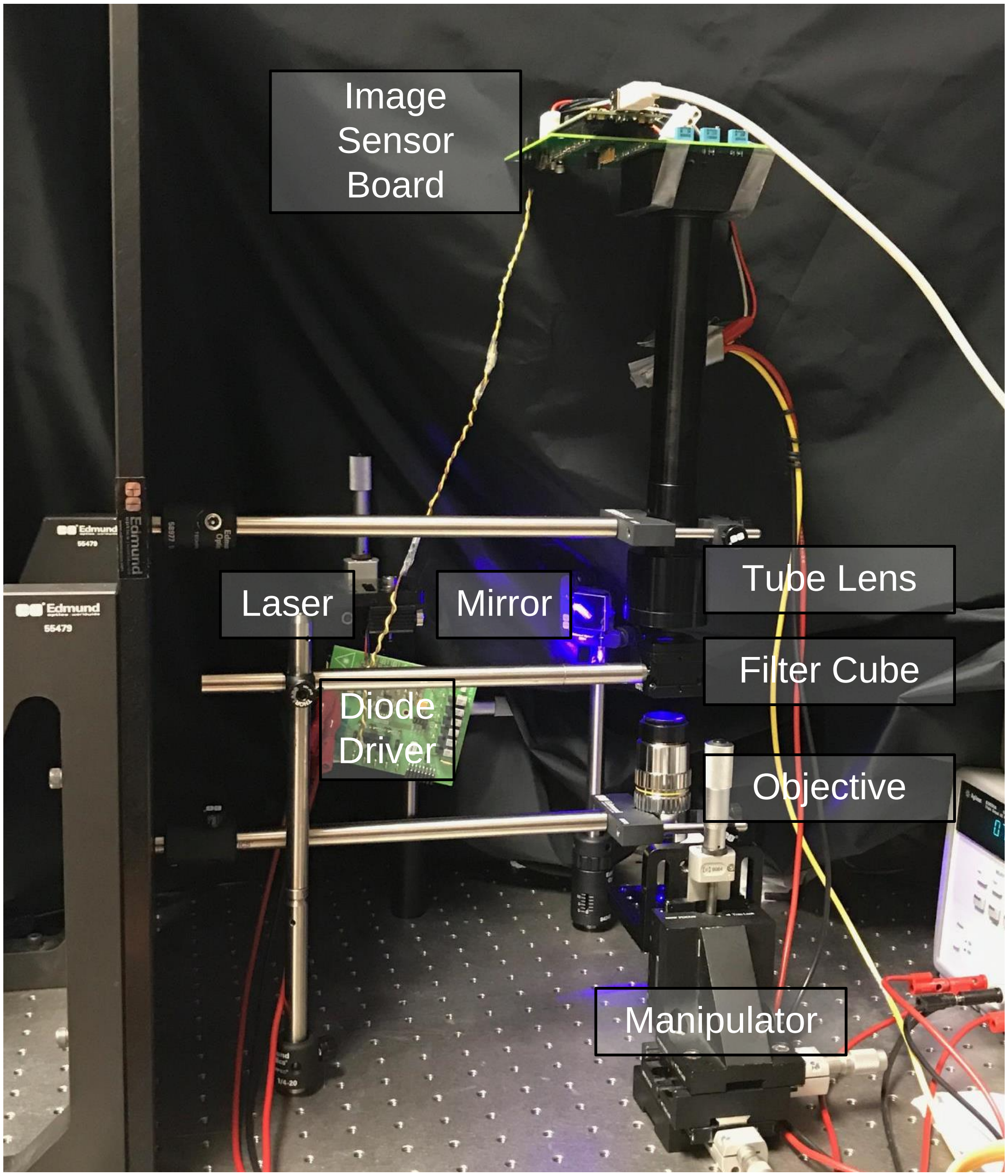


Fig. 4-12. FLIM measurement setup implemented on an optical table.

4.6 Experimental Results

To demonstrate the advantage of the proposed CMM scheme, we compare three different imaging methods using the same two-tap pixel sensor. The three methods are: two-sample RLD, multiple-sample CMM, and single-shot CMM. The two-sample RLD scheme takes only one frame image from the image sensor. The captured frame has two samples per pixel from the two complementary time gates. The multiple-sample CMM scheme takes 80 frames and calculates the CMM equation by post-processing in a computer. This provides the reference value of lifetime for comparison. Lastly, the single-shot CMM scheme (the proposed method) takes only single frame from the sensor.

The measurement was made for three optical signals, IRF, fluorescein, and GFP cells. IRF is measured as a baseline to calibrate the CMM measurement. The reason for measuring fluorescein is to evaluate and characterize the sensing system because its lifetime is well known as 4ns. Finally, GFP cells are measured to demonstrate the lifetime measurement of real samples. The GFP cells that we used are MDA-MB-231 breast cancer cells transfected by GFP.

4.6.1 Region of Interest

Fig. 4-13(a) shows a cell picture taken by the custom-built image sensor using a white external light source. GFP-transfected breast cancer cells are fixed on the dish for this measurement.

Since the image contains the cell region as well as the non-fluorescent region, a region of interest (ROI) must be defined to include only pixels with fluorescence. ROI is defined based on the fluorescence intensity image which is shown in Fig. 4-13(b). The intensity image is acquired by averaging the signals measured by each tap. Only pixels with intensity value larger than 12.5 are counted for characterization. Using this ROI, the lifetime value and standard deviation in time domain (noise) are calculated from the fluorescence lifetime image in Fig. 4-13(c).

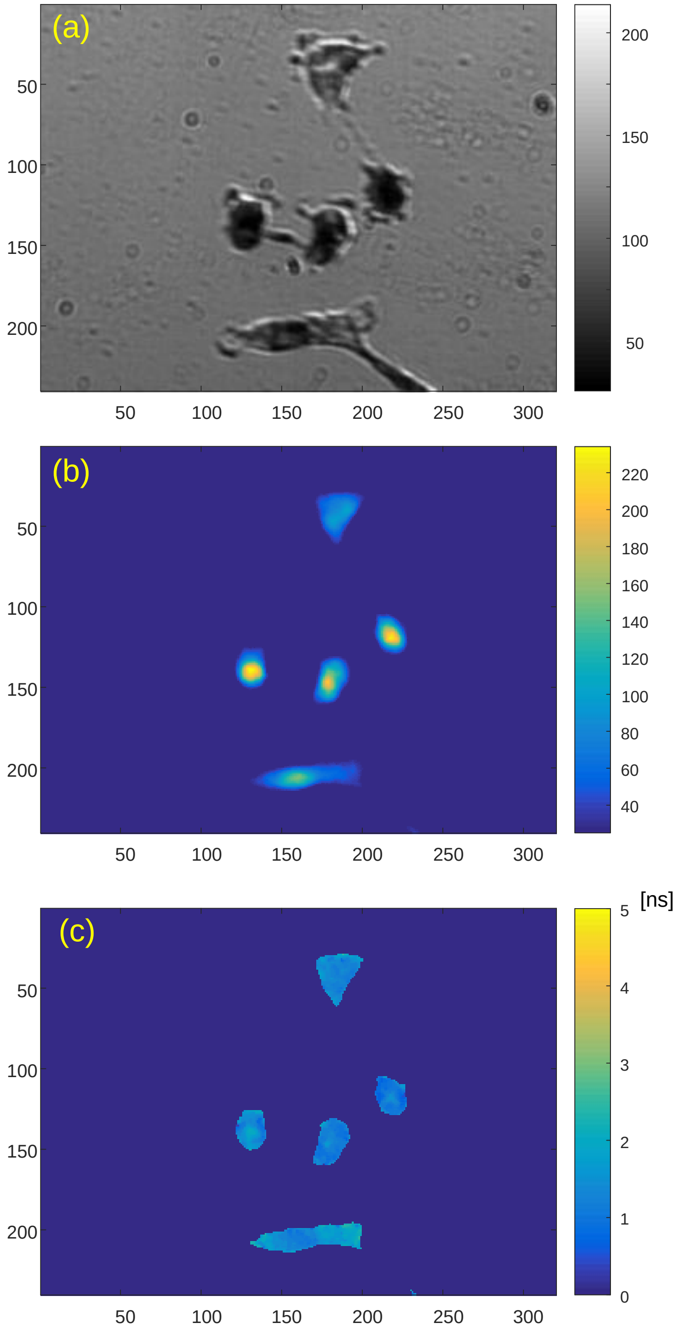


Fig. 4-13. GFP cell images (a) intensity (b) fluorescence (c) fluorescence lifetime.

4.6.2 Measurement Methods Comparison

Table 4-2 shows the lifetimes, random noise, and fixed-pattern noise measured using RLD, conventional CMM and proposed CMM. RLD and the proposed CMM has the same integration time for fair comparison. The conventional CMM measurement is made by acquiring 80 frames of the image with different time gates. To minimize the photo-bleaching during the measurement, we first captured the images using RLD and proposed CMM. The conventional CMM measurement is done after them. Note that the fluorescence signal gets too weak after one conventional CMM measurement so we could not get enough samples for random noise measurement.

The RLD reports an incorrect higher value for GFP lifetime, as it is expected in the simulation in Fig. 4-10. This error is due to the long tail of IRF which cannot be compensated in RLD method. The proposed method showed the best performance in fixed-pattern noise (FPN) and random noise measurement. The noise performance of conventional CMM should be the best in theory but the photo-bleaching during the measurement yielded very poor SNR.

We also estimated the best achievable resolution at the maximum fluorescence intensity. The measured noise depends on the intensity level as discussed in 3.4; the higher the signal the smaller the noise. The proposed CMM method has an estimated resolution of 123ps which is ~4 times better than the RLD method.

The measurement results showed that the proposed CMM method could provide high accuracy regardless of the width of IRF. The imaging speed of the proposed CMM is as fast as the two-tap RLD method, which is known to be the fastest fluorescence lifetime imaging. The fast imaging speed helped offer high SNR by preventing the samples from the photo-bleaching.

Table 4-2. Measured lifetime using RLD, full sample CMM, and proposed single shot CMM. Values in parenthesis are standard deviation.

unit [ns]	RLD		Conv. CMM		Proposed	
	Fluorescein	GFP	Fluorescein	GFP	Fluorescein	GFP
Fluorescein	4.08	**2.32**	4.139	1.589	4.03	1.41
Fixed Pattern Noise	**0.4**	**0.957**	**0.779**	**0.845**	**0.256**	**0.512**
Measured Noise	**0.849**		N/A		**0.22**	
Est. Resolution*	**0.475**		N/A		**0.123**	
Integration time	**250ms**		**20s**		**250ms**	

4.6.3 Fluorescence Lifetime Imager Comparison

Table 4-3 compares the proposed sensors with other silicon-based fluorescence lifetime image sensors. Only the proposed sensor achieved fast imaging speed (1>Hz), high resolution (>QVGA) and good accuracy (<500ps). In [56], the sensor achieved 2% of error when measuring 6.9ns lifetime, but the error increased to 67% for a longer lifetime of 53ns. This is expected in all RLD based sensor. The sensor reported in [62] has the longest integration time because it is based on TCSPC. It also has the smallest pixel array size due to a large complexity of the pixel and the readout circuits. Both RLD and conventional CMM measurement are performed in [52]. It also reported the accuracy improvement when using CMM algorithm, but it didn't report the integration time for CMM measurement. Assuming that 64 intensity frames are used for CMM calculation, we expect at least 5 seconds of integration time.

Table 4-3. Comparison of the proposed CMM with other silicon-based fluorescence lifetime image sensors.

	2013 TED [62]	**2013 Sensors** [56]	**2015 ISSCC** [52]		**Proposed CMM**
Method	**TCSPC**	**RLD**	**RLD**	**CMM**	**Single Shot CMM**
Measured Noise/Error	n.a.	2% @ 6.9ns **67% @ 53ns**	n.a.	n.a.	**0.22ns**
Time Resolution	n.a.	n.a.	n.a.	n.a.	**0.123ns**
Integration Time	**4s**	3ms / 2ms	83ms	**>5s (*est.)**	**250ms**
Pixel array	**32x32**	**256x256**	**256x512**	**256x512**	**336x256**
Pixel size [um]	25	6.3	11.2x5.6	11.2x5.6	5.9

4.7 Conclusions

A single-shot CMM method has been proposed to provide fluorescence lifetime imaging for high frame-rate, high accuracy, and robustness under photo-bleaching. The proposed CMM method can be readily applied to any two-tap pixels and does not increase the sensor complexity. The only required component is a delay controller which can be integrated in the timing generator or assembled using an off-the-shelf component. The simulation results showed that the proposed CMM method provides better accuracy than the RLD method at the same imaging speed. A microscopy system has been built to test this new measurement scheme. The experimental results verified the advantages of the proposed CMM method over the standard RLD and conventional multiple-sample CMM schemes. It has achieved a faster imaging speed than the full-sample CMM method by 80 times with smaller standard deviation in lifetime due to less photo-bleaching effect from single-shot measurement.

Chapter 5

Time-of-Flight Depth Camera with Adaptive Background Light Suppression

5.1 Introduction

Depth or three-dimensional (3-D) imaging is very critical in many computer vision systems. For example, the depth camera in micro-vehicles allows detecting fast-approaching obstacles for collision avoidance. It can also be used for mapping the surroundings or terrains for navigation and surveillance purposes. From the video sequence of the depth images, motion or gesture can be recognized for human-computer interaction (HCI) in video games, virtual and augmented realities [63], [64].

There are three challenges in realizing 3-D cameras for the applications mentioned above. First, the 3-D camera should be implemented in a small form factor. Depth cameras are usually larger than the conventional 2-D cameras because of the increased complexity. They usually require an illuminator, additional optics and sometimes mechanical scanner. Second, the digital post-processing should be minimized for video-rate outputs. Normally, mobile processors are not powerful enough for complex image processing in real-time. Finally, the camera should reliably operate in various lighting conditions. Most 3-D cameras reported up to date cannot operate in outdoor conditions because the strong sunlight disturbs the measurement [34], [37], [41], [42], [65]–[69].

In the rest of this section, the first two challenges will be addressed by evaluating two popular methods for depth imaging, triangulation and time-of-flight (TOF). Then, the problems in outdoor 3-D imaging will be explained followed by the proposed solution for BGL suppression in mobile 3-D cameras.

5.1.1 Triangulation

Though triangulation is one of the most popular for depth imaging method [70], it has three big problems to be used in mobile applications. First, the stereo vision or the passive triangulation suffers from intensive computation and geometric constraints. It has to process two different images from two cameras or a pair of pixel arrays in a camera [71]. From the images, depth information is extracted by finding the correspondence between them. This process requires a lot of computation and sometimes fails when it cannot find the geometrical correspondence.

This geometric constraints can be relieved by active triangulation [65]–[67]. In this method, one of two cameras is replaced with a patterned illumination. Thanks to the known pattern, the camera can easily find the correspondence in the emitted and the captured pattern. However, active triangulation has higher cost and form factor due to the use of mechanical scanners.

Finally, the dynamic range of the triangulation method is limited by the sizes of the focal plane and the pixel because the depth is translated into the lateral shifts in the focal plane. For example, a fine depth resolution of 1.5mm can be acquired at the cost of limited range intervals of 1m~1.1m [65]. To increase the dynamic range, the resolution should be dropped accordingly.

5.1.2 Principle of Time-of-Flight Depth Measurement

TOF offers the least computational overhead, no geometric constraints and a wide dynamic range with a good accuracy of centimeter ranges. Because of these good properties, the TOF has been preferred in real-time depth imaging and actively investigated in the past few years [34], [37], [39], [41], [42], [68], [69], [72]–[82].

Fig. 5-1(a) illustrates the basic operation of TOF measurement and its variant implementation schemes. The TOF system consists of a light source (LED or laser) and an image. The measurement begins with the emission of a short light pulse toward the subject. The light pulse is then reflected by the subject and then returns to the sensor. As soon as the sensor detects the returned light, it measures the time between the emission and the detection. The time is called the time-of-flight (TOF) from which, the depth (D) is given by

$$\mathrm{D} = (\mathrm{c} \cdot \mathrm{TOF})/2 \tag{11}$$

where c is the speed of the light.

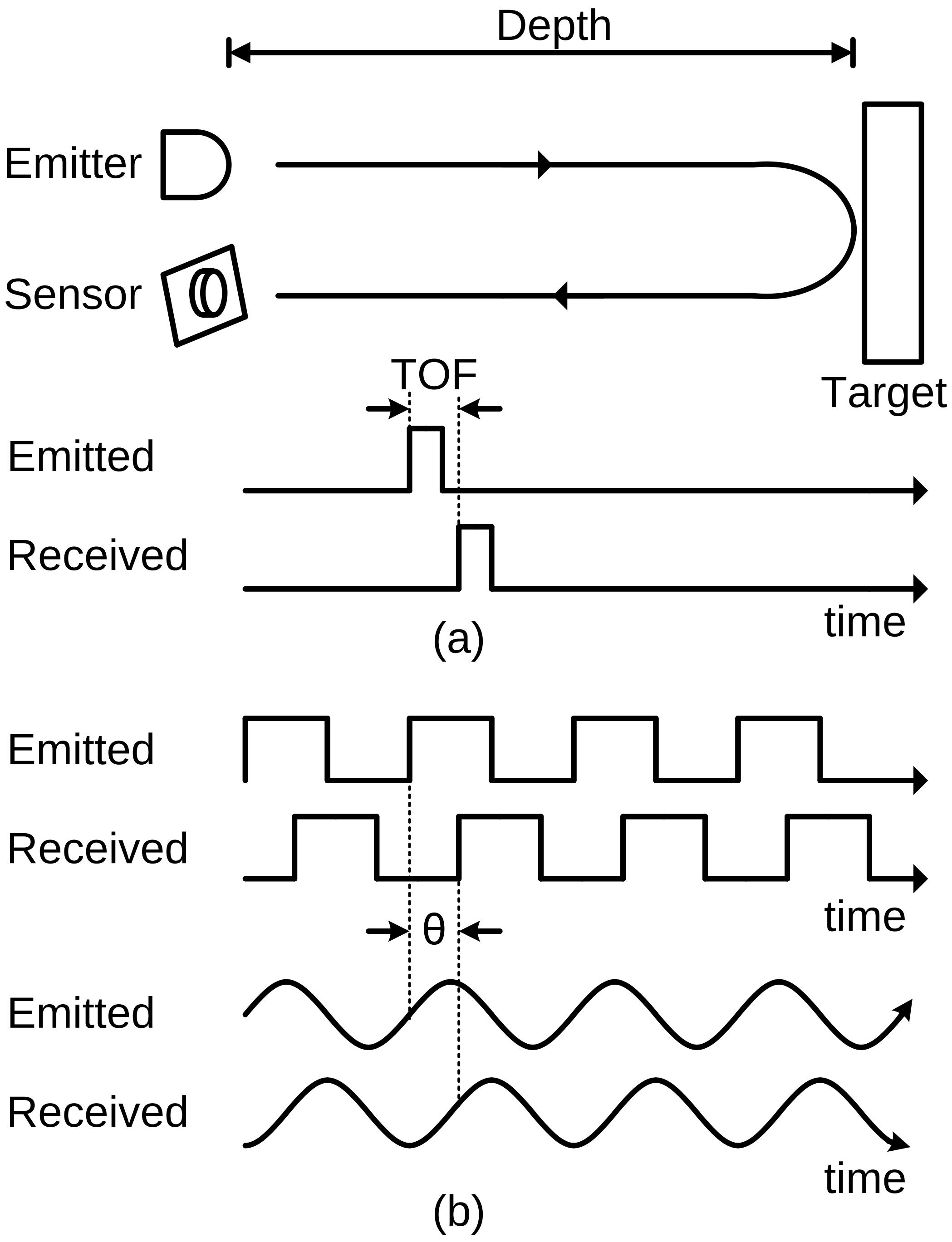


Fig. 5-1. Schematic diagram of principle of TOF sensor operation: (a) direct and (b) indirect TOF measurement schemes.

This method is called direct TOF and usually requires an extremely sensitive pixel since the light pulse can be significantly attenuated during the travel. Single-photon avalanche diode (SPAD) has been reported to implement the system [68], [69], [82].

However, the direct TOF has two major problems due to the employment of SPAD. First, the SPAD is vulnerable to the extraneous light. Due to the nature of single photon detection, SPAD is not able to receive the signal once triggered by the BGL. Second, the size of the pixel is not easy to scale with the technology. It is reported that the technology scaling degrades the performance of the SPAD [ref required].

Alternatively, the TOF can be indirectly measured using conventional less-sensitive pixels, if the synchronous measurements can be repeated to enhance the signal-to-noise ratio (SNR) as shown in Fig. 5-1(b). In this indirect measurement, the phase shift of the periodic light signal is measured using the demodulation pixels or lock-in pixels [34], [41], [42]. From the measured phase shift (θ), the TOF is given by

$$\mathrm{TOF} = \frac{\theta \cdot \mathrm{T}}{2\pi} = \frac{\theta}{2\pi \mathrm{f}} \tag{12}$$

where T and f are the period and frequency of the light, respectively.

