A dynamic vision sensor with direct logarithmic output and full-frame picture-on-demand

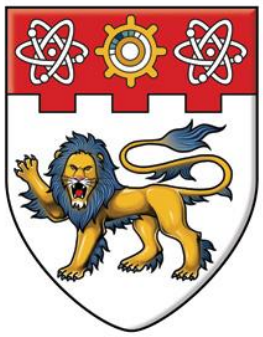
Guo, Menghan

2016

Guo, M. (2016). A dynamic vision sensor with direct logarithmic output and full-frame picture-on-demand. Master’ s thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/69258

https://doi.org/10.32657/10356/69258



NANYANG TECHNOLOGICAL UNIVERSITY

A DYNAMIC VISION SENSOR WITH DIRECT LOGARITHMIC OUTPUT AND FULL-FRAME PICTURE-ON-DEMAND

GUO MENG HAN

SCHOOL OF ELECTRICAL AND ELECTRONIC ENGINEERING

2016

A DYNAMIC VISION SENSOR

GUO MENG HAN

2016

A Dynamic Vision Sensor with Direct Logarithmic Output and Full-Frame Picture-On-Demand

Guo Meng Han

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Master of Engineering

2016

Acknowledgements

This thesis is the summary of my research work as a MENG student in school of EEE of Nanyang Technological University. Lots of people have supported and encouraged me along this way and without them all the project and this thesis would never be done favourably.

First of all, I would like to express my in depth gratitude to my supervisor, Professor Chen Shoushun, for all his invaluable guidance, insightful technical advices, continuous support, and patient encouragement throughout my study. It has been a pleasure studying in the smart sensor group and I learned a lot from him.

I would appreciate Yu Hang for his technical discussions in my research and assistances during my sensor tapeout and measurement. I would appreciate Ding Ruoxi for his strong support in testing the sensor. In addition, I would like to extend my acknowledgements to my other colleagues in smart sensor group, Dr. Zhao Bo, Dr. Qian Xinyuan, Zhang Xiangyu, Gibran Limi Jaya, Wang Yue, Huang Jing and Guo Heng for their assistances in my study and life.

I would like to thank Lim-Tan Gek Eng, Tay-Teo Boon Ping (Dorothy), David Robert Neubronner in VIRTUS and Chong-Eng Puay Cheng, Low Siew Kheng (Serene), Ng-Yap Poh Geok (Pamela), Tan Geok Lan (Janet), Tham Suet Mei in Satellite Research Centre for providing all the research resources.

Finally, I would like to express my special appreciation to my parents and some of my best friends in China and USA for their endless love and continuous support. All of this will encourage me to move on in the future life.

This page is intentionally left blank.

Contents

List of Figures

This page is intentionally left blank.

List of Tables

This page is intentionally left blank.

List of Abbreviations

AC	Alternating Current
ADC	Analog-to-Digital Convertor
AER	Address-Event-Representation
APS	Active Pixel Sensor
ATIS	Asynchronous Time-based Image Sensor
BW	Bandwidth
CCD	Charge-Coupled Device
CDS	Correlated Double Sampling
CIS	CMOS Image Sensor
CMOS	Complementary Metal-Oxide-Semiconductor
DAC	Digital-to-Analog Converter
DAVIS	Dynamic and Active Pixel Vision Sensor
DC	Direct Current
DNL	Differential Non-Linearity
DPS	Digital Pixel Sensor
DR	Dynamic Range
DVS	Dynamic Vision Sensor
FC	Folded Cascode
FOV	Field of View
FPGA	Field Programmable Gate Array
FPN	Fixed-Pattern-Noise
FSR	Full-Scale Range
GBW	Gain Bandwidth Product
GUI	Graphic User Interface
INL	Integral Non-Linearity
LHP	Left-Hand-Plane
LSB	Least Significant Bit
MSB	Most Significant Bit
OTA	Operational Transconductance Amplifier
PC	Personal Computer

PM	Phase Margin
PPS	Passive Pixel Sensor
PWM	Pulse Width Modulation
RFC	Recycling Folded Cascode
RHP	Right-Hand-Plane
RMS	Root-Mean-Square
ROI	Region of Interest
SAR	Successive Approximation Register
SC	Switched-Capacitor
SNR	Signal-to-Noise Ratio
SPI	Serial Peripheral Interface
TCDS	Time-domain Correlated Double Sampling
USB	Universal Serial Bus

Abstract

Inspired by the working mechanism of human retina, the dynamic vision sensor (DVS) has become a research hotspot recently because of its certain advantages over the conventional frame-based image sensor. Instead of working in integration mode, the pixel in DVS continuously monitors local light intensity change through a front-end logarithmic photo detector and reports an event if certain threshold is reached. The output of DVS is not frames, but a stream of asynchronous address-events. The speed of the sensor is not limited by any traditional concept such as exposure time and frame rate. It can detect fast motion which is traditionally captured by expensive, high-speed cameras running at thousands of frames per second, but with significantly reduced amount of data.

The early DVSs only output binary address-events. In other words, only the row and column address of the detected events are reported by the sensor. The absence of absolute light intensity information, however, severely restricts their applications in lots of image processing algorithms, such as object recognition, classification and so on. This leads to the subsequent trend to integrate DVS with the “snapshot” function. Two pioneering works combine DVS with extra in-pixel light intensity measurement circuits based on pulse width modulation (PWM) and active pixel sensor (APS) scheme respectively. Unfortunately, both designs rely on the time-consuming integration method to achieve intensity information and thus, results in undesired discordance between the very fast address-events output and the postponed intensity readout, which neutralizes the potential advantages of the event-based system.

In this thesis, a new DVS which directly takes the voltage of the in-pixel logarithmic photo detector as its intensity information is presented. This voltage responds to incident light continuously and therefore no additional exposure time is needed. Combined with a column-parallel area-saving SAR ADC array, the sensor is capable of reporting the fast address-events as well as their corresponding intensity information simultaneously. Wide intra-scene dynamic range is also achieved due to the logarithmic compression of light intensity. On the other hand, an external control signal is also added which can force all the pixels to produce request signals and in turn a full-array on-demand reference picture can be obtained. This also enables the

sensor to output continuous frames, which makes it compatible with mainstream image processing algorithms. By means of some improvements in circuit design at pixel, column and system level, the image quality of the sensor is ensured with favourable signal-to-noise ratio (SNR) and fixed-pattern-noise (FPN). A 160×192 prototype has been implemented using AMS 0.35 μm 2P4M CMOS technology. The pixel occupies 30×30 μm^2 with 14.5 % fill factor. The total chip area is 7.3×7.4 mm^2 and the average power consumption is about 85 mW with 40 MHz clock.

Chapter 1

Introduction

1.1 Background and Motivation

High-speed cameras are important demands in many application areas such as television, industry, scientific research and even military purpose. They are widely adopted as indispensable diagnostic instruments to assist the audiences, engineers or researchers to figure out some high-speed processes, which happen too fast to be seen directly by the human eyes. In fact, each time a fast motion is required to be captured or tracked for further analysis, the high-speed camera is a valuable tool to implement such function. Some application examples are shown in Fig. 