5.1.3 Indirect TOF Measurement under BGL Illumination

In this subsection, the indirect TOF is further investigated under the BGL illumination. Fig. 5-2 shows the conceptual diagram of light pulses in TOF measurements with the presence of background light. We assume that optical square wave is emitted from the infrared (IR) light source. The square wave, $\mathrm{p}(\mathrm{t},\theta)$ with the phase delay (θ) and the period (T) is given as

$$\mathrm{p}(\mathrm{t},\theta) = \mathrm{rect}\left(\frac{\mathrm{t}}{\mathrm{T}/2}\right) * \sum_{\mathrm{n}} \delta\left(\mathrm{t} - \mathrm{nT} - \frac{\mathrm{T}}{2\pi} \cdot \theta\right). \tag{13}$$

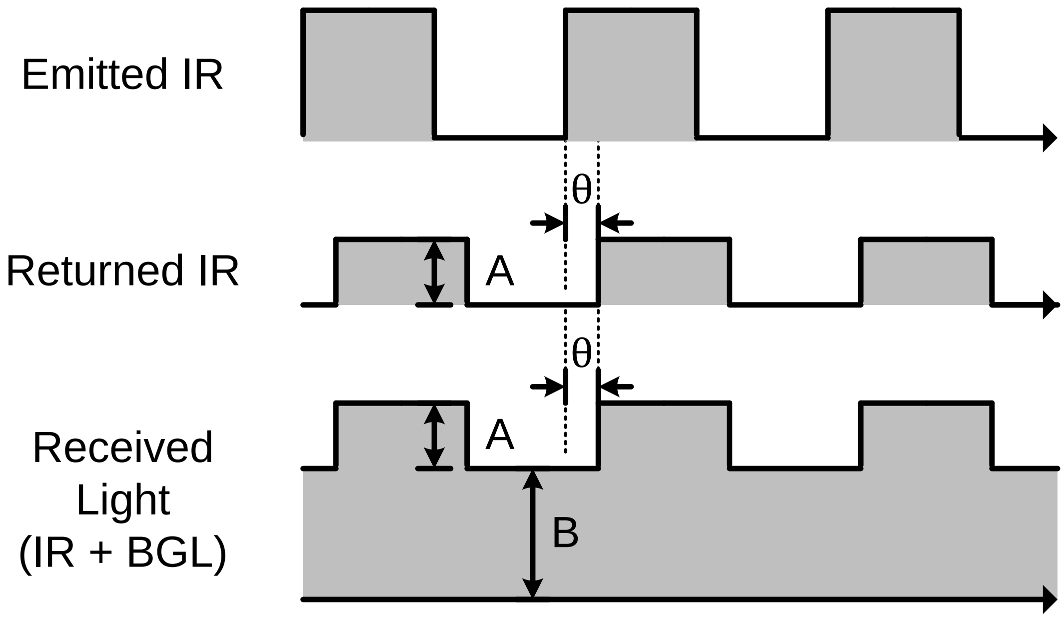


Fig. 5-2. Emitted and received light (infrared) waveforms in TOF measurements when exposed to background light illumination.

As a reference signal with no phase delay, the emitted IR is defined as $p(t, 0)$. After the roundtrip, the measured IR will be attenuated and delayed from the emitted IR. Arriving at the photodiode, IR signal generates the photocurrent by $A \cdot p(t, \theta)$ with a conversion factor of A and the delay of θ. Considering the additional photocurrent contributed by background light, the total photo current is given as

$$I_{PH}(t) = B + A \cdot p(t, \theta) \tag{14}$$

where B represents the BGL component which is assumed constant over the measurement period.

If B is zero, a simple circuit can be used to find θ as shown in Fig. 5-3 [41], [42]. The circuit consists of a pinned-photodiode (PPD) and two transfer gates (TX). The photocurrent, I_{PH}, can be steered to one of the floating diffusion (FD) capacitors, C1 or

C2, by selecting the TX signals. The photocurrent is then integrated into charges in the FD capacitors. The integrated charge with $\mathrm{p}(\mathrm{t},\phi)$ signal is given by

$$\mathrm{Q}(\phi) = \int \mathrm{A} \cdot \mathrm{p}(\mathrm{t},\theta) \cdot \mathrm{p}(\mathrm{t},\phi)\,\mathrm{dt}. \tag{15}$$

Complimentary gating signals, $\mathrm{p}(\mathrm{t},0)$ and $\mathrm{p}(\mathrm{t},\pi)$, are typically selected for TX1 and TX2, resulting in the integrated charges of $\mathrm{Q}(0)$ and $\mathrm{Q}(\pi)$. These two quantities can be processed to find the phase shift as

$$\theta = \pi \cdot \frac{\mathrm{Q}(\pi)}{\mathrm{Q}(0) + \mathrm{Q}(\pi)}\,. \tag{16}$$

The existence of BGL will add a constant to both measured charges resulting in an error during the phase estimation. To eliminate the BGL component, the difference of the charges is defined as

$$\Delta\mathrm{Q}(\phi) = \mathrm{Q}(\phi) - \mathrm{Q}(\phi + \pi). \tag{17}$$

The differential quantity contains two unknown variables, A and θ. Subsequently, the demodulation process should be performed again in order to solve for two variables. As can be seen in the Fig. 5-4, phase delay, θ can be uniquely determined from the two differential charges, $\Delta\mathrm{Q}(0)$ and $\Delta\mathrm{Q}(\pi/2)$ as

$$\theta = \begin{cases} \frac{\pi}{2} \cdot \left(1 + \frac{-\Delta\mathrm{Q}(0)}{|\Delta\mathrm{Q}(0)| + \left|\Delta\mathrm{Q}\left(\frac{\pi}{2}\right)\right|}\right), & \text{if } \Delta\mathrm{Q}\left(\frac{\pi}{2}\right) \geq 0 \\ \frac{\pi}{2} \cdot \left(3 + \frac{+\Delta\mathrm{Q}(0)}{|\Delta\mathrm{Q}(0)| + \left|\Delta\mathrm{Q}\left(\frac{\pi}{2}\right)\right|}\right), & \text{if } \Delta\mathrm{Q}\left(\frac{\pi}{2}\right) < 0 \end{cases} \tag{18}$$

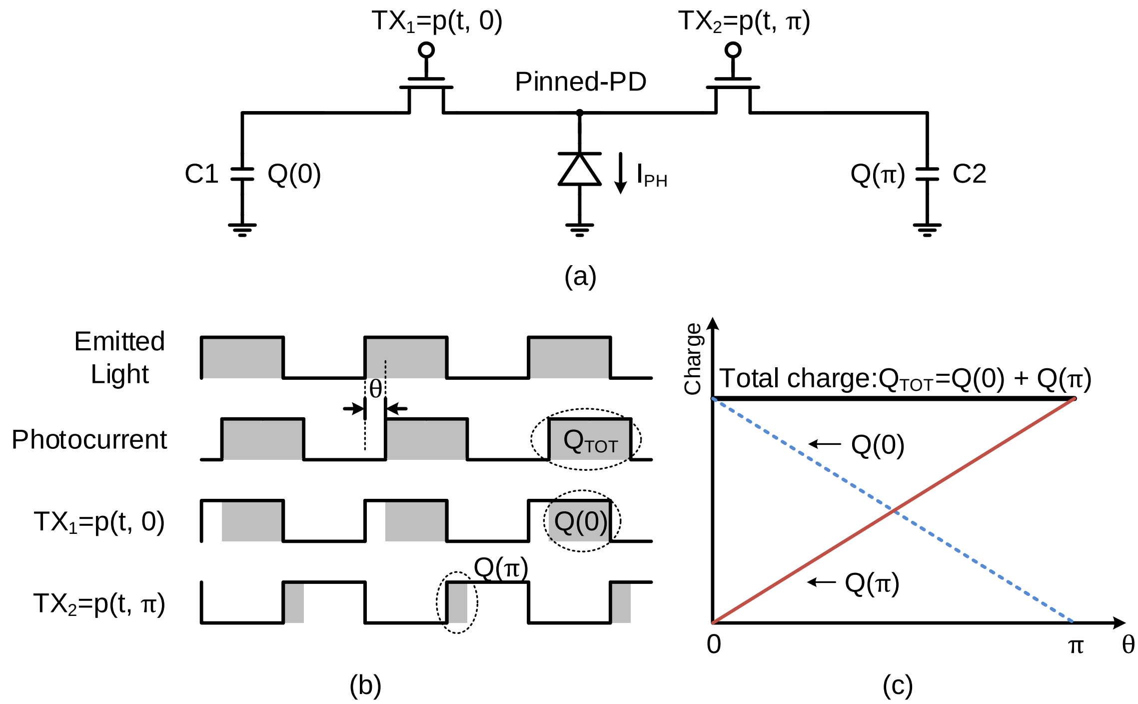


Fig. 5-3. Demodulation of the received light to measure phase delay when there is no background illumination: (a) demodulation pixel using a pinned-photodiode and two transfer gates, (b) waveforms of the emitted light, photocurrent and TX signals, and (c) integrated charges in the demodulation pixel as a function of phase delay.

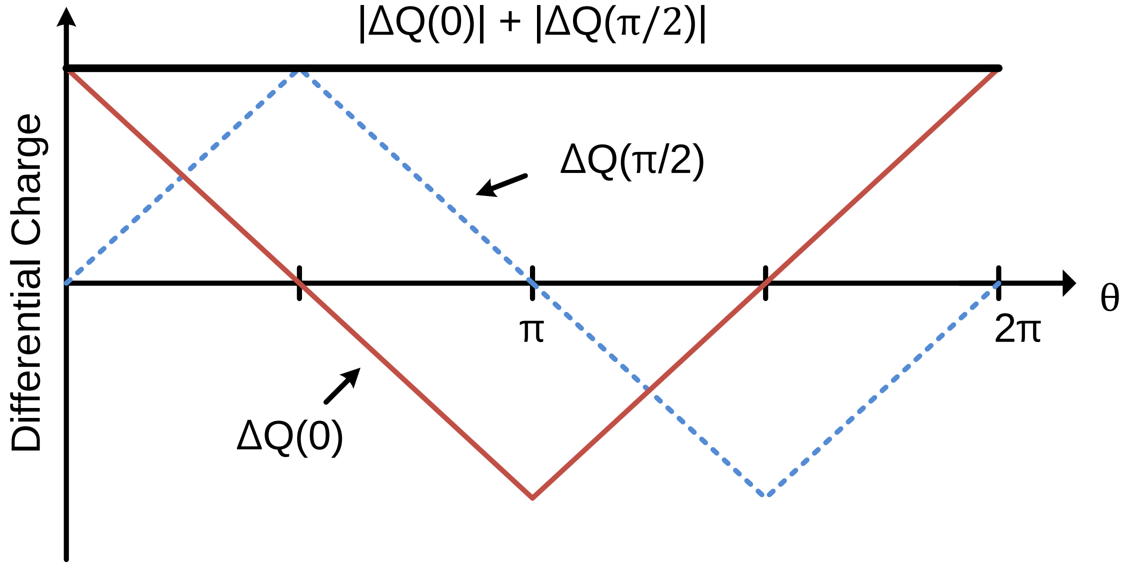


Fig. 5-4. Determination of phase delay from two differential charges, $\Delta(0)$ and $\Delta Q(\pi/2)$, under background illumination

It is worth to note that (16) and (18) assumes that the waveform of the photocurrent is a perfect square wave. The actual waveform will be band-limited due to the finite bandwidth of the emitter and the photodiode. This results in the nonlinear response of the sensor especially when θ is close to 0 or π [41].

5.1.4 Background Light Problem

On the other hand, the suppression of the background light (BGL) such as sunlight is very challenging in mobile applications. Some 3-D cameras for automotive applications are able to suppress the BGL using special schemes in pixel circuits [81], [82]. These cameras, however, suffer from a low spatial resolution due to the large pixel size, making it hard to be embedded in mobile devices. Sometimes they require a mechanical scanning device which significantly increases the size of the system.

Since the measurement involves an active lighting, it can be significantly interfered by the background light (BGL). An optical bandpass filter can suppress some BGL components, but further suppression is typically necessary to prevent the saturation under strong BGL.

In a SPAD-based TOF sensor [82], the signal photons could be differentiated from those from BGL by utilizing spatiotemporal coherence of the signal photons. This approach requires multiple SPADs to find the correlation of photons in each pixel. As a result, it had a large pixel size of 150×50 μm^2 and a line sensor with only 32 pixels was implemented. Although a spatial resolution of 340×96 was demonstrated for automotive applications, this approach would not be adequate for mobile applications due to the mechanical scanner.

In non-SPAD based sensors, various BGL suppression schemes have been investigated [72]–[78], [81]. These schemes can be categorized in four types: (1) minimum-charge transfer (MCT) [72], [73], (2) current-mode subtraction [81], (3) charge-domain subtraction [74]–[77], and (4) voltage-mode subtraction [78], [79].

The MCT employs a large intermediate storage to store both BGL and signal charges. Then, the signal charges are transferred to the smaller sensing node with the help of column-level monitoring circuit. A high conversion gain could be achieved by small sensing node while a large BGL can be tolerated by the large intermediate storage. However, the MCT scheme cannot screen the BGL charges which exceed the well-capacity of the intermediate node, and the operation is slow due to the limited bandwidth of the monitoring circuit.

The current-mode scheme [81] employs a current source in each pixel to compensate for the BGL-induced photocurrent. The measurement results show that the scheme only works with large signal currents (> 30pA), and hence a large pixel size (> $100\mu m$) is inevitable.

The charge-domain scheme uses active circuits [74], anti-parallel capacitors [75], flipping the integration capacitors [76], or injecting hole packets [77] for common-mode charge elimination. All these approaches require in-pixel capacitors which significantly increase the pixel size. Though a smaller MOSCAP can be used [77], the use of PMOS devices result in a large pixel size due to the well-spacing rule.

Finally, the voltage-mode scheme suppresses the BGL using correlated-double-sampling (CDS) [78] or a unity-gain buffer [79] in the pixel. These additional circuits increase the complexity, power consumption and the pixel size.

In the next section, I propose an Adaptable Background light Cancellation (ABC) scheme for real-time outdoor 3-D imaging [38], [39]. The scheme provides the adaptable spatiotemporal resolution in accordance with the BGL level. This allows maintaining the maximum resolution under the given BGL constraints, i.e., full resolution without BGL, and reduced resolution under high BGL. For the maximal flexibility, the BGL suppression circuits are implemented in column-level as suggested in [83].

5.2 Proposed Design

In this dissertation, I propose a column-level BGL suppression scheme [39] which can suppress the BGL over 100klx without significant increase in pixel size. The scheme is mostly implemented in column-level so that the circuits can be shared to minimize the area overhead. The main idea of the scheme is the dynamic adjustment of resolution to attain a higher BGL suppression. The resolution control is done by pixel-binning and super-resolution techniques, providing multiple spatial and temporal resolutions on demand. This, in turn, enables an adaptable imaging for the optimal performance in various conditions by compromising the resolution, frame-rate, and BGL suppression. To the best of our knowledge, this is the first demonstration of the 3-D camera that achieves the BGL suppression over 100klx without using in-pixel compensation circuitry. Though only 336×256 pixels are implemented in the prototype sensor, the array size can be easily extended thanks to the small pixel size of 5.9µm.

5.2.1 System Architecture

Fig. 5-5 shows the system architecture of the prototype sensor. It consists of the sensor chip and control electronics. The sensor chip has an array of 336×256 pixels. The pixel array can be reconfigured for various spatiotemporal resolutions for an adaptive imaging. The BGL suppression circuit block is implemented in the column-level. It is composed of 336 integrators and an analog frame memory with 84×64 cells. The integrator performs double-delta sampling (DDS), amplification, and integration of differential charges. The differential mode is integrated in the analog frame memory, while the common mode caused by the BGL is rejected. The following column-parallel ADCs digitize the signals stored in the analog memory. The ADC outputs are latched in the line memory and sequentially transferred to the PC via FPGA for the depth calculation and displaying.

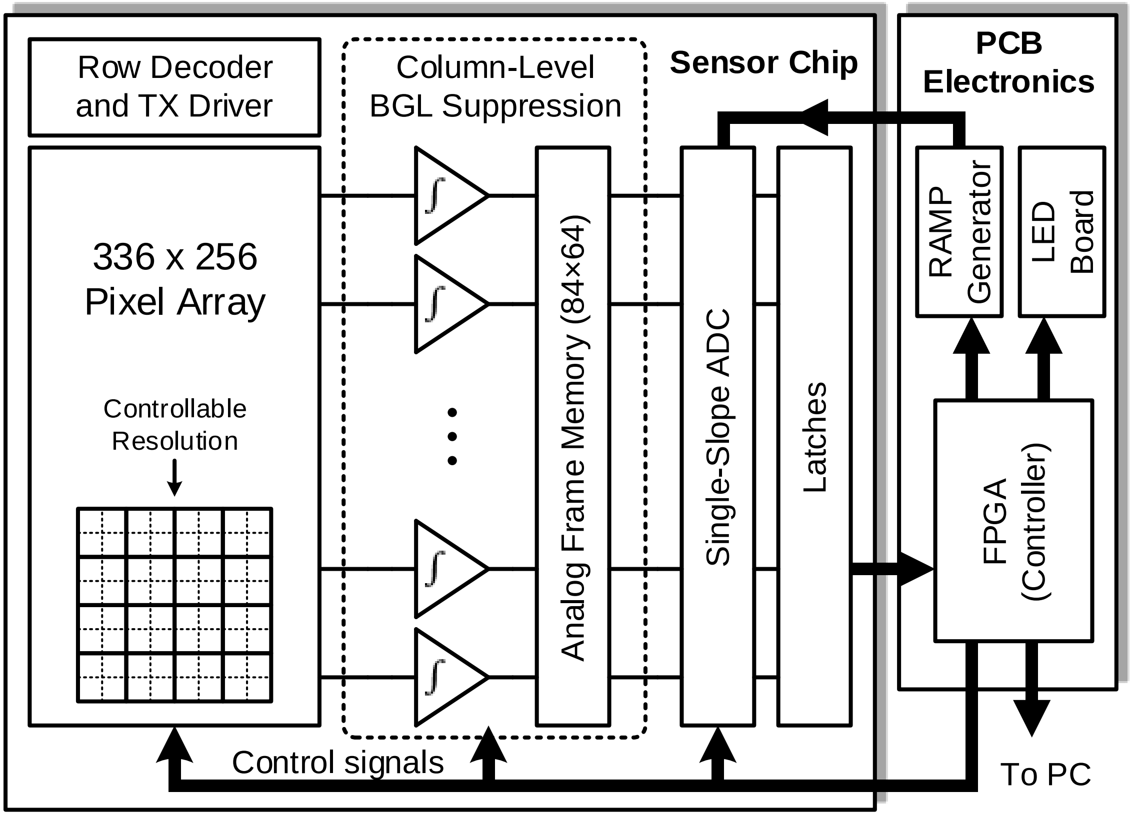


Fig. 5-5. System architecture of the prototype sensor.

5.2.2 Operation Principle

The biggest problem of the BGL is the sensor saturation. The strong BGL will generate a huge photocurrent which can saturate the sensor in a very short amount of time. Short integration time can prevent the saturation but with degraded SNR due to the small number of photon collected. Therefore, repeated measurement of the multiple sampling should be applied to enhance the SNR.

In the proposed scheme shown in Fig. 5-6, the integration time is divided into several short sub-integration times. After each sub-integration, the difference integrator accumulates the differential output, V_{SIG}, into the frame memory. Then, the pixel is reset for the next sub-integration. As a result, the saturation can be prevented by periodic reset and the SNR can be increased from the repeated measurement.

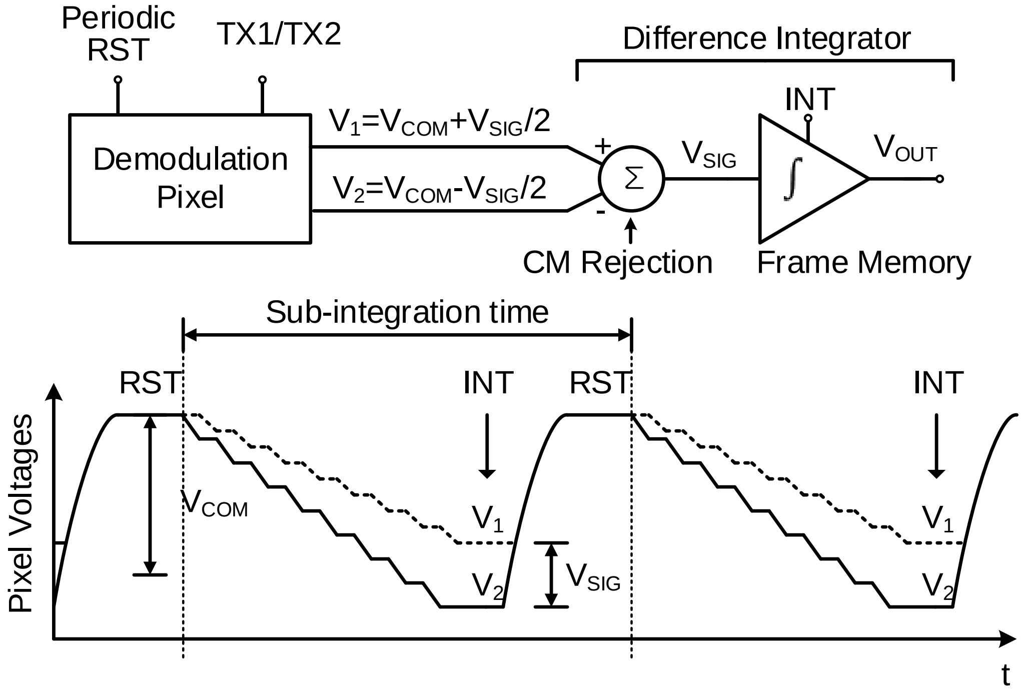


Fig. 5-6. Proposed BGL suppression scheme.

The proposed scheme requires a frame memory which can be implemented in either analog or digital signal domain. We chose the on-chip analog frame memory instead of the digital memory because the digital approach cannot be applied to high spatiotemporal resolution sensors for two reasons. First, the required ADC speed and power consumption should be increased by the number of sub-integration cycles. For example, when we implement 16 sub-integrations per frame as in our prototype sensor, the digital approach requires 16 times faster ADCs. Second, quantization noise can degrade the signal-to-noise ratio (SNR) when the pixel signal from each sub-integration is relatively small. On the other hand, the noise contribution from our approach (i.e. the analog integration) is negligible since the operation comes free from the DDS without additional circuits. The detailed circuit implementation will be described in the next section.

It is worth to review the BGL suppression principles in the literature [74]–[78] to compare with this approach. In some approaches, BGL portion is compensated either periodically [74]–[76] or conditionally [77], leaving the signal charges untouched. This eliminates the need for the second integrator, but is not adequate for a large array size. For example, the passive compensation requires the use of a big integration capacitor and a large photodiode (64~125 µm in pixel size) [75], [76] to minimize the charge injection and sharing. The active compensation demands more power and pixel size (109 µm) [74]. An open-loop compensation with PMOS devices is used in [77], but the use of PMOS increases the pixel size (14 µm). The scheme used in [78] also requires in-pixel capacitors and active circuits, resulting in a big pixel size (29.1 µm). In contrast, the proposed BGL suppression scheme is implemented in columns; thus, the second integrator can be shared, making the pixel circuit small.

5.2.3 Adaptable BGL Suppression with Resolution Control

3-D cameras can be exposed to various environments when mounted in mobile devices. In extreme cases, the BGL can be higher than 100klx in bright sunny days. This will saturate the pixel in less than a few milliseconds, and the BGL suppression scheme should be activated.

Unfortunately, there are only a few BGL suppression schemes that can overcome the BGL over 100klx and their pixel structures are complicated and huge (109~125 µm) [74], [75]. Due to this limitation, they cannot provide high resolution imaging even in indoors where no BGL suppression is required. Therefore, it is desirable to implement an adjustable BGL suppression scheme to adapt the sensor operation to environmental changes. Similar concept has been applied for a multi-resolution imaging by optimizing spatial and temporal resolutions [84].

In the proposed scheme, the spatiotemporal resolution can be adaptively controlled to boost the BGL suppression for different environmental conditions. The flexibility comes from the column-level implementation where the array is scanned row by row, as in rolling shutter operation. The sub-integration time or the BGL suppression interval can be adjusted by pixel binning in vertical direction. For example, when every four pixels are merged vertically, the effective number of rows decreases by four times. As a result, the sub-integration time becomes shorter, increasing the maximum suppressible level of BGL by four times.

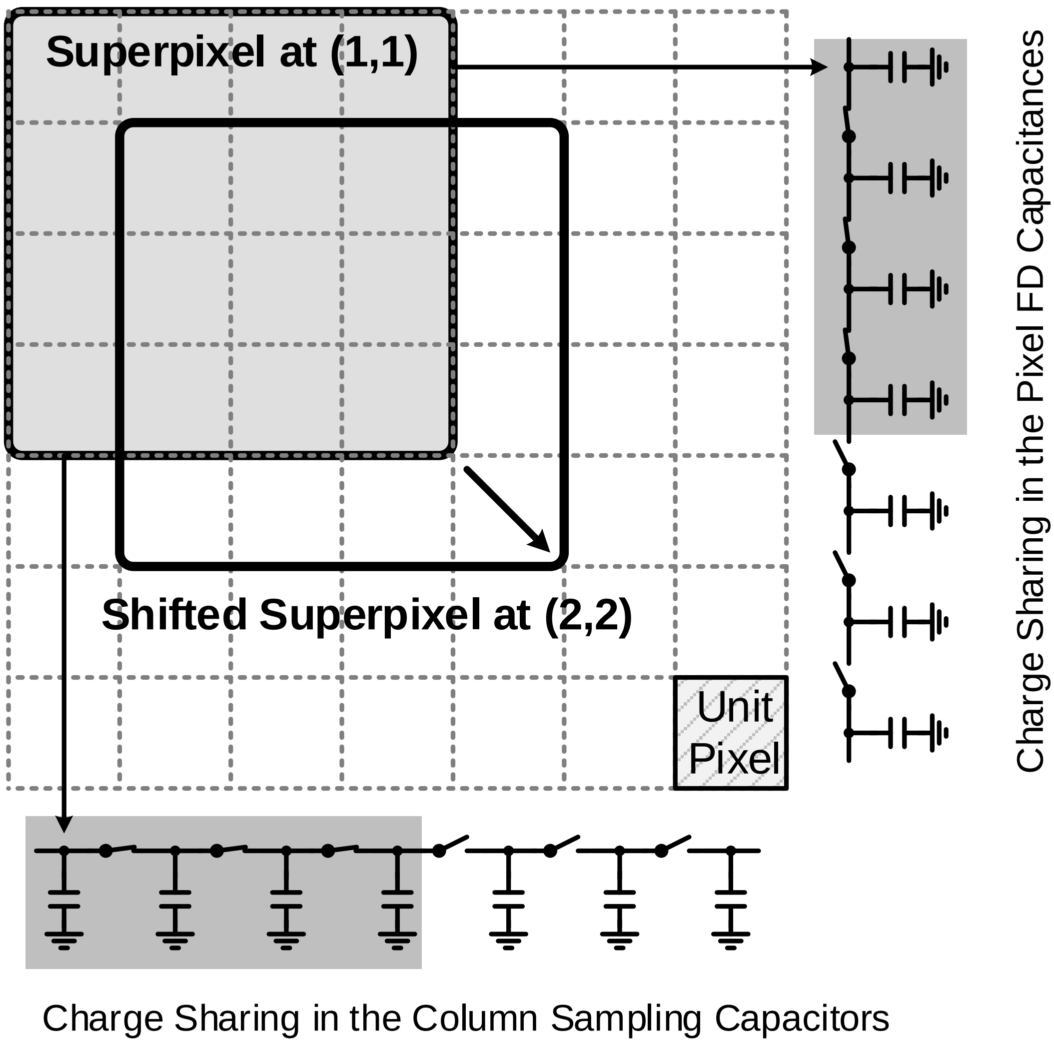


Fig. 5-7. Pixel binning and sub-pixel-shift for adaptive resolution control.

This resolution controllability provides trade-off between the resolution and the BGL suppression for adaptable operation as depicted in Fig. 5-8. The top-right corner indicates a reference point with the maximum spatial (x-axis) and temporal (y-axis) resolution when no pixel-binning is applied. As explained previously, the pixel-binning can boost the BGL suppression at the cost of lower spatial resolution. This BGL suppression boosting at the reduced resolution is depicted in Fig. 5-8(a).

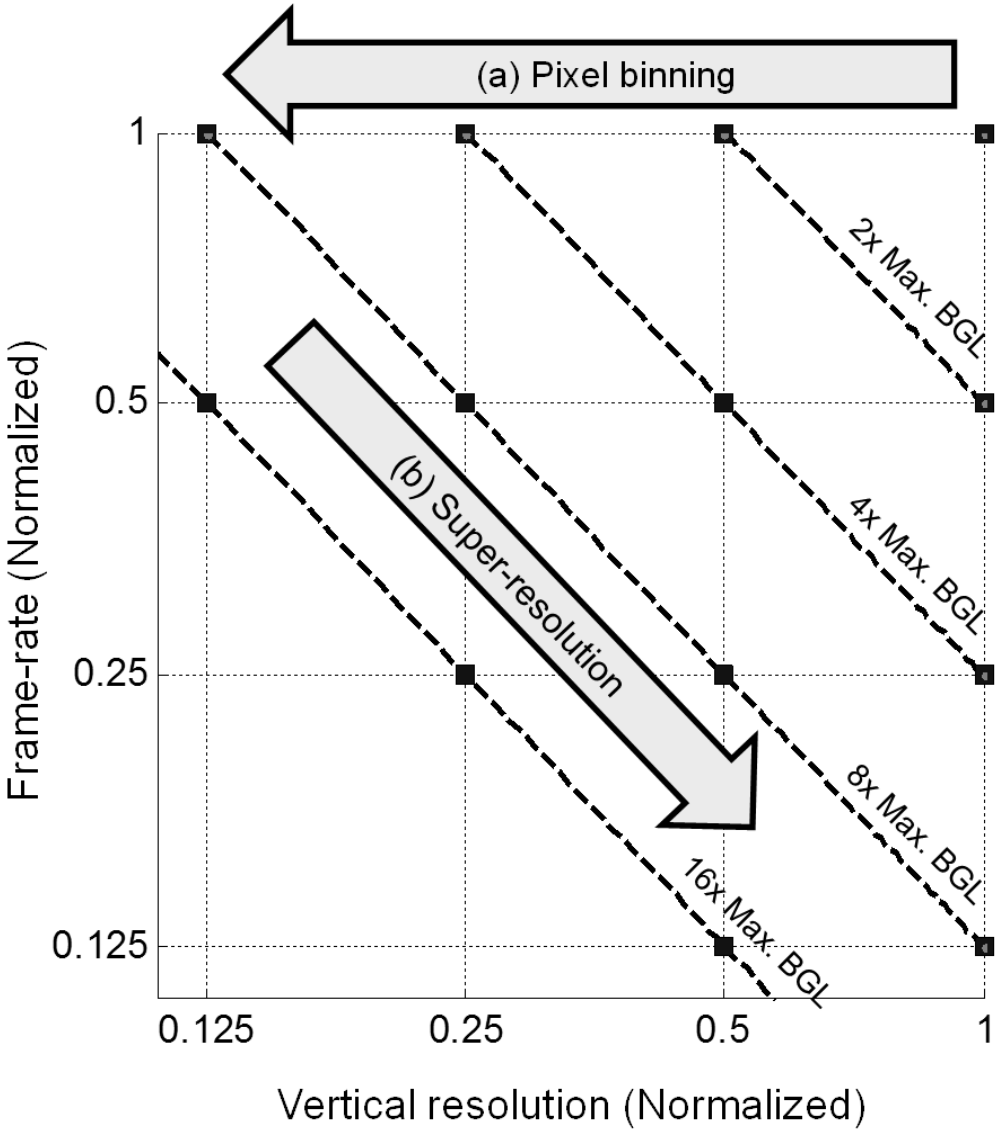


Fig. 5-8. Spatiotemporal resolution control by (a) pixel binning and (b) super-resolution for adaptive imaging.

The reduced spatial resolution can be compensated by a technique called super-resolution, which uses multiple low resolution images with sub-pixel displacements to reconstruct a higher resolution image [85]. This technique has also been applied to time-of-flight depth imaging and showed the edge improvement [86]. However, capturing image sequences with sub-pixel-shift usually requires mechanical scanners which increases the system complexity and size [87], [88]. To avoid this problem, we employed an electrical pixel-shift scheme as shown in Fig. 5-7. The figure shows how the superpixel (merged pixel) can be shifted from the coordinate (1, 1) to (2, 2). The minimum shift length is a unit-pixel-pitch; therefore, the original resolution of 336×256 can be fully recovered by this technique at the expense of a lower frame-rate as depicted in Fig. 5-8(b).

The contour plot of the maximum suppressible BGL is shown as dotted lines in Fig. 5-8. Each line indicates the same spatiotemporal resolution, or the frame-rate-resolution product. This clearly shows the trade-off between the spatiotemporal resolution and BGL suppression performance. The spatiotemporal resolution reduction has another advantage in terms of shot noise. Since the effective well-capacity per pixel increases by pixel-binning or multiple captures, the SNR degradation due to shot noise from the strong BGL can be compensated. Consequently, the proposed scheme can adaptively adjust its operation for various illumination conditions, providing optimal performance.

5.3 Circuit Implementation

5.3.1 Pixel Operation and Circuits

Fig. 5-9 shows the schematic, layout and timing diagram of pixel circuits. The pixel consists of a demodulation circuit and readout transistors. The demodulation circuit is composed of a PPD and two TGs (TX1 and TX2) [42]. The TGs steer the photocurrent into one of the two FD nodes according to the TG signals. The photocurrent is integrated in the FD nodes as charges, developing a voltage signal on the capacitance. This FD node voltage is then read by a source follower, SF. The current sink load of the source follower is placed outside of the pixel array and is shared by the pixels in the same column. The row select (RS) transistors are placed between the source follower and the column line to access one row at a time. The FD nodes can be reset through reset (RST) transistors.

To implement the vertical pixel-binning operation, three transistors are added to the conventional structure [42]: two for charge sharing, and one for x-y addressing. The x-y addressing transistor is equivalent to an AND gate and the pixel can be programmed to define a superpixel by CBIN (column binning) and RBIN (row binning) signals, enabling the multi-resolution feature as in our previous work [84]. From the simple pixel architecture as shown in Fig. 5-9(a), we could achieve a pixel-pitch of 5.9μm with a fill factor of 24% (Fig. 5-9(b)).

Fig. 5-9(c) shows the timing diagram of pixel operation. In this particular example, the LED array is synchronized with TX1 signal, and four pixels are vertically binned.

First, the binning configuration is programmed after the TX1 and TX2 signals are applied to the TGs to demodulate the photocurrent. In this example, the gates of the binning switches in four rows are programmed by writing the CBIN values through RBIN. The charges on the FD nodes in the neighboring pixels are shared when the binning switches are turned on. Then, the row select (RS) transistor is turned on to read both signal and reset levels for DDS operation. Right after the demodulation, the pixel output voltage, VPIX is the signal level. VPIX becomes the reset level after FD nodes are reset by RST signals. After that, binning configuration can be conditionally erased if a new configuration is needed in the next integration period.

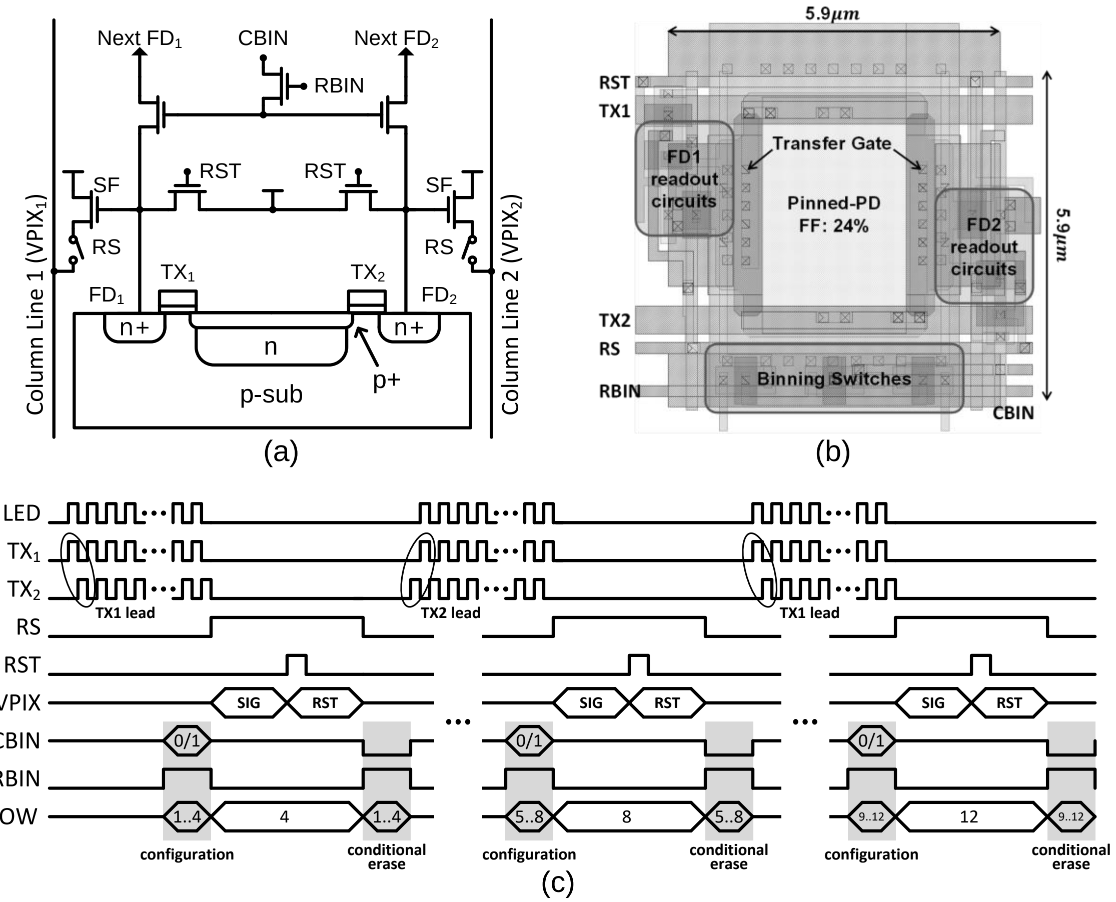


Fig. 5-9. (a) Unit pixel schematic, (b) pixel layout, and (c) timing diagram of pixel operation.

For the signal integrity, TX signals are turned off during the readout in order to prevent possible interference through parasitic couplings. This, however, causes uneven distribution of BGL signals to FD nodes. In other words, the BGL will contribute to the differential signal and cannot be cancelled by the proposed circuits. In order to solve this problem, the leading edge of TX signals is alternated as in Fig. 5-9(c). The BGL produced in the blank period is alternately integrated in each FD node in turn. The rest rows operate in the same manner.

5.3.2 Column-Level BGL Suppression Circuits

Fig. 5-10 shows the BGL suppression circuit and its timing diagram. The circuit consists of four sampling capacitors, an analog multiplexer and a switched-capacitor integrator. The sampling capacitor is to store reset and signal levels from the two pixel outputs.

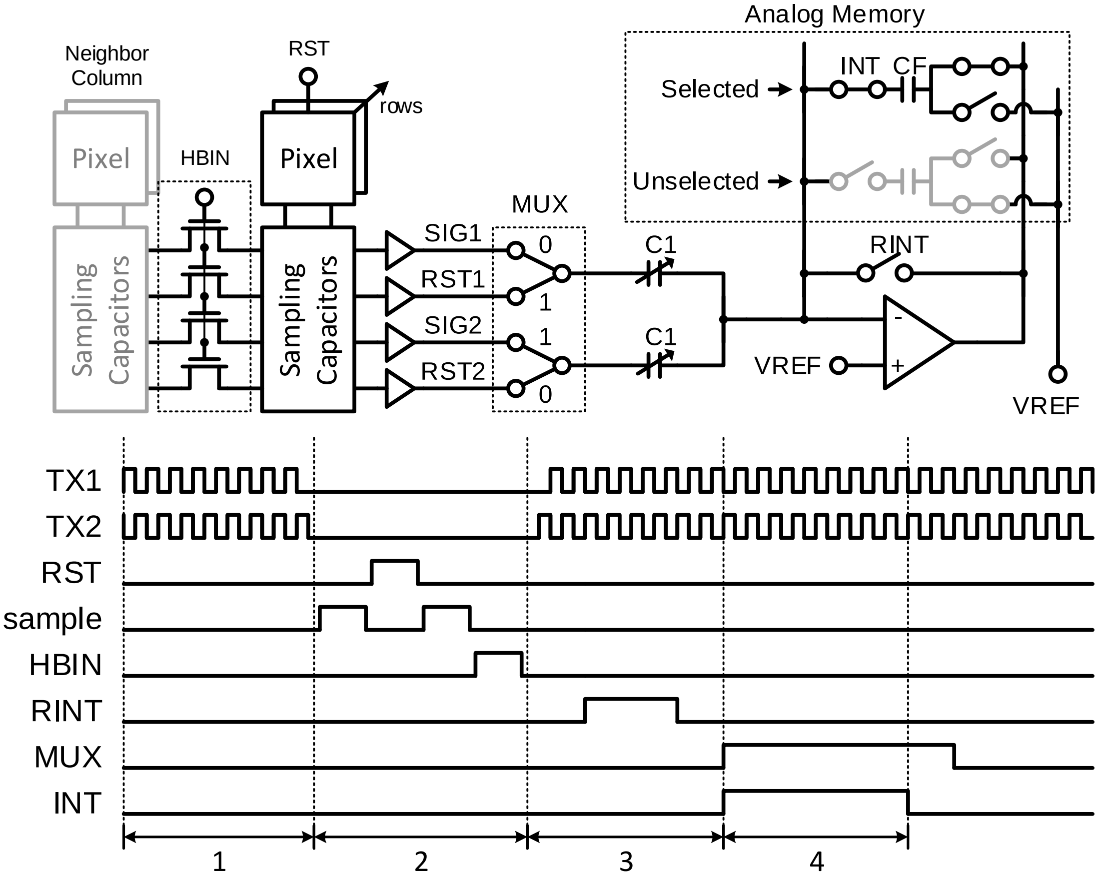


Fig. 5-10. BGL suppression circuit and its timing diagram.

During the phase 1, the pixel is demodulated at 12.5MHz to integrate the signal. In the phase 2, the signal level of the pixel is sampled, followed by the reset levels for DDS. After the sampling, HBIN becomes high to perform the pixel binning between the sampling capacitors in the next column. In the phase 3, the integrator is reset (RINT) while SIG1, RST2 and the offset voltage of the OPAMP are sampled in the input capacitors. Then, in the phase 4, INT signal transfers the charges into the feedback capacitor, CF, while the two signals, RST1 and SIG2, are applied. After this operation, the incremental voltage across the feedback capacitor becomes:

$$\Delta V_{CF} = \frac{C1}{CF}[(RST1 - SIG1) - (RST2 - SIG2)]. \tag{19}$$

As can be seen in (19), the integrator integrates the differential signal of each pixel with DDS operation. Although the DDS operation cannot remove the reset noise introduced in the first sub-integration time, it removes all the other reset noise between sub-integration times, the DC offsets of the pixel, and the flicker noise of the source follower. The value of input capacitance (C1) can be digitally selected, providing a programmable gain.

The feedback capacitor CF in the integrator is selected from the capacitor bank which serves as the analog frame memory. The frame memory consists of 84×64 cells and it can hold the entire frame signals when 4×4 pixel-binning is applied. There are 64 cells assigned for every four columns, and only one of them is selected at once for the integration. As shown in the Fig. 5-10, the selected capacitor forms the feedback loop for the integration, whereas it is isolated from the circuit when not selected. One plate of the unselected capacitor is biased with a DC voltage to prevent any floating capacitors.

The overhead of the implemented circuit could be minimized by sharing resources. The area required for the frame memory requirement was relieved by 16 times in our implementation by pixel-binning operation and resolution recovery technique. The DDS and programmable gain amplification are achieved for free during the integration with no additional area and power consumption. The only overhead is the increased power due to the multiple capturing with repeated sampling and integration. However, this overhead is still less than the conventional multiple capturing schemes because the sensor requires A/D conversion only once in each frame regardless of the number of sub-frames thanks to the analog approach.

5.3.3 Single-Slope ADC Circuits

Column-parallel 10-bit single-slope ADCs with a conversion time of $20\mu s$ have been implemented in the sensor. Fig. 5-11 shows the schematic and timing diagram of the ADC. Unlike the ADCs in the conventional image sensors, the input voltage ranges from negative to positive voltages depending on the sign of the differential mode of the pixel output. Therefore, the operation should be slightly different from the conventional single-slope ADCs. The ADC consists of an offset-cancelled comparator, a latch and a global counter. To cover the input range from negative to positive values, a ramp signal initially steps down to the negative full-scale (-0.5 V) and ramps up to the positive full-scale (0.5 V).

The A/D conversion starts with offset cancellation of the comparator. The operation achieves offset-cancellation in both input and output storages [89]. The offset of the preamplifier is generated by a unity gain feedback (R1), and is sampled in the input capacitor along with the input reset voltage. Then, the output offset is sampled (R2). After storing the offsets, the input signal is sampled (S), and the ramp signal starts. The comparator generates a latch signal when the ramp signal crosses the sampled voltage.

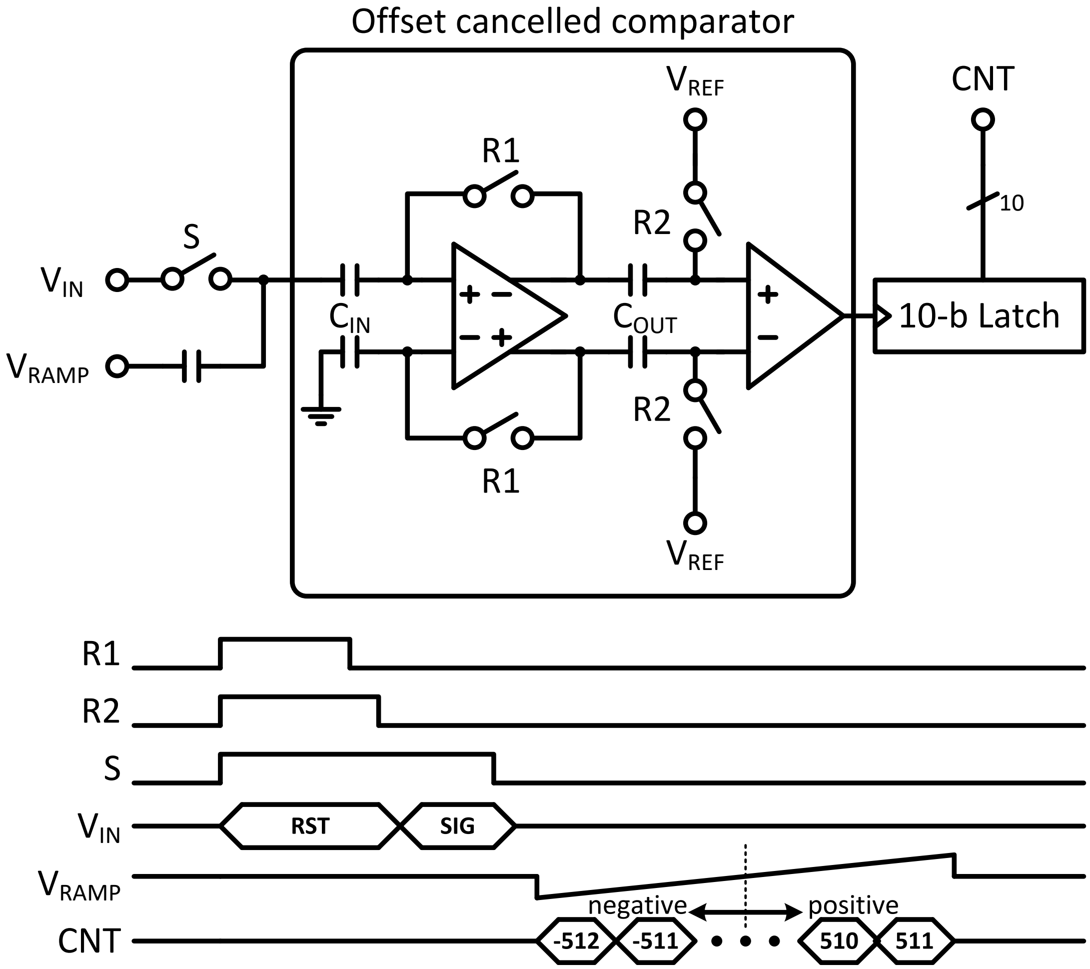


Fig. 5-11. Single-slope ADC with an offset-cancelled comparator.

5.4 Experimental Results

The prototype chip has been fabricated using a 0.11μm 1P4M CMOS image sensor process and has been fully characterized. The core size is 3mm×4.1mm except for the test patterns as shown in Fig. 5-12.

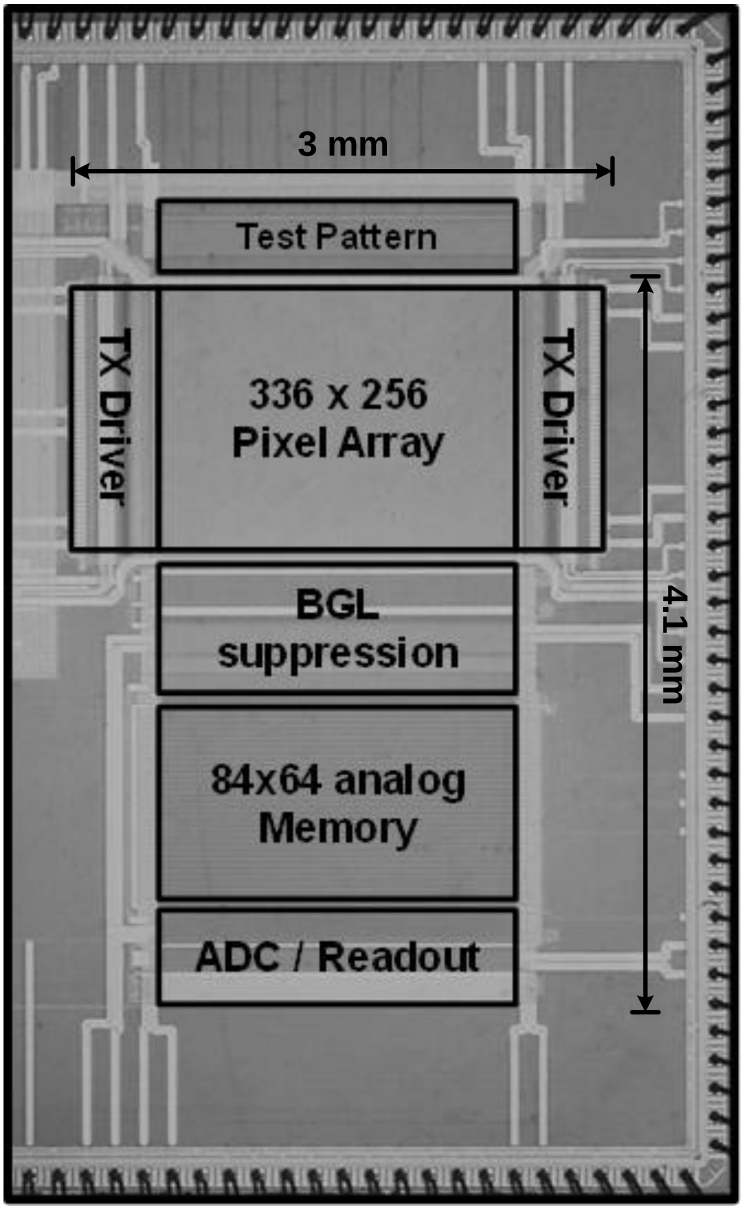


Fig. 5-12. Die microphotograph.

5.4.1 Experiment Setup

To test the sensor chip, ninety-six IR-LEDs with the wavelength of 850-nm were used for pulsed illumination at the modulation frequency of 12.5MHz. The emitted optical power was measured as $650\mu W/cm^2$ at 1-m distance from the LED array.

The sensor was equipped with a lens and an IR-pass filter. The lens had the focal length of 8mm and the F-number of 1.4. An IR bandpass filter with 40nm bandwidth was placed between the lens and the sensor in order to reject all spectral components but the emitted light from the LED array, i.e. IR at 850nm. The optical filter helped improve the BGL suppression by the factor of 10.

5.4.2 Imaging Modes and Post Processing

Both 2-D and 3-D images could be captured using the fabricated sensor. When capturing 2-D images, only a half of the pixel circuit was used. The half pixel is identical to the conventional 4T pixel structure. Thanks to the 4T pixel operation, true correlated double sampling (CDS) could be performed for better image quality. All 336 ADCs were used to digitize the image at 336×256 resolution.

Unlike the 2-D imaging, the 3-D imaging utilizes the full pixel circuits. The differential signal from the pixel could be digitized by 336 ADCs to provide the full resolution of 336×256 when BGL cancellation scheme is not used.

When BGL cancellation is used, 4×4 pixel binning was performed for two reasons as stated in 5.3.2. First, it improves the image quality from high shot noise caused by BGL. Second, the size of the frame memory could be reduced by 16 times. Despite the reduced resolution from the binning operation, the full-resolution image could be attained by super-resolution technique after obtaining 16 images from the sub-pixel shifting in x- and y-directions. In this mode, only one fourth of the column circuits were activated to save the unnecessary power consumption.

For the depth calculation, only two quantities, $\Delta Q(0)$ and $\Delta Q(\pi/2)$ are need according to the equation (18). However, in the experiments, $\Delta Q(\pi)$ and $\Delta Q(3\pi/2)$ were also measured in order to eliminate the mismatch in the charge transfer efficiency between the transfer gates. Each measurement was made during the integration time of 25ms, resulting in the total measurement time of 100ms. The 25ms-long integration time was divided into sixteen 1.56ms-long sub-integration times.

After the depth calculation, no additional processing but a simple offset calibration was applied to compensate the clock skew over the entire pixel array.

5.4.3 Linearity and Noise Performance

Fig. 5-13(a) shows the measured depth over the range from 0.75 to 4m. The measured depth agrees well with the actual depth. The data points were acquired using a white planar subject at different locations. The non-linearity (NL) over the measured range was 1%, where NL is defined as the average difference between the measured and actual depths divided by the measurement range [37].

The noise of the sensor was measured from the 100 pixels at the center of the 100 consecutive depth images. The measured noise and its theoretical estimation are plotted in Fig. 5-13(b). The theoretical noise or the standard deviation of the depth, ΔD can be obtained using the error propagation equation [80]. The depth error due to the noise is given by

$$\Delta D = \frac{c}{8 \cdot f_{MOD}} \cdot \frac{\sqrt{N_R^2 \cdot (N_0^2 + N_1^2) + N_0 \cdot N_1 (N_0 + N_1)}}{(N_0 + N_1)^2} \qquad (20)$$

where f_{MOD} is the frequency of (de)modulation signals, N_R^2 is the sum of reset and readout noise powers, and N_0 and N_1 represent the number of charges of $\Delta Q(0)$ and $\Delta Q(\pi/2)$, respectively. In the noise calculation, we considered the light attenuation according to the distance, that is, N_0 and N_1 were scaled down by D^2.

The measurement and the theoretical calculation match very well. First in the shorter distance where the shot noise nominates, the noise increases almost linearly. On the other hand, in the longer distance, the noise increases drastically as the readout noise becomes more dominant. The temporal noise was measured as low as 1.47cm-rms at 1-m.

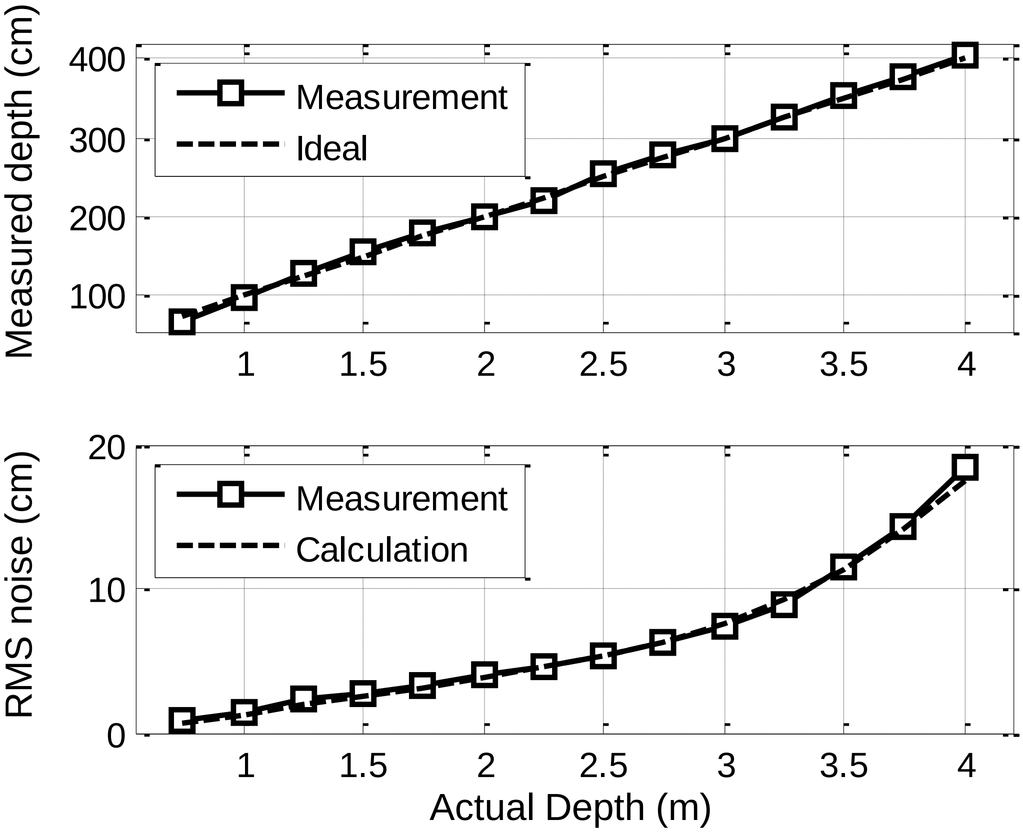


Fig. 5-13. (a) Measured depth and (b) RMS noise versus actual depth.

5.4.4 Background Light Cancellation Performance

The performance of the BGL cancellation scheme was characterized in terms of BGL to signal ratio (BSR), which is the same concept but reciprocal of the signal to BGL ratio (SBR) defined in [75]. Under the maximum attainable BGL during experiment (100klx), BSR was measured as 100 with the use of optical filtering, and 16 without it, respectively.

We also characterized the accuracy of the sensor in terms of the random temporal noise and the static offset under various BGL conditions as shown in Fig. 5-14. The random noise increased along with the strength of the BGL due to the increased shot noise. The maximum noise was measured as 3.65cm-rms under the BGL of 100klx.

The static offset tended to increase as the BGL with the maximum value of 2.6cm at 100klx BGL. The main cause of this static error is the imperfect BGL cancellation. The leftover BGL components after the cancellation will be added to the numerator and denominator of (18). Higher BGL added bigger offsets to both the numerator and denominator, resulting in higher errors in depth estimation. The imperfect BGL rejection can be characterized by BGL rejection ratio (BLRR), which is defined as the ratio between the responsivity of a pixel to BGL and that to signal light [90]. The BLRR of the prototype sensor was measured as -59dB.

From the measurement, the sensor managed an acceptable accuracy under extreme BGL conditions for most TOF applications which require centimeter-range accuracies.

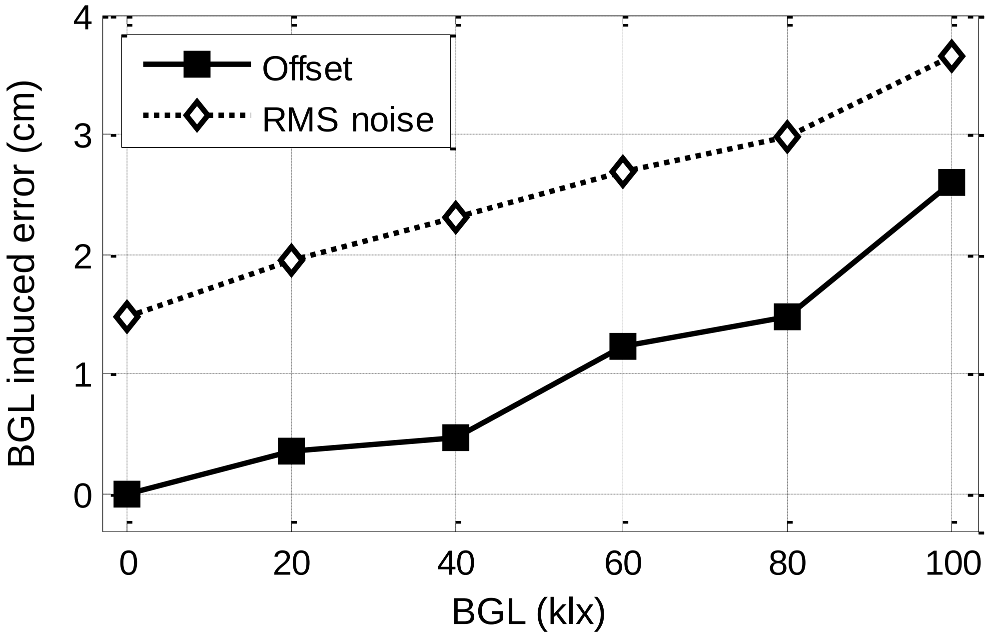


Fig. 5-14. Accuracy vs. BGL.

5.4.5 Sample Images and Superresolution

Fig. 5-15 shows the captured depth images under various BGL conditions to evaluate the implemented BGL suppression scheme. Without the scheme, the sensor was saturated when exposed to BGL higher than 10klx, whereas the sensor could operate properly under 100klx with the BGL suppression scheme turned on. In outdoor experiments, a 3-D image was successfully captured on a bright sunny day where the BGL was measured as 100klx [39].

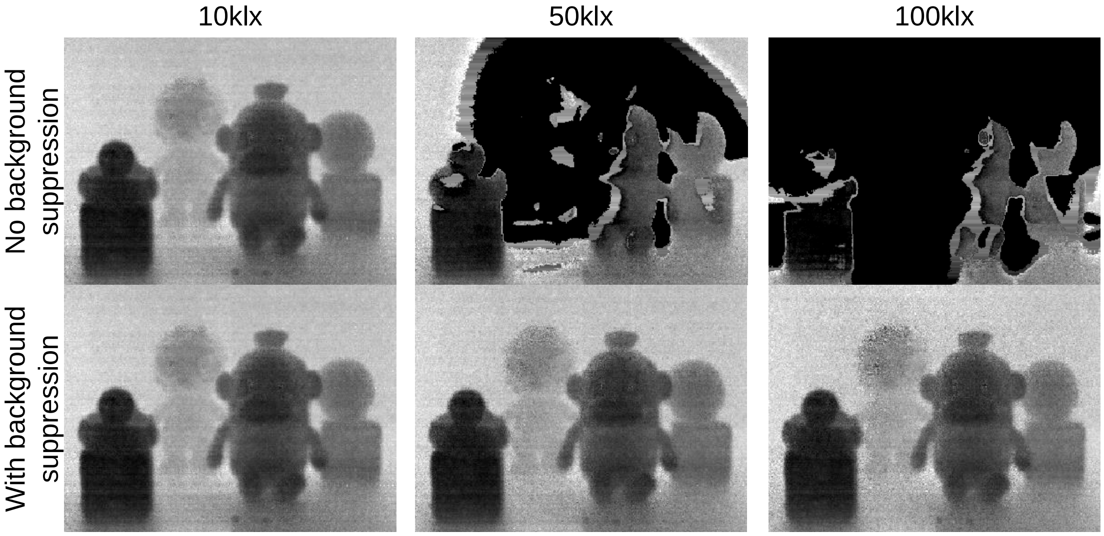
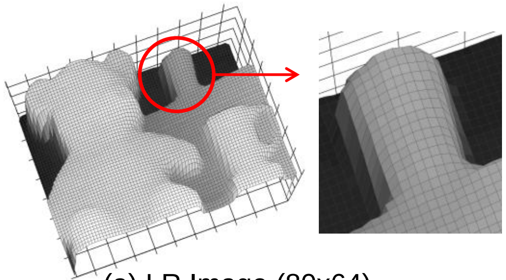
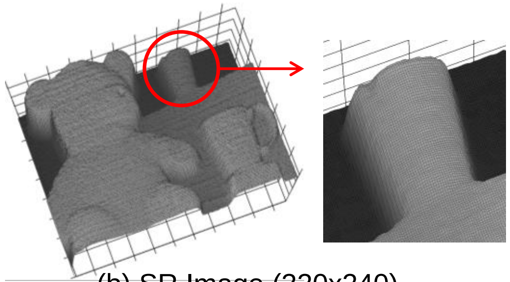


Fig. 5-15. Captured images under 10klx, 50klx, and 100klx BGL conditions with and without applying the proposed BGL suppression scheme.

The proposed super-resolution recovery and the electrical superpixel-shift are also shown in Fig. 5-16. A recovered image with the full resolution of 336×256 (Fig. 5-16(b)) was acquired by weaving 16 frames of low resolution images (Fig. 5-16(a)). A cup with the curved surface is magnified for better visual representation of the resolution improvement. Finally, a 3-D textured image was constructed as shown in Fig. 5-16(c). Both 2-D and 3-D images could be captured from the same sensor by time-division multiplexing as in [34]. Since both images share the same optics, they can be easily fused without any intensive post-processing.



(a) LR Image (80x64)



(b) SR Image (320x240)

(c) SR Image with 2D Texture

Fig. 5-16. Resolution enhancement by super-resolution: (a) low-resolution (LR) image after 4x4 pixel binning, (b) super-resolution (SR) image with a fully recovered resolution from 16 LR images, and (c) SR image with 2-D texture.

5.4.6 Summary and Remarks

The sensor performance is summarized and compared with the state-of-the-art TOF sensors in Table 5-1 and Table 5-2, respectively. The proposed sensor achieved a very small pixel size of $5.9\mu m$. The noise, non-linearity, and dynamic range performance are comparable to other sensors except [82], which is not suitable for mobile applications due to the requirement of mechanical scanners. Also, the fabricated sensor can provide adaptive operation which allows the optimal BGL suppression on demand by changing the spatiotemporal resolution.

The sensor performance can be further improved in terms of area, spatiotemporal resolution and power consumption. The area overhead of the BGL suppression scheme can be relieved by either using a technology with a higher capacitor density or a 3-D stacking option. Higher speed ADCs can help increase the frame-rate and the resolution. To decrease the power consumption, more effort should be on pixel design, because demodulation involves high frequency switching of the transfer gates in the entire array.

5.5 Conclusions

A TOF depth sensor with a column-level BGL suppression scheme has been presented for high-resolution outdoor 3-D imaging in mobile applications. The adaptable spatiotemporal resolution control has been realized by pixel-binning and super-resolution. A prototype chip was fabricated in an array of 336×256 pixels using 0.11μm CMOS technology, and achieved a very small pixel size of $5.9\mu m$. The fabricated chip was tested in both indoor and outdoor conditions. The sensor successfully demonstrated its ability to suppress background illumination over 100klx at full resolution. The measured non-linearity is 1% over a range of 0.75~4m. RMS noise is measured as 1.47cm at 1m distance. With its adaptable operation and high BGL suppression capability, the proposed sensor is suitable for 3-D mobile imaging applications such as gesture/motion recognition and depth profile.

Table 5-1. Sensor Configuration and Performance Summary.

Process	$0.11\mu m$ 1P4M CMOS Image Sensor
Power	680mW @ 3.3V (analog) / 1.5V (digital)
Core area	3.0mm×4.1mm
(De)Modulation	850nm LED operated at 12.5MHz 650μW/cm^2 optical power at 1m distance 50% demodulation contrast at 12.5MHz
Lens	F# 1.4 / focal length of 8mm
Resolution	Variable (max. 336×256)
Pixel	$5.9\mu m \times 5.9\mu m$ with a fill factor of 24%
Integration time	100ms (25ms / phase) during measurement
Frame-rate	10fps for measurement (max. 160fps)
Range	75cm~400cm
RMS noise	1.47cm measured at 1m distance
BGL suppression	>100klx (65mW/cm^2 at 1m distance)
BGL to signal ratio	>100 (65mW / 650μW)
BGL-induced offset	<2.6cm for 0~100klx BGL

Table 5-2. Performance Comparison.

	[34]	[41]	[75]	[78]	[82]	This work
Technology	$0.11\mu m$ CIS	$0.35\mu m$ CMOS	$0.6\mu m$ CMOS	$0.18\mu m$ CMOS	$0.18\mu m$ CMOS	$0.11\mu m$ CIS
BGL suppression scheme	Opt. filter	Opt. filter Burst mode	Pixel-level charge subtraction	Pixel-level CDS	Temporal-spatial correlation	Column-level suppression / Opt. filter
Pixel pitch	$14.6\mu m$	$15\mu m$	$125\mu m$	$29.1\mu m$	$25\mu m$	$5.9\mu m$
Fill factor	38.5%	19%	66%	34%	70%	24%
Maximum Resolution	500×274	336×252	16×16	160×120	340×96 w/ scanning	336×256 superresolution
Frame-rate	11	30	16	27.8~80	10	10~160*
BGL suppression	n.a.	n.a.	150klx	20klx	80klx	>100klx
BGL to signal ratio	n.a.	n.a.	1000	n.a.	n.a.	>100
Range	0.75~4.5m	1.8~3.3m	0.1~3.2m	1~4.5m	10~100m	0.75~4m
Non-linearity	0.93%	1.73%	-1~2cm	Unknown	0.37%	1.01%
RMS noise	1.35cm @ 1m	Min. 2.35cm	1cm @ 1m	10cm @ 1m	<10cm @ 100m	1.47cm @ 1m

* Spatial resolutions are 336×256 and 84×64 at 10fps and 160fps, respectively

Chapter 6

Conclusion and Future Work

This thesis presents time-resolved image sensors for time-of-flight depth imaging and fluorescence lifetime imaging. This work especially intends to meet the challenging requirements of the emerging applications such as real-time fluorescence lifetime imaging and outdoor depth imaging. We proposed two time-resolved imaging schemes: a single-shot center-of-mass method pixel for high-speed and high-accuracy fluorescence lifetime imaging, and an adaptive background cancellation scheme for time-of-flight depth imaging.

6.1 Summary

The single-shot center-of-mass method (CMM) is proposed to provide both the high speed and high accuracy using two-tap pixels. Conventional approaches such as rapid lifetime determination (RLD) and multi-sampling CMM (MS-CMM) could only be designed for either high speed or high accuracy. To overcome this trade-off, the proposed method incorporates the signal processing unit into the integration node of the pixel; thus, the pixel can oversample the optical signal for high accuracy without increasing data transmission rate. As opposed to the conventional two-tap pixels, the proposed approach does not require any pixel circuit change. The only additional hardware is a digitally controllable delay line in the system level to shift the laser trigger signal during the integration time. The measurement result showed that the proposed approach provides the same imaging speed as RLD, while achieving higher accuracy than RLD and MS-CMM. In theory, the MS-CMM should provide the best accuracy, but the photo-bleaching effect degraded the signal quality during MS-CMM measurement. This, again, proves the advantage of the proposed sensor.

We also proposed an adaptive background light cancellation scheme for outdoor time-of-flight depth imaging. Unlike the conventional approaches, the background light suppression circuits are implemented in the column-level of the sensor. This greatly reduces the pixel size, resulting in higher pixel density. The employment of column-level approach also enables reconfigurability of the scheme; the level of background suppression can be enhanced by compromising spatiotemporal resolution. The spatiotemporal resolution is controlled through pixel-binning and super-resolution techniques. The added flexibility allows the sensor to achieves optimum performance under various environmental condition. Thanks to the column-level approach, the fabricated sensor achieves the smallest pixel size of 5.9μm among all reported time-of-flight sensors so far. The sensor is tested with an indoor lamp as well as outdoor sunlight; the sensor can achieve over 100klx background light suppression.

6.2 Suggestions for Future Work

In this work, we introduced three methods for real-time portable time-resolved imaging. We suggest the following research topics for further:

- Contact imaging: Contact imaging is an interesting application for fluorescence lifetime imagers. Contact imagers are compatible with most microfluidics platforms and need almost no optics, making the total system very compact.
- Hybrid imaging in endoscopy: Since the time-of-flight depth sensor and the fluorescence lifetime imaging share the same camera architecture, both imaging modes can be utilized for better usability. For example, time-of-flight depth sensing provides a useful images for minimally invasive surgery (MIS) [91] while the fluorescence lifetime imager can offer real-time diagnoses or cancer detection. Since an endoscope has a very limited space, a hybrid camera approach for both applications will be more feasible than putting two cameras in the endoscope.
- Delay controller: In many cases, there are need for adjusting the delay of the light pulses. For example, the single-shot CMM pixel in this thesis relies on delay control of the light emitter. The delay can also be controlled to calibrate the system statically or adaptively for better synchronization between the emitter and the sensor.
- Embedded adaptive controller: In this thesis, we presented an adaptive background light cancellation scheme. An embedded controller with ambient light sensor can be integrated into the system to adaptively reconfigure the sensor.
- Light emitter: System performance is heavily influenced by the light source, especially the pulse width of the emitter pulse and the optical power. Therefore, it is important to investigate this topic in a more extensive study. In this thesis, we used a custom-built LED array board for time-of-flight imaging, and a VCSEL

laser board for fluorescence lifetime imaging. If light sources are available for higher power and shorter pulse width, the frame-rate and the time resolution can be improved.

- System on a chip: For ultimate portability needed in applications such as implantable devices and endoscopic applications, the form factor of the sensing system should be in millimeter scale. Integrating the CMOS sensor, optics, and VCSEL module with all required peripherals in a small package will make this system more practical for mobile applications or endoscopy.
- Multi-tap pixel: We propose a next generation pixel structure that offers advantages of both TCSPC and gated imaging such as high speed and high accuracy even under low light conditions. We present the preliminary research in the appendix. The multi-tap pixel has multiple transfer gates and floating diffusions to implement multiple time-gates in each pixel. Time-to-space conversion is employed to overcome the pixel's geometrical problem for multiple gate implementation. A compressive analog-to-digital converter is also proposed to compress the data from the pixel. The compression is configurable to balance the ADC speed and the I/O speed for the best overall imaging speed. This research requires a close relationship with the foundry because development of a pixel needs a full understanding of the process.

Appendix A

Preliminary Research: Multi-Tap Pixel

A.1 Introduction

In time-resolved imaging, time and intensity are generally resolved sequentially because the sensors are not able to resolve both time and intensity simultaneously. For example, time-correlated single-photon counting (TCSPC) first resolves the time of each photons and then it resolves the intensity by counting the number of photons in each time bin. On the other hand, the time gating first measures the intensity of the optical signal within the time-gate while the timing information is resolved from multiple measurements from many time gates.

As the name suggests, TCSPC resolves the timing of individual photons but only one photon at a time. When a photon is detected by a single-photon detector, the time is instantaneously resolved by a time-to-digital converter (TDC). This process is repeated many times to resolve times for all photons. When all photons are measured, the number of photons at a specific time is counted to resolve the intensity of that particular time bin. This method is the most precise among all time-resolved imaging methods; however, the measurement and data transmission time is too long for real-time application because of the excessive repetition and large amounts of data. Although using a laser with high repetition-rate and data compression scheme may relieve the problem, the state-of-the-art TCSPC chip achieved only 400k pixels per second [49].

In contrast, time-gating measures the number of photons during a specified time-gate. In each measurement, the intensity or the photon count is resolved by an analog-to-digital converter (ADC). The measurement is repeated with different time gates to resolve time information. Because this method offers the fastest imaging speed, it can be used in real-time applications as in [48]. However, this method suffers from low signal-to-noise-ratio (SNR) because only a small portion of the photons are resolved at time, leaving other photons outside of the time-gate. Though wider time gates can get more photons time-resolved, the SNR is not improved because of the lower time resolution of wider time gate.

From the discussion on the two time-resolved imaging methods, it is obvious that there is a fundamental limit in the time-resolved imaging, i.e. the trade-off between imaging speed and the SNR.

Efforts has been made to overcome this trade-off but none of them were successful in imaging applications. For example, a super-pixel was made with 1024 SPADs to improve the photon detection rate of TCSPC method by 8 [60]. However, this system is designed for a point measurements and it cannot be applied to an imaging system due to its large pixel size (720um). The imaging measured in pixel-rate does not improve in this case because the increase of the pixel size is more than the speed improvement.

To improve the SNR of time-gated imaging, the number of simultaneous time-gates in one integration time should be increased. This increased number of gates will provide better time-resolution for each photon. In CMOS technology, pixels with up to 4 taps [92], [93] are reported; however, the SNR improvement of 4-tap gating from 2-tap gating is not significant. Moreover, further increasing the tap count is very difficult due to physical layout of the pixel and additional clock generation and distribution.

In this chapter, we propose a multi-tap pixel that can easily achieve more than 10 taps, for high SNR and high speed time-resolved imaging. We also propose a compressive analog-to-digital converter (ADC), where the data is compressed during analog-to-digital conversion in order to address the speed limitation of data conversion and transmission. A prototype sensor with 1x256 pixels is designed and fabricated using a CMOS process and it has been fully characterized.

A.2 Proposed Multi-Tap Pixel

Before discussing the multi-tap pixel, we first review the operation of the two-tap pixel. Fig. A-1 shows a cross-section of a typical two-tap pixel [34]–[42] consisting of two transfer gates (TX), two floating diffusion (FD) nodes and a pinned photodiode at the center. In this figure, the photo-generated electron is directed toward FD_1 through TX_1. In general, the electrons can be transferred to any FD node by controlling the transfer gates.

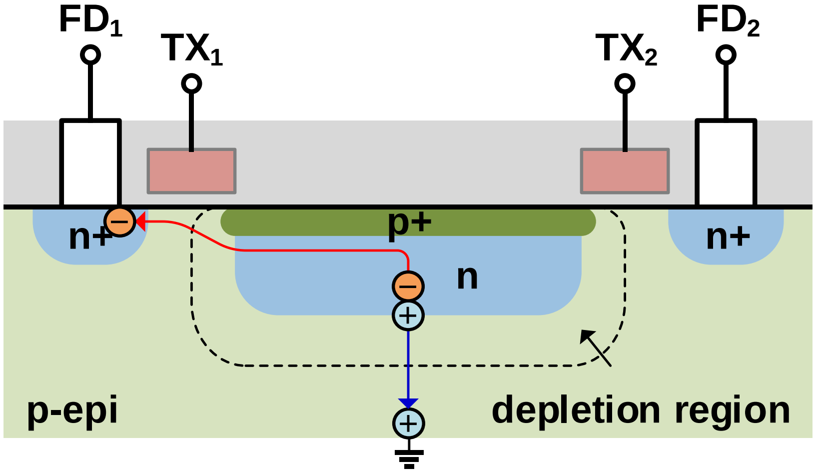


Fig. A-1. Cross-section of two-tap pixel in CMOS process

More taps can be added to this conventional structure; however, implementing more than four taps is impractical for two reasons. First, the pixel requires as many clock signals as it has transfer gates. For example, to resolve the total range of 16-ns in 1-ns resolution, we need 16 non-overlapping clock signals with 1ns pulse-width. Distributing the signals to the entire pixel array is very challenging because of large load capacitance and long clock lines.

Second, the number of taps is constrained by the physical space. The taps (FD and TX pair in this particular pixel) must be located right next to the photodetector. To add more taps to the device, the perimeter of the photodetector has to increase as shown in Fig. A-2. However, increasing the perimeter or the detector size also increases the travel length of the electrons and hence the time resolution is not improved. In addition, the asymmetric placement of the taps will result in non-symmetric gating behavior.

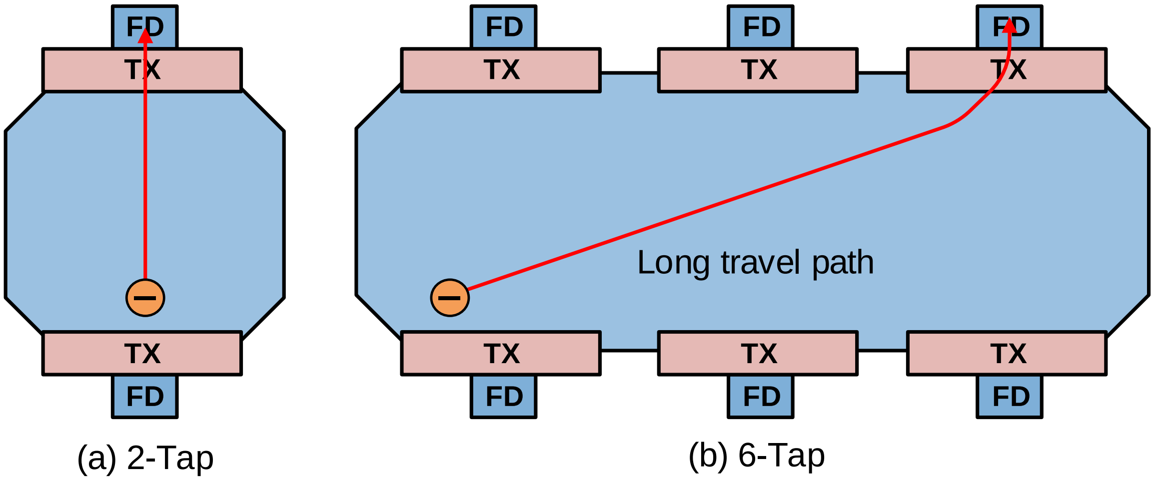


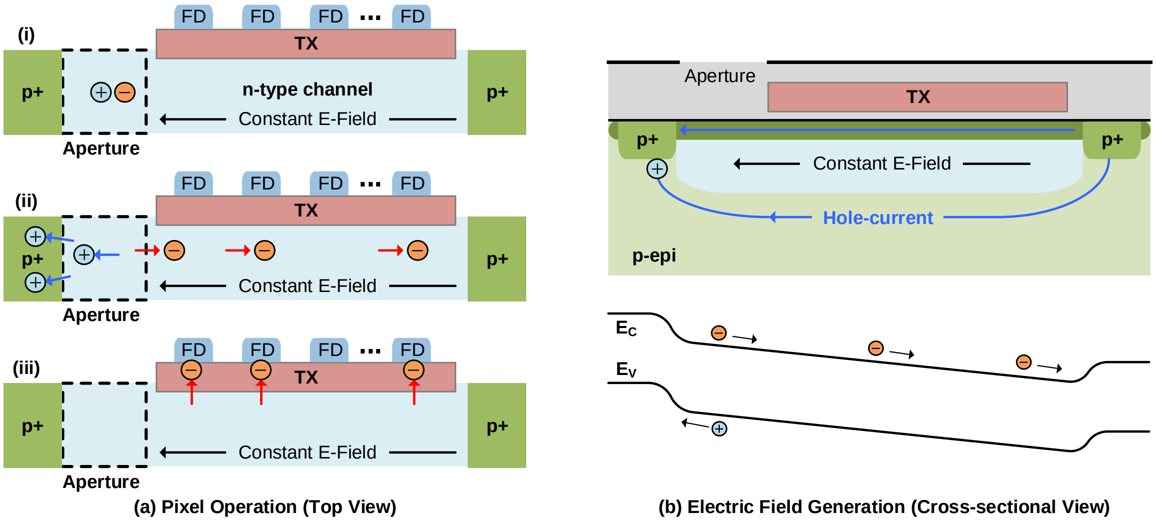
Fig. A-2. Implementing multiple taps in conventional structure (a) 2-tap (b) 6-tap

A.2.1 Structure and Operation of Proposed Multi-Tap Pixel

We propose a CMOS-compatible time-gating pixel structure that can easily extend the number of taps. The pixel is based on time-to-space conversion (SPC) originally used in streak camera [94] where the time is resolved in the spatial domain; i.e. light intensities at different times are mapped to different locations, and then spatially resolved.

Fig. A-3(a) shows the top-view layout and the operation of the proposed multi-tap pixel. The pixel consists of an n-type buried region (aperture and channel), multiple pairs of FDs (floating diffusion) and TX (transfer gate), and two substrate contacts (p+).

The pixel operation is performed in three steps. First, the photons are detected in the aperture region as shown in Fig. A-3(a)(i). Each detected photon generates an electron-hole pair, which are then separated by the electric field, shown in Fig. A-3(a)(ii). Holes are drained by the adjacent p+ contact and electrons drift to the right at a constant velocity in response to the constant electric field. In a short time, electrons are located in different places depending on their generation time; i.e. the earlier the photon generates the electron, the farther the electron travels. Finally, TX is turned on and the distributed electrons are transferred to the closest of the spatially distributed taps, as shown in Fig. A-3(a)(iii). In the end, each FD node stores the intensity information at a specified time, resolving both intensity and time simultaneously.



(a) Pixel Operation (Top View)

(b) Electric Field Generation (Cross-sectional View)

Fig. A-3. Proposed multi-tap pixel (a) operation (b) electric field generation in channel

Fig. A-3(b) shows how the constant electric field is generated in the channel region. First, two p+ contacts are biased to produce a hole current through the substrate. The current creates a potential gradient that aligns with the current path. This potential gradient in the substrate also causes the gradient of pinning potential in the channel region, as shown in the energy band diagram. Therefore, a constant electric field in the channel is formed.

The proposed pixel has two advantages over the conventional structure. First, the number of taps is not limited by the physical space because the taps can be laid out apart from the photodetector. Second, the pixel requires only one TX signal regardless of the number of taps because the photon arrival time is resolved in spatial domain rather than in the original time domain.

So far, we have qualitatively described the operation principle of the proposed pixel. Now, we will quantitatively analyze the pixel to find out the resolution limit of the pixel, or the minimum time that the pixel can resolve.

The time resolution of the proposed pixel depends on the speed of the electron in the channel and the spatial resolution of the taps. Higher electron speed means higher time-to-space conversion gain and hence better time resolution. Denser placement of taps gives better time resolution as well because of the time-to-space conversion. Therefore, high electron speed and device density are desirable for high time resolution.

The speed of the electron or the time-to-space conversion gain (A_{TS}) depends on the semiconductor material and the doping concentration of the channel. In a semiconductor, the electron speed is given as the product of electron mobility (μ_e) and the electric field (E) in the region.

$$A_{TS} = \mu_e \cdot E = \mu_e \cdot V/L$$

where, V and L is the applied voltage across the channel and the length of the channel, respectively. The electron mobility in lightly doped silicon is around $500 cm^2/V \cdot s$, so the equation for electron mobility in the channel is approximately

$$\mathrm{A_{TS;Si}} = 50 \cdot \mathrm{V/L}$$

where the voltage and the length is in [V] and [μm]. Let N be the number of taps, ΔL be the tap spacing, and ΔV be the voltage across the adjacent taps. Time resolution is then given as

$$\Delta \mathrm{t} = 20\mathrm{mV} \cdot \Delta \mathrm{L}^2/\Delta \mathrm{V}.$$

. In most modern CMOS processes, the taps can be laid out in a pitch as small as 1μm. The voltage across the channel can be as high as .5V while maintaining a proper TX operation. Assuming that N taps are placed with a 1μm pitch along the channel, the time resolution is then

$$\Delta \mathrm{t} = \frac{20\mathrm{mV}}{0.5/\mathrm{N}} = 0.04 \cdot \mathrm{N}.$$

Since the time resolution depends on the number of taps, the resolutions of 320ps, 640ps, and 1.32ns are achievable with 8-, 16-, and 32-tap pixels, respectively.

A.2.2 Device Simulation of Multi-Tap Pixel

A device simulation is performed using Synopsys Sentaurus to verify the operation of the proposed pixel. Fig. A-4 shows the device structure created for the simulation. To avoid extremely long simulation time, only four taps are implemented in the structure. In the figure, GND and NGND represent the two p+ taps to induce hole-current (substrate current). GND and NGND are biased at zero and negative voltage, respectively. PDP is the p+ layer at the surface to make the buried channel (PDN). PDN is below the PDP layer and not visible in the figure. A direct contact (DRN) to the buried PDN is biased at VDD to deplete the channel. The direct contact can be replaced with a tap with always-on TX. The doping concentration of the buried channel region for simulation is shown in Fig. A-5.

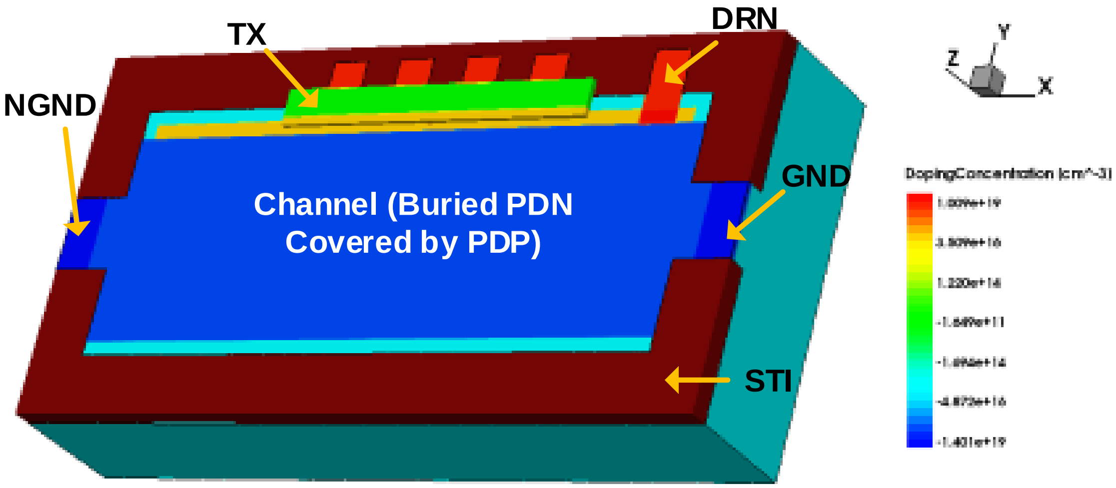


Fig. A-4. Device structure for simulation of a multi-tap pixel

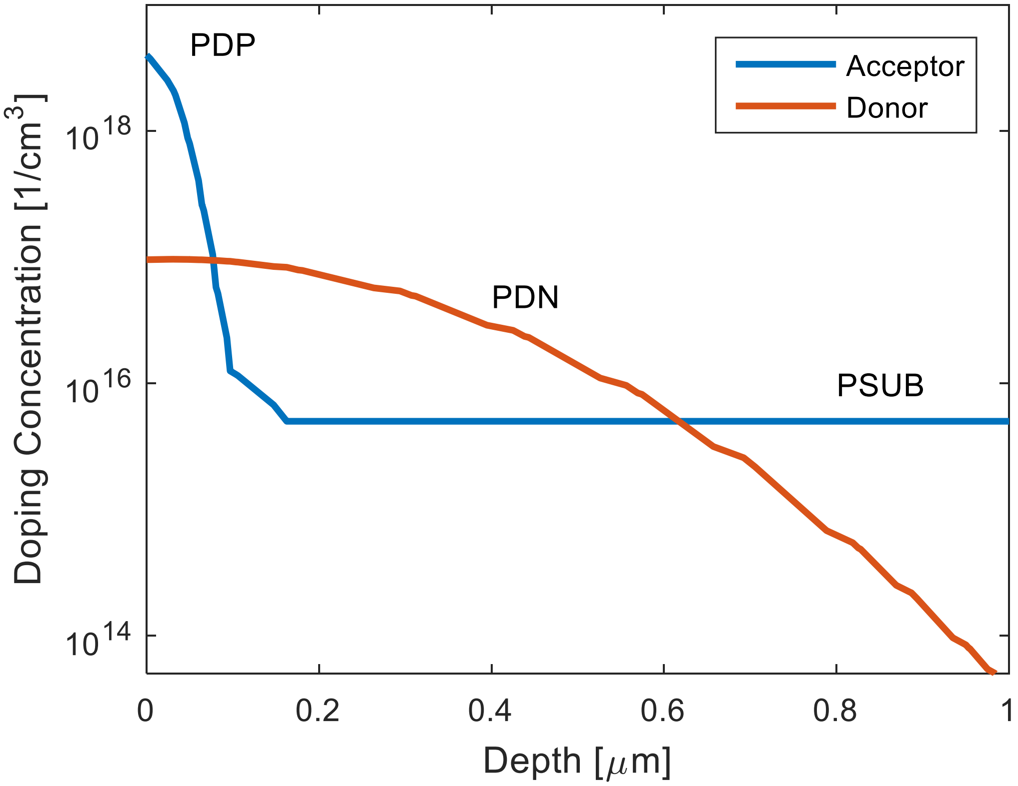


Fig. A-5. Doping concentration assumed for the channel region

At the beginning of the simulation, NGND is initially biased at 0V and then the bias is ramped down to -0.5V. This modulates the body potential and the channel potential as described before. Fig. A-6 shows the x-axis electrostatic potential profile of the PDN region at z=0.3μm and y=2.5μm. It is clear that the potential of PDN is modulated by the substrate potential. The slope, or electric field, is measured to be 42mV/μm from the simulation. Any electrons in the channel region will travel at a constant velocity towards GND and then will be finally drained to DRN.

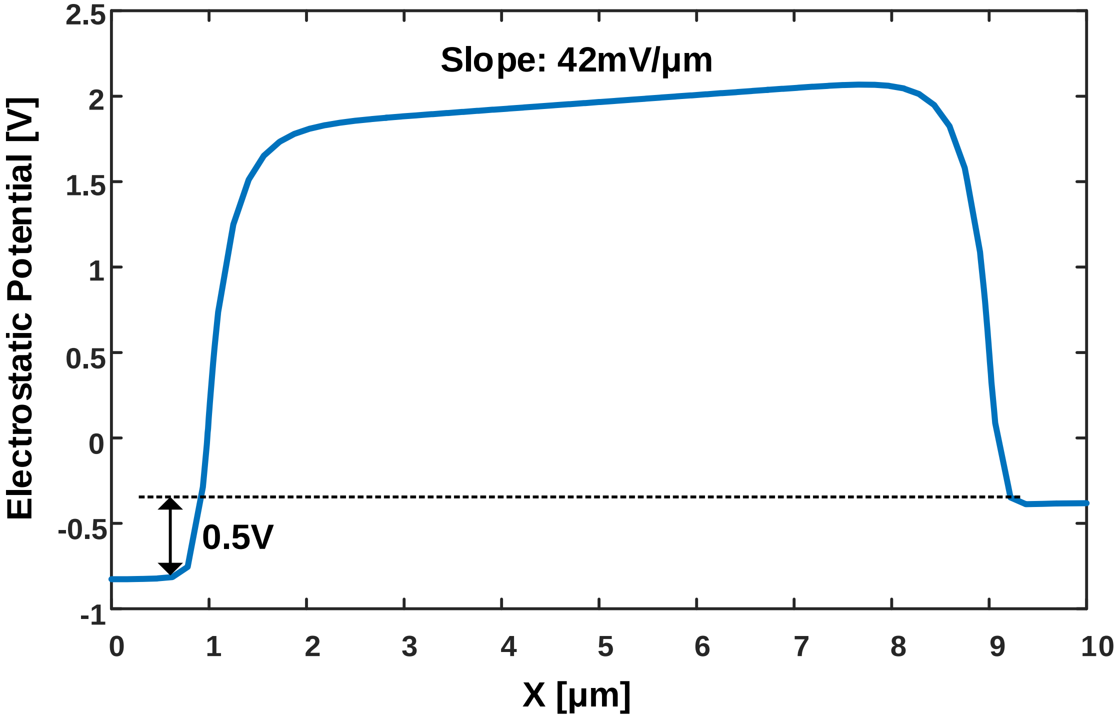


Fig. A-6. Modulated electrostatic potential along the channel (PDN region).

After photoelectrons are spatially distributed based on to their generation time (or photon absorption time), TX gate is turned on and all electrons are captured in the closest tap. This charge transfer is verified by simulation as shown in Fig. A-7. It shows the cross-section of the device at x = 6.75μm. When TX is off (Fig. A-7(a)), the channel region is isolated from the floating diffusion. The activation of TX (Fig. A-7(b)) creates an inversion layer under the gate and the charges in the channel region are transferred to the floating diffusion.

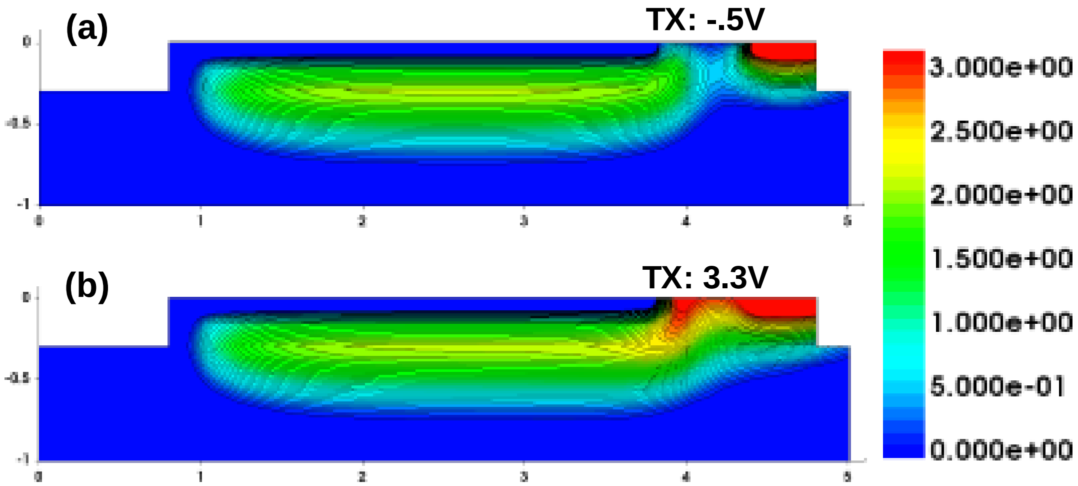


Fig. A-7. Potential contour plot in the cross-section of the device: (a) when TX is off and (b) when TX is on

A.3 Compressive Analog-to-Digital Converter

Though the proposed multi-tap pixel provides high time resolution, the large amount of data from the pixel limits the frame-rate of the sensor; i.e. the time-resolved sensor with N-tap pixels produces N intensity images whereas the conventional intensity imager produces only one image. For a high frame-rate or imaging speed, both analog-to-digital conversion (ADC) and data transmission speed must be increased. However, the speed of the ADC and I/O circuits is limited by the given technology node.

In this section, we propose a time-resolved imaging compression scheme for a higher imaging speed. First, we explain the data redundancy in time-resolved imaging and introduce the center of the mass method (CMM) as a compression scheme. Then, we discuss the optimal implementation level of the compression scheme. Finally, we propose a compressive analog-to-digital converter design.

A.3.1 Data Compression Scheme

The data redundancy in the time-resolved imaging is large because most time-resolved imagers generate multiple intensity images instead of a single time image. Therefore, the processing of intensity images on-chip for a single time image will greatly reduce the data rate. The center of mass method (CMM) is a good candidate for the processing for both the time-of-flight and the fluorescence lifetime imaging. CMM estimates the average photon arrival times (τ) as the following equation:

$$\tau = \frac{\Sigma\, n \cdot f[n]}{\Sigma\, f[n]} \tag{A.1}$$

where f[n] is the n-th intensity image.

The processing or the data compression scheme can be implemented in either analog or digital domain as shown in Fig. A-8. The analog processing, in Fig. A-8(a), is implemented before the ADC and reduces the data by N times, relieving the requirement for both ADC and I/O; however, the analog processing is not only difficult to implement but also vulnerable to noise and process variations. On the other hand, digital processing shown in Fig. A-8(b) provides higher accuracy but the ADC must handle N times more data, limiting the overall imaging speed.

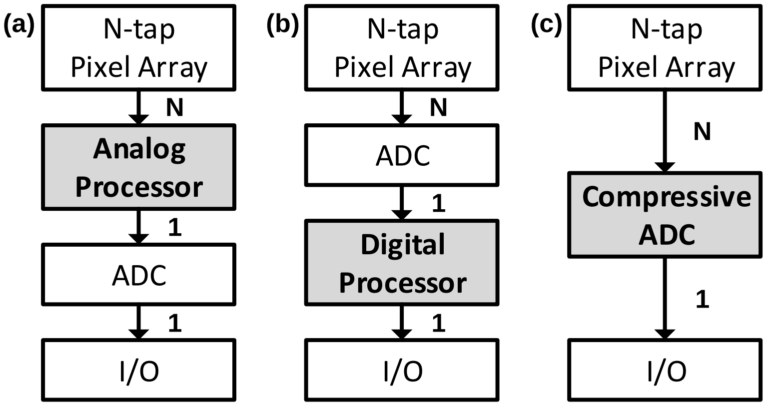


Fig. A-8. Implementation level of the data compression scheme (a) analog domain (b) digital domain (c) mixed-signal domain

To take advantages of both analog and digital approaches, we propose a mixed-mode implementation resulting in a compressive ADC. A two-step coarse-fine ADC is proposed for both the data processing and fast A/D conversion. The coarse ADC is implemented as a delta-sigma modulator which performs the multiplying integration in (A.1) and the fine ADC is implemented as a SAR ADC for high conversion speed.

The delta-sigma modulator can integrate any sequence of the input signals. It can calculate the numerator and the denominator of the equation (A.1). The denominator $\Sigma\, f[n]$ is calculated by adding the input signals sequentially. The numerator $\Sigma\, n \cdot f[n]$ seems more complex due to the multiplication but it can be easily implemented using a multiple-sampling technique; i.e. the term $n \cdot f[n]$ is calculated by adding $f[n]$ for n times. Therefore, for N-tap pixel, the number of integration cycles N_{DS} for the numerator is given as:

$$N_{DS} = \frac{N(N+1)}{2}. \tag{A.2}$$

This integration in the delta-sigma modulation also quantizes the integrated value in the resolution of $log_2 N_{DS}$ bits. The digital output of this modulator is not sufficient in most applications (e.g. 7 bits for N=16 and 5 bits for N=8); therefore, the residue of this ADC is further quantized by the fine SAR ADC for additional bits.

A.3.2 Compression Ratio: Optimizing ADC and I/O Speed

The CMM processes N images and produces 1 image, giving N:1 compression ratio. This also reduces the number of AD conversions by N, but each AD conversion time will increase due to the increased number of integration cycles according to (A.2). Therefore, there exists an optimal compression ratio that will maximize the overall speed of the imaging.

Compression ratio can be adjusted by dividing the CMM equation shown in (A.1). To illustrate this, we consider the CMM equation with N=16. The numerator (NUM) and the denominator (DEN) can then be expressed as

$$\mathrm{NUM} = \mathrm{f}(1..16) \cdot [1..16]'$$

$$\mathrm{DEN} = \mathrm{sum}\big(\mathrm{f}(1..16)\big)$$

At N=16, the number of cycles for the conversion is 136. If the ADC time is longer than the I/O time, we can reduce the ADC time by dividing the formula into following expressions:

$$\mathrm{NUM}_1 = \mathrm{f}(1..8) \cdot [1..8]', \qquad \mathrm{NUM}_2 = \mathrm{f}(9..16) \cdot [1..8]'$$

$$\mathrm{DEN}_1 = \mathrm{sum}\big(\mathrm{f}(1..8)\big), \qquad \mathrm{DEN}_2 = \mathrm{sum}\big(\mathrm{f}(9..16)\big)$$

$$\mathrm{NUM} = \mathrm{NUM}_1 + \mathrm{NUM}_2 + 8 \cdot \mathrm{DEN}_2$$

$$\mathrm{DEN} = \mathrm{DEN}_1 + \mathrm{DEN}_2$$

In this case, the number of cycles for the conversion is reduced to 36. As a result, the each ADC time is reduced by approximately 3.8 and the data size increases by 2. Therefore, the overall ADC time is reduced by ~2 (3.8/2), whereas the I/O time increases by 2. Therefore, in general, the optimal row time can be obtained from the IO time and ADC time as $\sqrt{\mathrm{T_{IO}} \cdot \mathrm{T_{ADC}}}$. Using this optimal row time, the optimal number of division can be chosen.

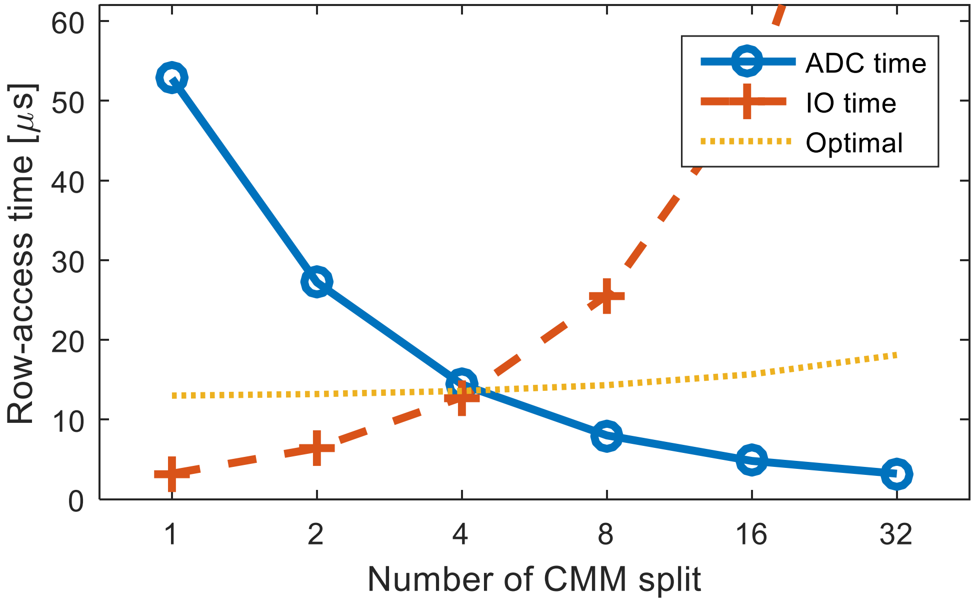


Fig. A-9. Row-access time of ADC and I/O with respect to the number of CMM splits

The effect of CMM splits is illustrated in Fig. A-9. In this example, it is assumed that the pixel has 32 taps, a row has 256 pixels, and ADC and I/O are clocked at 10MHz and 80MHz, respectively. If we choose not to use the compression at all (division of 32 in the figure), the overall speed is limited by the I/O time. The other extreme case is to directly implement the CMM equation (division of 1 in the figure), the speed is limited by the ADC time. The optimal point is found at the division of 4 in this particular case and the overall imaging speed can be improved by 8 compared to the case of no compression.

A.4 Prototype Sensor Design

Fig. A-10 shows the architecture of the multi-tap pixel sensor. This prototype has a pixel array consisting of 256 pixels in a row. The pixel array has 256×16 analog signals because each pixel has 16-taps. Two ADC arrays are laid out above and below the pixel array. 128 ADCs above the pixel array are dedicated to the pixels in even columns (i.e. 0,2,4,…,254), while the rest 128 ADCs below are for the pixels in odd columns (i.e. 1,3,5,…,255). This configuration gives an enough pitch for column-parallel ADC layout, which is twice the pixel pitch. The digital outputs from the ADCs are first latched in a digital memory. After each A/D conversion, the outputs stored in the memory are sequentially accessed by the column scanner and then sent outside via the sense amplifier and a 2:1 multiplexer. After that, the tap scanner selects the next tap and this process is repeated until the pixel values in all taps are read out.

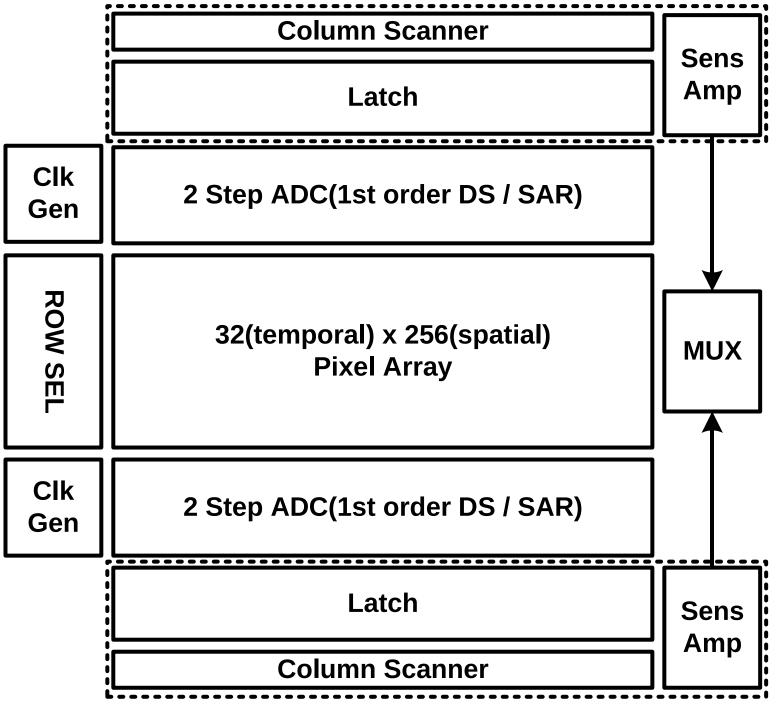


Fig. A-10. Architecture of the prototype sensor with proposed multi-tap pixel.

A.4.1 Pixel Design

Fig. A-11 shows the layout and the schematic of a multi-tap pixel for the prototype chip. Each pixel contains 16 taps, each of which has four transistors for active sensing:(1) the transfer gate (TX) transfers electrons from the channel to the floating diffusion (FD), (2) the reset transistor (RX) resets the initial voltage of the tap, (3) the source follower (DX) is used to amplify the signal, and finally (4) the select transistor (SX) is used to address the tap.

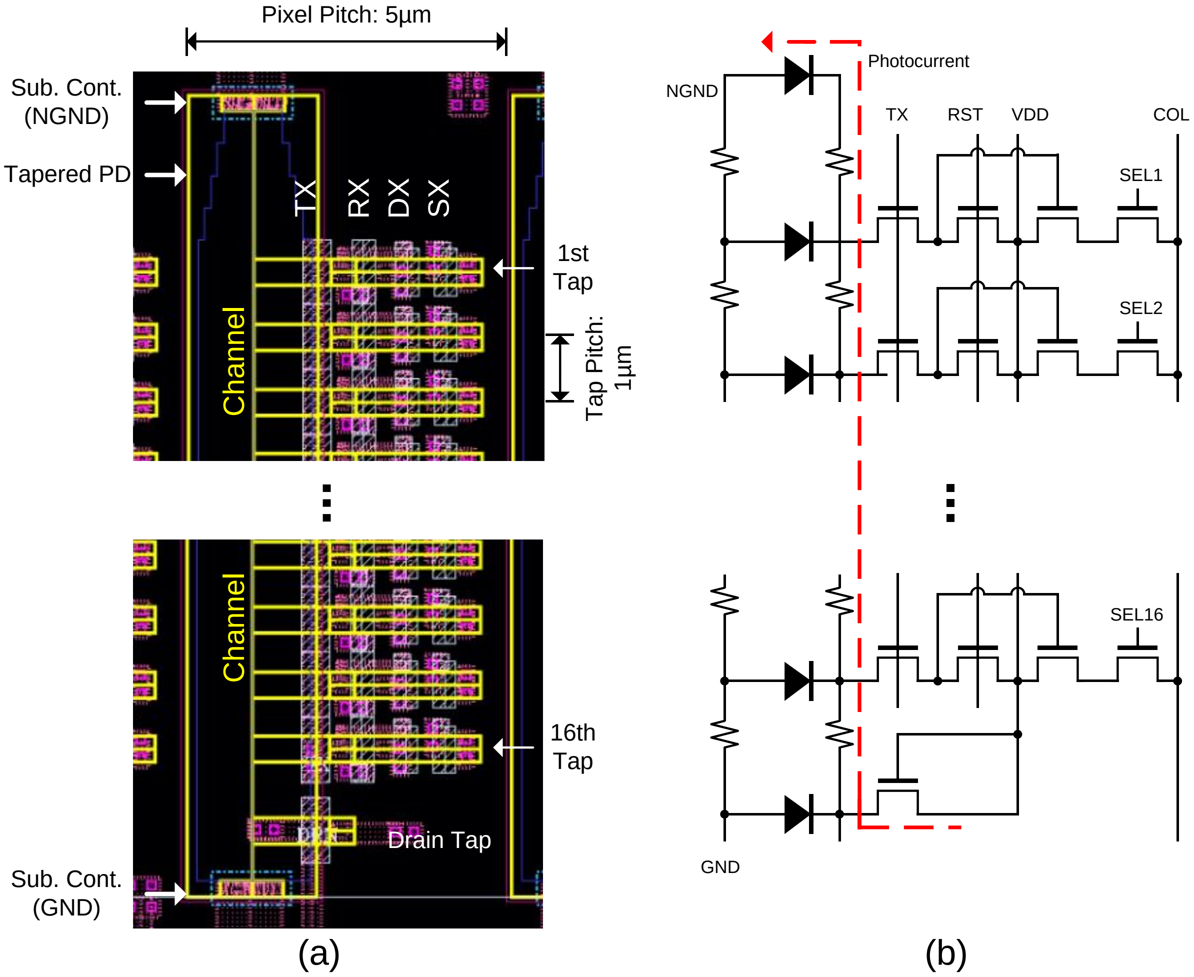


Fig. A-11. (a) Layout and (b) schematic of a multi-tap pixel for the prototype.

The photodiode used in this layout is a long vertical pinned-photodiode, which, in the used CMOS image sensor process, is composed of a substrate and two photodiode doping layers (i.e. PDN and PDP). Only the top part (tapered PD) of the region is exposed to the light and the rest of the pixel region (channel) is covered by a shield metal. The PDN layer in the PD region is tapered in order to generate a built-in lateral electric field which helps to accelerate the movement of charges from PD to the channel. There are two p-substrate contact regions, at the top and the bottom of each pixel, respectively. These substrate contacts are used to generate a linearly-modulated potential in the channel region by flowing a small static current along the channel.

The pixel operates in the following sequence. First, a global reset signal (RST) is applied to the entire pixel array. It sets the initial voltage of the FD nodes before photocurrent integration. After the reset, the voltage of each FD is readout. Then the sensor start measuring the light signals: light emitter illuminates the object and the reflected or fluorescence light is returned to the sensor depending on its application. The FD levels are readout again to perform correlated double sampling (CDS).

A.4.2 ADC Design

There two major requirements for the proposed compressive ADC. First it should be able to perform a weighted summation for CMM calculation. Second, it should provide a faster conversion speed than the single-slope ADC, which is conventional in state-of-the-art CMOS image sensors.

Two meet both requirements, we propose a two-step ADC for the prototype chip. For the first step conversion, a delta-sigma ADC is chosen to realize the weighted summation. Since the maximum gain of the weight is 8, the required number of integration cycles is 36. However, 36 cycles of delta-sigma ADC offer only 5-bit quantization step. In order to increase the resolution to 10~11 bits, the ADC has to run extra ~1000 steps. To provide 5 additional bit without increasing the ADC time, we first sample the residue signal after the delta-sigma ADC provides 5-bit coarse data. This residue is then resolved further by the second ADC for fine resolution. We chose SAR ADC as the second step ADC due to its low power consumption and high speed operation.

The block diagram of this two-step ADC is shown in Fig. A-12. The integrator and the first loop with the DAC forms the first order delta-sigma modulator. The output of the delta-sigma modulator is fed to the 6-bit counter for decimation. After the delta-sigma modulation, the residue is sampled onto the capacitor. The capacitor is now isolated from the integrator breaking the delta-sigma loop. The capacitor, quantizer and the successive approximation register (SAR) with the second DAC forms the SAR ADC. This SAR operation provides additional 5-bit for the fine conversion.

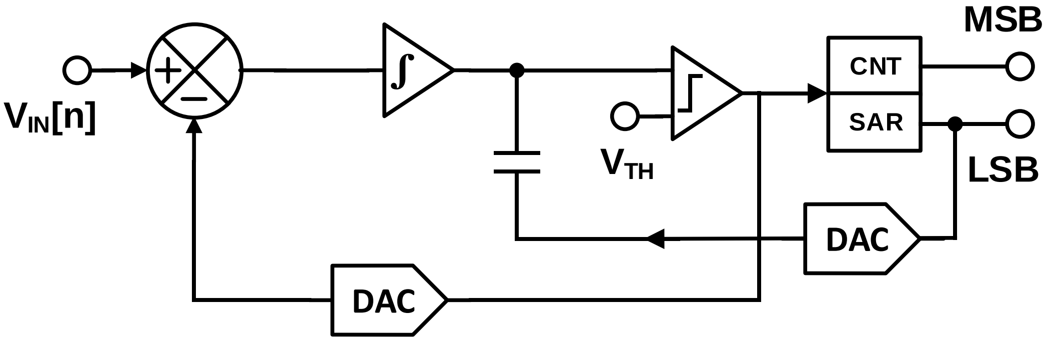


Fig. A-12. Block diagram of proposed two step ADC.

Fig. A-13 shows the first-order delta modulator, which is composed of switched capacitor op-amps. In this circuit, the capacitor, C_{OS}, and two switches between C_{OS} and C2 remove the OPAMP input offset as well as the low frequency noise. Capacitors C1, C2 and the rest of the switches form a switched capacitor integrator. The integrator uses two non-overlapping clocks, P1 and P2. In the first phase, when P1 is on, the input voltage VIN is sampled in the input capacitor C1. In the next phase, VB and VDAC are subtracted. As a result, the voltage equal to (VIN – VB – VDAC) is integrated onto C2. The voltage VIN – VDAC corresponds to the summation node in Fig. A-13. VB term is added to shift the common mode of the pixel output to the ADC's input common mode. The integrator's output is then compared to a reference voltage to perform 1-bit quantization at the rising edge of P1. Based on the comparator's output, VDAC is chosen from either the positive or the negative reference to be subtracted during the next integration phase, P2.

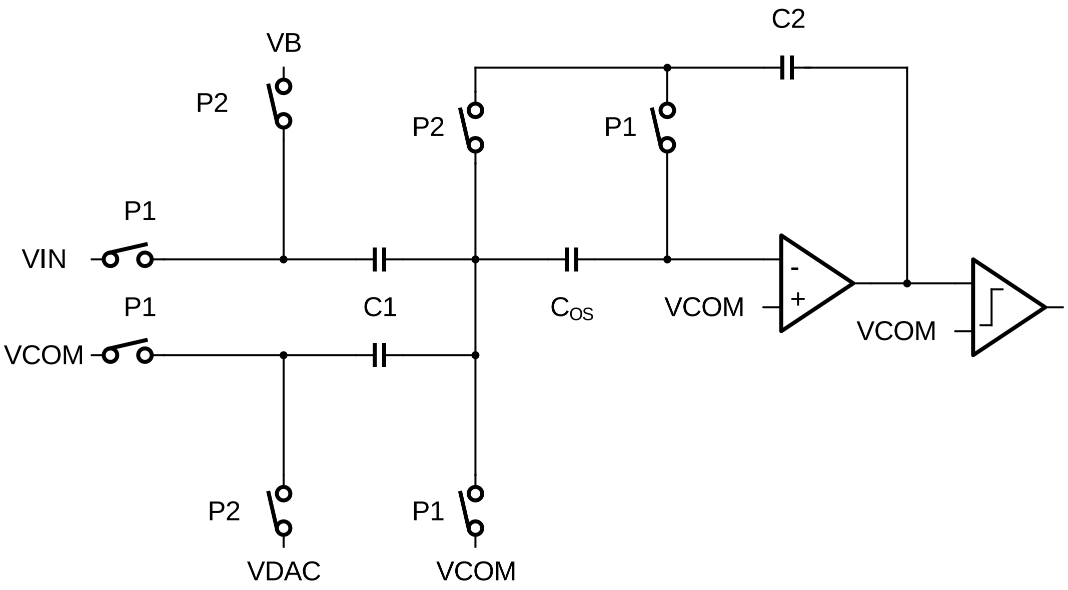


Fig. A-13. Schematic of the first order delta-sigma modulator.

Fig. A-14 shows the schematic of the proposed SAR ADC. The residual voltage VRES from the sigma-delta modulation is first sampled onto the entire capacitors. During this sampling phase, the bottom plates of all capacitors are connected to the common mode voltage, VCOM, which is the average of the positive and negative references, VH and VL. After the sampling, the voltage of the top plate is quantized by a comparator, generating the most significant bit (MSB). Based on the comparator value, either VH or VL is connected to the capacitor corresponding to each bit, 8C for MSB and C for LSB. This process is repeated for five cycles, generating 5-bit conversion in total.

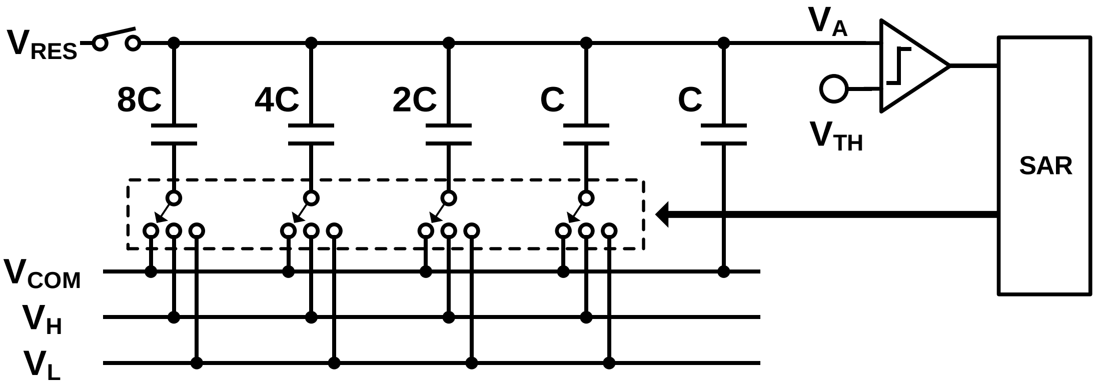


Fig. A-14. Schematic of the SAR ADC.

A.5 Experimental Results

In order to characterize the time quantization step in multi-tap pixels, the charge velocity in the channel region has been measured because the charge velocity in the channel is directly related to the time resolution as described in the section A.2.1.

Fig. A-15 shows the test setup to measure the charge speed in the channel. The photo-generated charges travel from the photodiode toward the distal end during the delay time between the LED and TX pulses as shown in Fig. A-15(b). The average travel length of the charges is proportional to the delay as shown in Fig. A-15(c). The charge velocity is defined as the travel length divided by the delay; therefore, a loop-back measurement with varying delays (step of 363ps) is performed to measure the charge velocity.

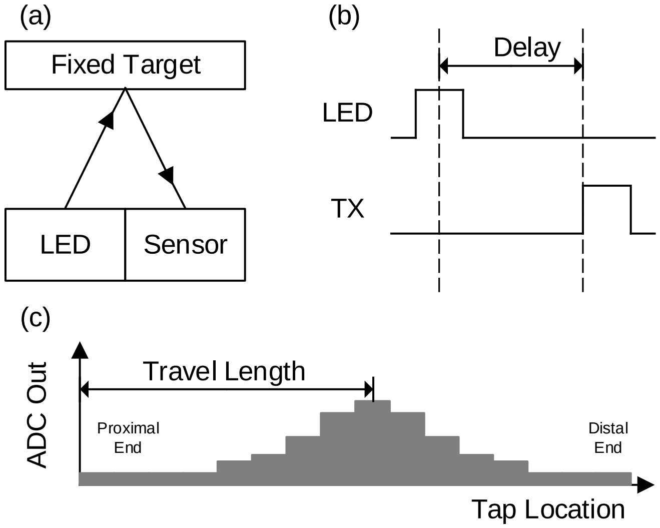


Fig. A-15. Loop-back test for charge velocity measurement; (a) loop-back configuration (b) timing diagram (c) example of measured signals

Fig. A-16 shows the measured images from the loop-back test with 0.5V bias across the substrate contacts. X-axis is the pixel location (from 0 to 255) and y-axis represents the tap location of each pixel. The images are acquired at delay values ranging from 0 to

11.62ns with 363ps step. In Fig. A-16, only 5 subsampled images with 2.9ns step are shown. The charge movement is visualized in each image (a snapshot of the photo-electrons at each time).

Charge velocity can be calculated from these snapshots; however, two major issues prevent the precise measurement acquisition. First, the image suffers from a large fixed-pattern noise even after CDS. This suggests that there may be potential wells in the middle of the channel where the charges are captured. Those captured charges are continuously integrated in the vicinity integration node, contributing to the large fixed-pattern noise.

Second, no significant charge movement is observed after around the 8th tap. For example, the charge movement is not clear after the 3rd image. This suggests that there is no lateral electric field around that location or there might be a potential barrier. The root cause is still unknown, and process and device simulation studies may reveal the actual reason in the future. As a result, the calculated charge velocity is constant only until the charges reach the 8th tap. After that, the charge velocity rapidly drops toward zero.

To mitigate these issues, only a subset of the first images was used to calculate the speed. For example, the images with delays less than 5.08ns were used in this example. In addition to that, the tap values from 256 pixels are averaged in order to reduce the FPN.

Fig. A-17 shows the measured charge velocity at different substrate bias conditions. As expected, the larger the bias (or electric field), the faster the charge speed. However, the measured charge velocity is approximately 7 times lower than the expected value. There are 3 possible explanations for this problem.

1) The resistance of the channel is too low compared to the contact and wire resistance. In this case, only a little portion of the bias is applied to the channel area.
2) The resistance is non-uniform across the channel, leading to a non-uniform electric field. If the electric field is concentrated at a short distance, then the other channel regions may not have sufficient electric field to accelerate the electrons.
3) The channel region after pixel reset is not fully depleted. If there are remaining electrons in the channel, they will flat out the electric potential as well as the electric field in the channel.

To identify the root cause, a future simulation study and device investigation is required with the actual process and device parameters. Since this requires an active engagement of the engineers from the foundry, we could not fully identify the root cause. We suggest continuing this research in the future with the support of the foundry company.

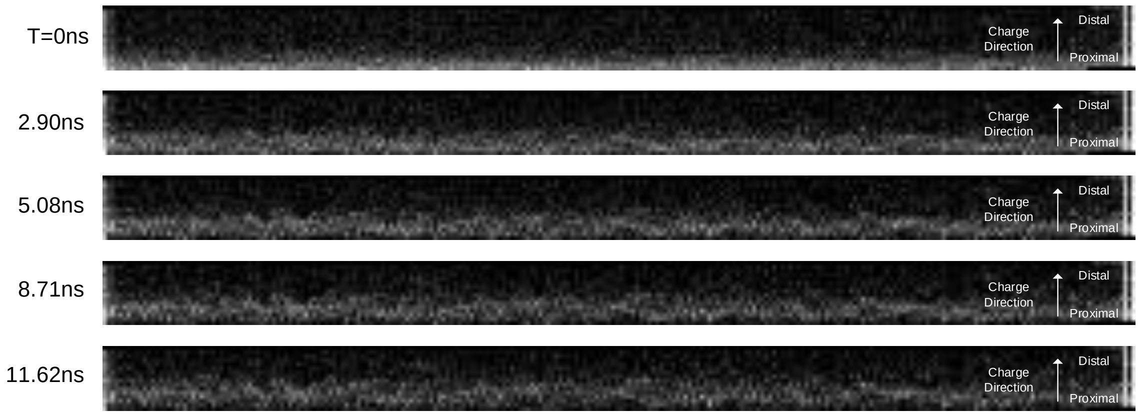


Fig. A-16: Loop-Back Test: Measured Images. (0.5V across the channel region)

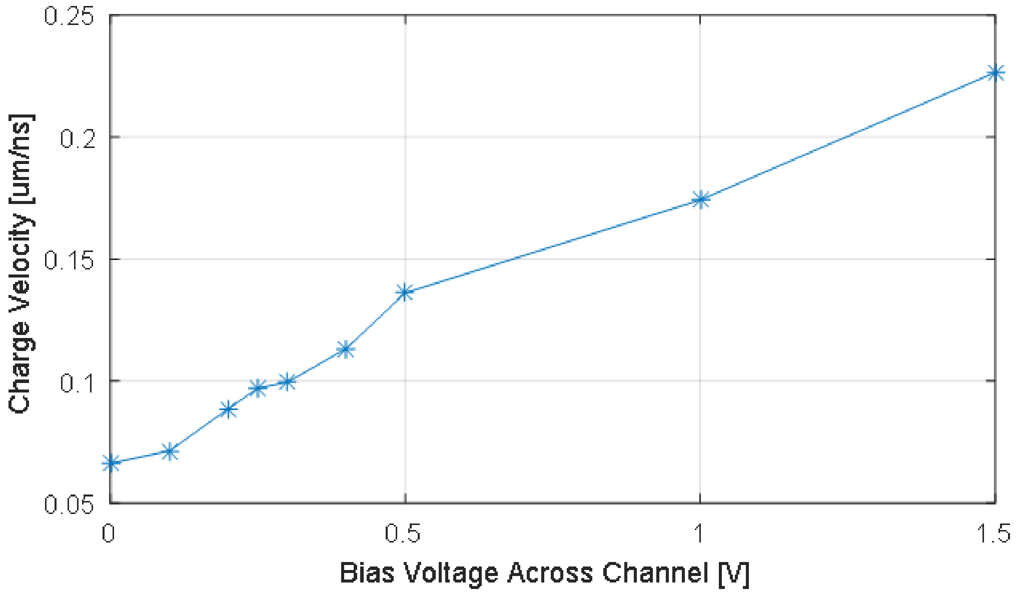


Fig. A-17: Charge Velocity vs. Bias Voltage across the Channel (or electric field).

A.6 Summary

In this chapter, we proposed the multi-tap pixel based on time-to-space conversion. It enables fast photon acquisition as well as high signal-to-noise ratio especially under low-light illumination. We also proposed the compressive ADC to reduce the data size to increase the overall imaging speed by balancing ADC and I/O speed. The proposed pixel is simulated using Sentaurus, and fabricated in CMOS image sensor technology. The experimental results showed the time-to-space conversion gain of $0.07\sim0.225\mu m/ns$ with the substrate bias of 0~1.5V. Unfortunately, the measured results underperform compared to the expected values. We have come up with causes of this problem, but none of them can be confirmed without acquiring confidential information from the foundry. We suggest continuing this research in the future with the support of foundry.

BIBLIOGRAPHY

[1] L. Marcu, P. M. W. French, and D. S. Elson, Fluorescence Lifetime Spectroscopy and Imaging: Principles and Applications in Biomedical Diagnostics. CRC Press, 2014.

[2] M. Y. Berezin and S. Achilefu, “Fluorescence lifetime measurements and biological imaging.,” Chem. Rev., vol. 110, no. 5, pp. 2641–84, May 2010.

[3] K. L. Diehl and E. V Anslyn, “Array sensing using optical methods for detection of chemical and biological hazards.,” Chem. Soc. Rev., vol. 42, no. 22, pp. 8596–611, Dec. 2013.

[4] C.-W. Chang, D. Sud, and M.-A. Mycek, “Fluorescence lifetime imaging microscopy.,” Methods Cell Biol., vol. 81, no. null, pp. 495–524, Jan. 2007.

[5] P. J. Tadrous, J. Siegel, P. M. W. French, S. Shousha, E.-N. Lalani, and G. W. H. Stamp, “Fluorescence lifetime imaging of unstained tissues: early results in human breast cancer.,” J. Pathol., vol. 199, no. 3, pp. 309–17, Mar. 2003.

[6] H. Fatakdawala et al., “Multimodal in vivo imaging of oral cancer using fluorescence lifetime, photoacoustic and ultrasound techniques.,” Biomed. Opt. Express, vol. 4, no. 9, pp. 1724–41, Jan. 2013.

[7] V. Masilamani et al., “Optical biopsy of benign and malignant tissue by time resolved spectroscopy.,” Technol. Cancer Res. Treat., vol. 12, no. 6, pp. 559–63, Dec. 2013.

[8] S. Keereweer et al., “Optical image-guided cancer surgery: challenges and limitations.,” Clin. Cancer Res., vol. 19, no. 14, pp. 3745–54, Jul. 2013.

[9] J. McGinty et al., “Wide-field fluorescence lifetime imaging of cancer.,” Biomed. Opt. Express, vol. 1, no. 2, pp. 627–640, Jan. 2010.

[10] A. L. Vahrmeijer, M. Hutteman, J. R. van der Vorst, C. J. H. van de Velde, and J. V Frangioni, “Image-guided cancer surgery using near-infrared fluorescence.,” Nat. Rev. Clin. Oncol., vol. 10, no. 9, pp. 507–18, Sep. 2013.

[11] M. Hansard, S. Lee, O. Choi, and R. Horaud, Time-of-Flight Cameras : Principles , Methods and Applications. Springer, 2012.

[12] C. C. Enz and G. C. Temes, “Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization,” Proc. IEEE, vol. 84, no. 11, pp. 1584–1614, 1996.

[13] B. Fowler, M. D. Godfrey, and S. Mims, “Reset noise reduction in capacitive sensors,” IEEE Trans. Circuits Syst. I Regul. Pap., vol. 53, no. 8, pp. 1658–1669, Aug. 2006.

[14] E. R. Fossum, “CMOS image sensors: electronic camera on a chip,” in Proceedings of International Electron Devices Meeting, pp. 17–25.

[15] J. Choi, S. Park, J. Cho, and E. Yoon, “A 1.36μW adaptive CMOS image sensor with reconfigurable modes of operation from available energy/illumination for distributed wireless sensor network,” in 2012 IEEE International Solid-State Circuits Conference, 2012, pp. 112–114.

[16] M. F. Snoeij, A. J. P. Theuwissen, and J. H. Huijsing, “A 1.8V 3.2μW Comparator for Use in a CMOS Imager Column-Level Single-Slope ADC,” in 2005 IEEE International Symposium on Circuits and Systems, 2005, pp. 6162–6165.

[17] M. F. Snoeij, A. J. P. Theuwissen, K. A. A. Makinwa, and J. H. Huijsing, “A CMOS Imager With Column-Level ADC Using Dynamic Column Fixed-Pattern Noise Reduction,” IEEE J. Solid-State Circuits, vol. 41, no. 12, pp. 3007–3015, Dec. 2006.

[18] S. Yoshihara et al., “A 1/1.8-inch 6.4 MPixel 60 frames/s CMOS Image Sensor With Seamless Mode Change,” IEEE J. Solid-State Circuits, vol. 41, no. 12, pp. 2998–3006, Dec. 2006.

[19] M. Mase, S. Kawahito, M. Sasaki, and Y. Wakamori, “A 19.5b dynamic range CMOS image sensor with 12b column-parallel cyclic A/D converters,” in 2005 IEEE Solid-State Circuits Conference, 2005, pp. 350–352.

[20] J.-H. Park, S. Aoyama, T. Watanabe, K. Isobe, and S. Kawahito, “A High-Speed Low-Noise CMOS Image Sensor With 13-b Column-Parallel Single-Ended Cyclic ADCs,” IEEE Trans. Electron Devices, vol. 56, no. 11, pp. 2414–2422, Nov. 2009.

[21] S. Lim et al., “A 240-frames/s 2.1-Mpixel CMOS Image Sensor With Column-Shared Cyclic ADCs,” IEEE J. Solid-State Circuits, vol. 46, no. 9, pp. 2073–2083, Sep. 2011.

[22] Y. Chae et al., “A 2.1 M Pixels, 120 Frame/s CMOS Image Sensor With Column-Parallel ΔΣ ADC Architecture,” IEEE J. Solid-State Circuits, vol. 46, no. 1, pp. 236–247, Jan. 2011.

[23] H. T. D.X.D Yang, A.E. Gamal, B. Fowler, “A 640x512 CMOS image sensor with ultrawide dynamic range floating-point pixel-level ADC,” IEEE Journal of Solid State Circuits, 1999. [Online]..

[24] D. Kim, J. Cho, S. Lim, D. Lee, and G. Han, “A 5000S/s Single-Chip Smart Eye-Tracking Sensor,” in 2008 IEEE International Solid-State Circuits Conference - Digest of Technical Papers, 2008, pp. 46–594.

[25] S. J. Carey and P. Dudek, “Vision chip with high accuracy analog S2I cells,” in 2014 14th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA), 2014, pp. 1–2.

[26] D. Kim, Y. Chae, J. Cho, and G. Han, “A Dual-Capture Wide Dynamic Range CMOS Image Sensor Using Floating-Diffusion Capacitor,” IEEE Trans. Electron Devices, vol. 55, no. 10, pp. 2590–2594, Oct. 2008.

[27] A. Boukhayma, A. Peizerat, and C. Enz, “A Sub-0.5 Electron Read Noise VGA Image Sensor in a Standard CMOS Process,” IEEE J. Solid-State Circuits, vol. 51, no. 9, pp. 2180–2191, Sep. 2016.

[28] C. Lotto, P. Seitz, and T. Baechler, “A sub-electron readout noise CMOS image sensor with pixel-level open-loop voltage amplification,” in 2011 IEEE International Solid-State Circuits Conference, 2011, pp. 402–404.

[29] S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts, “A logarithmic response CMOS image sensor with on-chip calibration,” IEEE J. Solid-State Circuits, vol. 35, no. 8, pp. 1146–1152, Aug. 2000.

[30] S.-W. Han, S.-J. Kim, J. Choi, C.-K. Kim, and E. Yoon, “A High Dynamic Range CMOS Image Sensor with In-Pixel Floating-Node Analog Memory for Pixel Level Integration Time Control,” in Symposium on VLSI Circuits, Digest of Technical Papers., 2006, pp. 25–26.

[31] T. Hamamoto and K. Aizawa, “A computational image sensor with adaptive pixel-based integration time,” IEEE J. Solid-State Circuits, vol. 36, no. 4, pp. 580–585, Apr. 2001.

[32] D. Stoppa, A. Simoni, L. Gonzo, M. Gottardi, and G.-F. Dalla Betta, “Novel CMOS image sensor with a 132-dB dynamic range,” IEEE J. Solid-State Circuits, vol. 37, no. 12, pp. 1846–1852, Dec. 2002.

[33] D. Stoppa, M. Vatteroni, D. Covi, A. Baschirotto, A. Sartori, and A. Simoni, “A 120-dB Dynamic Range CMOS Image Sensor With Programmable Power Responsivity,” IEEE J. Solid-State Circuits, vol. 42, no. 7, pp. 1555–1563, Jul. 2007.

[34] S.-J. Kim, B. Kang, J. D. K. Kim, K. Lee, C.-Y. Kim, and K. Kim, “A 1920×1080 3.65μm-pixel 2D/3D image sensor with split and binning pixel structure in 0.11μm standard CMOS,” in 2012 IEEE International Solid-State Circuits Conference, 2012, pp. 396–398.

[35] K. Seong-Jin, J. D. K. Kim, H. Sang-Wook, K. Byongmin, L. Keechang, and K. Chang-Yeong, “A 640×480 Image Sensor with Unified Pixel Architecture for 2D/3D Imaging in 0.11μm CMOS,” in VLSI Circuits (VLSIC), 2011 Symposium on, 2011, pp. 92–93.

[36] S.-J. Kim, S.-W. Han, B. Kang, K. Lee, J. D. K. Kim, and C.-Y. Kim, “A Three-Dimensional Time-of-Flight CMOS Image Sensor With Pinned-Photodiode Pixel Structure,” IEEE Electron Device Lett., vol. 31, no. 11, pp. 1272–1274, Nov. 2010.

[37] S.-J. Kim, J. D. K. Kim, B. Kang, and K. Lee, “A CMOS Image Sensor Based on Unified Pixel Architecture With Time-Division Multiplexing Scheme for Color and Depth Image Acquisition,” IEEE J. Solid-State Circuits, vol. 47, no. 11, pp. 2834–2845, 2012.

[38] J. Cho et al., “A 3-D Camera With Adaptable Background Light Suppression Using Pixel-Binning and Super-Resolution,” IEEE J. Solid-State Circuits, vol. 49, no. 10, pp. 2319–2332, Oct. 2014.

[39] J. Cho et al., “A 5.9µm-Pixel 2D/3D Image Sensor with Background Suppression over 100klx,” in Symposium on VLSI Circuits, Digest of Technical Papers, 2013, pp. 6–7.

[40] K. Yasutomi, S. Itoh, and S. Kawahito, “A Two-Stage Charge Transfer Active Pixel CMOS Image Sensor With Low-Noise Global Shuttering and a Dual-Shuttering Mode,” IEEE Trans. Electron Devices, vol. 58, no. 3, pp. 740–747, Mar. 2011.

[41] 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 Sens. J., vol. 7, no. 12, pp. 1578–1586, Dec. 2007.

[42] D. Stoppa, N. Massari, L. Pancheri, M. Malfatti, M. Perenzoni, and L. Gonzo, “An 80×60 range image sensor based on 10µm 50MHz lock-in pixels in 0.18µm CMOS,” in 2010 IEEE International Solid-State Circuits Conference, Digest of Technical Papers, 2010, pp. 406–407.

[43] H. Tian, “NOISE ANALYSIS IN CMOS IMAGE SENSORS,” Stanford University, 2000.

[44] J. Philip and K. Carlsson, “Theoretical investigation of the signal-to-noise ratio in fluorescence lifetime imaging,” J. Opt. Soc. Am. A, vol. 20, no. 2, p. 368, Feb. 2003.

[45] K. Yasutomi, T. Usui, S.-M. Han, T. Takasawa, K. Kagawa, and S. Kawahito, “7.5 A 0.3mm-resolution Time-of-Flight CMOS range imager with column-gating clock-skew calibration,” in 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014, pp. 132–133.

[46] D.-U. Li, B. Rae, R. Andrews, J. Arlt, and R. Henderson, “Hardware implementation algorithm and error analysis of high-speed fluorescence lifetime sensing systems using center-of-mass method.,” J. Biomed. Opt., vol. 15, no. 1, p. 17006, Jan. 2010.

[47] Y. Sun et al., “Fluorescence lifetime imaging microscopy for brain tumor image-guided surgery.,” J. Biomed. Opt., vol. 15, no. 5, p. 56022, Jan. 2010.

[48] Y. Sun et al., “Fluorescence lifetime imaging microscopy: in vivo application to diagnosis of oral carcinoma,” Opt. Lett., vol. 34, no. 13, p. 2081, Jun. 2009.

[49] R. M. Field, S. Realov, and K. L. Shepard, “A 100 fps, Time-Correlated Single-Photon-Counting-Based Fluorescence-Lifetime Imager in 130 nm CMOS,” IEEE J. Solid-State Circuits, vol. 49, no. 4, pp. 867–880, Apr. 2014.

[50] M. Perenzoni, N. Massari, D. Perenzoni, L. Gasparini, and D. Stoppa, “11.3 A 160×120-pixel analog-counting single-photon imager with Sub-ns time-gating and self-referenced column-parallel A/D conversion for fluorescence lifetime imaging,” in 2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers, 2015, pp. 1–3.

[51] R. M. Field and K. L. Shepard, “A 100-fps fluorescence lifetime imager in standard 0.13-μm CMOS.” pp. C10–C11, 2013.

[52] M.-W. Seo et al., “A 10.8ps-time-resolution 256×512 image sensor with 2-Tap true-CDS lock-in pixels for fluorescence lifetime imaging,” in 2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers, 2015, pp. 1–3.

[53] D. D.-U. Li et al., “Time-domain fluorescence lifetime imaging techniques suitable for solid-state imaging sensor arrays.,” Sensors (Basel)., vol. 12, no. 5, pp. 5650–69, Jan. 2012.

[54] S. Burri, F. Powolny, C. Bruschini, X. Michalet, F. Regazzoni, and E. Charbon, “A 65k pixel, 150k frames-per-second camera with global gating and micro-lenses suitable for fluorescence lifetime imaging,” in SPIE Photonics Europe, 2014, p. 914109.

[55] Z. Li et al., “A Time-Resolved CMOS Image Sensor With Draining-Only Modulation Pixels for Fluorescence Lifetime Imaging,” IEEE Trans. Electron Devices, vol. 59, no. 10, pp. 2715–2722, Oct. 2012.

[56] L.-E. Bonjour, D. Beyeler, N. Blanc, and M. Kayal, “CMOS Demodulation Image Sensor for Nanosecond Optical Waveform Analysis,” IEEE Sens. J., vol. 13, no. 5, pp. 1487–1497, May 2013.

[57] K. K. S. and, A. Periasamy*, H. Ashworth*, and J. N. D. H. Snow, “Error Analysis of the Rapid Lifetime Determination Method for Double-Exponential Decays and New Windowing Schemes,” 1999.

[58] Y. Won, S. Moon, W. Yang, D. Kim, W.-T. Han, and D. Y. Kim, “High-speed confocal fluorescence lifetime imaging microscopy (FLIM) with the analog mean delay (AMD) method,” Opt. Express, vol. 19, no. 4, p. 3396, Feb. 2011.

[59] S. Moon, Y. Won, and D. Y. Kim, “Analog mean-delay method for high-speed fluorescence lifetime measurement,” Opt. Express, vol. 17, no. 4, p. 2834, Feb. 2009.

[60] D. Tyndall, B. Rae, D. Li, J. Richardson, J. Arlt, and R. Henderson, “A 100Mphoton/s time-resolved mini-silicon photomultiplier with on-chip fluorescence

lifetime estimation in 0.13µm CMOS imaging technology," in 2012 IEEE International Solid-State Circuits Conference, 2012, pp. 122–124.

[61] Y. J. Won, W.-T. Han, and D. Y. Kim, "Precision and accuracy of the analog mean-delay method for high-speed fluorescence lifetime measurement," J. Opt. Soc. Am. A, vol. 28, no. 10, p. 2026, Sep. 2011.

[62] L. Pancheri, N. Massari, and D. Stoppa, "SPAD Image Sensor With Analog Counting Pixel for Time-Resolved Fluorescence Detection," IEEE Trans. Electron Devices, vol. 60, no. 10, pp. 3442–3449, Oct. 2013.

[63] V. Ganapathi, C. Plagemann, D. Koller, and S. Thrun, "Real time motion capture using a single time-of-flight camera," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010, pp. 755–762.

[64] M. Van den Bergh and L. Van Gool, "Combining RGB and ToF cameras for real-time 3D hand gesture interaction," in 2011 IEEE Workshop on Applications of Computer Vision (WACV), 2011, pp. 66–72.

[65] Y. Oike, M. Ikeda, and K. Asada, "A 120×110 Position Sensor With the Capability of Sensitive and Selective Light Detection in Wide Dynamic Range for Robust Active Range Finding," IEEE J. Solid-State Circuits, vol. 39, no. 1, pp. 246–251, Jan. 2004.

[66] Y. Oike, M. Ikeda, and K. Asada, "Design and Implementation of Real-Time 3-D Image Sensor With 640×480 Pixel Resolution," IEEE J. Solid-State Circuits, vol. 39, no. 4, pp. 622–628, Apr. 2004.

[67] Y. Oike, M. Ikeda, and K. Asada, "A 375×365 high-speed 3-D range-finding image sensor using row-parallel search architecture and multisampling technique," IEEE J. Solid-State Circuits, vol. 40, no. 2, pp. 444–453, Feb. 2005.

[68] 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, vol. 40, no. 9, pp. 1847–1854, Sep. 2005.

[69] C. Niclass, C. Favi, T. Kluter, M. Gersbach, and E. Charbon, "A 128×128 Single-Photon Imager with on-Chip Column-Level 10b Time-to-Digital Converter Array Capable of 97ps Resolution," in IEEE International Solid-State Circuits Conference, Digest of Technical Papers, 2008, pp. 44–594.

[70] F. Blais, "Review of 20 years of range sensor development," in Electronic Imaging, 2003, pp. 62–76.

[71] S. Koyama, K. Onozawa, K. Tanaka, and Y. Kato, "A 3D vision 2.1Mpixel image sensor for single-lens camera systems," in IEEE International Solid-State Circuits Conference, Digest of Technical Papers, 2013, pp. 492–493.

[72] T. Oggier et al., “Novel pixel architecture with inherent background suppression for 3D time-of-flight imaging,” in Proceedings of SPIE, 2005, vol. 5665, no. 1, pp. 1–8.

[73] B. Büttgen and P. Seitz, “Robust Optical Time-of-Flight Range Imaging Based on Smart Pixel Structures,” IEEE Trans. Circuits Syst. I Regul. Pap., vol. 55, no. 6, pp. 1512–1525, Jul. 2008.

[74] G. Zach and H. Zimmermann, “A 2×32 range-finding sensor array with pixel-inherent suppression of ambient light up to 120klx,” in IEEE International Solid-State Circuits Conference, Digest of Technical Papers, 2009, p. 352–353,353a.

[75] G. Zach, M. Davidovic, and H. Zimmermann, “A 16×16 Pixel Distance Sensor With In-Pixel Circuitry That Tolerates 150 klx of Ambient Light,” IEEE J. Solid-State Circuits, vol. 45, no. 7, pp. 1345–1353, Jul. 2010.

[76] M. Davidovic, G. Zach, K. Schneider-Hornstein, and H. Zimmermann, “Range finding sensor in 90nm CMOS with bridge correlator based background light suppression,” in Proceedings of ESSCIRC, 2010, pp. 298–301.

[77] Y. M. Wang et al., “Compact Ambient Light Cancellation Design and Optimization for 3D Time-of-Flight Image Sensors,” in International Image Sensor Workshop, 2013, vol. 2, no. 1.

[78] M. Perenzoni, N. Massari, D. Stoppa, L. Pancheri, M. Malfatti, and L. Gonzo, “A 160×120-Pixels Range Camera With In-Pixel Correlated Double Sampling and Fixed-Pattern Noise Correction,” IEEE J. Solid-State Circuits, vol. 46, no. 7, pp. 1672–1681, Jul. 2011.

[79] J. Shin, B. Kang, K. Lee, and J. D. K. Kim, “A 3D image sensor with adaptable charge subtraction scheme for background light suppression,” in IS&T/SPIE Electronic Imaging, 2013, pp. 865907-865907–7.

[80] R. Lange and P. Seitz, “Solid-state time-of-flight range camera,” IEEE J. Quantum Electron., vol. 37, no. 3, pp. 390–397, Mar. 2001.

[81] T. Ringbeck, T. Möller, and B. Hagebeuker, “Multidimensional measurement by using 3-D PMD sensors,” Adv. Radio Sci., vol. 5, pp. 135–146, 2007.

[82] C. Niclass, M. Soga, H. Matsubara, S. Kato, and M. Kagami, “A 100-m Range 10-Frame/s 340×96-Pixel Time-of-Flight Depth Sensor in 0.18-μm CMOS,” IEEE J. Solid-State Circuits, vol. 48, no. 2, pp. 559–572, Feb. 2013.

[83] J. Cho, S. Park, J. Choi, and E. Yoon, “Area-efficient and low-power implementation of vision chips using multi-level mixed-mode processing,” in 2014 14th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA), 2014.

[84] J. Choi, S.-W. Han, S.-J. Kim, S.-I. Chang, and E. Yoon, “A Spatial-Temporal Multiresolution CMOS Image Sensor With Adaptive Frame Rates for Tracking the

Moving Objects in Region-of-Interest and Suppressing Motion Blur," IEEE J. Solid-State Circuits, vol. 42, no. 12, pp. 2978–2989, Dec. 2007.

[85] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," IEEE Signal Process. Mag., vol. 20, no. 3, pp. 21–36, May 2003.

[86] S. Schuon, C. Theobalt, J. Davis, and S. Thrun, "LidarBoost: Depth superresolution for ToF 3D shape scanning," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 343–350.

[87] M. Ben-Ezra, A. Zomet, and S. K. Nayar, "Video super-resolution using controlled subpixel detector shifts.," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 6, pp. 977–87, Jun. 2005.

[88] N. Ahuja, "A Refractive Camera for Acquiring Stereo and Super-resolution Images," in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, vol. 2, pp. 2316–2323.

[89] B. Razavi and B. A. Wooley, "Design techniques for high-speed, high-resolution comparators," IEEE J. Solid-State Circuits, vol. 27, no. 12, pp. 1916–1926, 1992.

[90] M. Perenzoni and D. Stoppa, "Figures of Merit for Indirect Time-of-Flight 3D Cameras: Definition and Experimental Evaluation," Remote Sens., vol. 3, no. 12, pp. 2461–2472, Nov. 2011.

[91] J. Penne et al., "Time-of-Flight 3-D endoscopy.," Med. Image Comput. Comput. Assist. Interv., vol. 12, no. Pt 1, pp. 467–74, Jan. 2009.

[92] T. Oggier et al., "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger TM)." [Online]. Available: http://www.mesa-imaging.ch/pdf/CSEM_manuscript_5249-65.pdf. [Accessed: 11-Nov-2011].

[93] T. Takasawa, K. Yasutomi, S. Aoyama, K. Kagawa, and S. Kawahito, "A Time-of-Flight Range Image Sensor With Background Canceling Lock-in Pixels Based on Lateral Electric Field Charge Modulation," IEEE J. Electron Devices Soc., vol. 3, no. 3, pp. 267–275, May 2015.

[94] J. Itatani, F. Quéré, G. Yudin, M. Ivanov, F. Krausz, and P. Corkum, "Attosecond Streak Camera," Phys. Rev. Lett., vol. 88, no. 17, p. 173903, Apr. 2002.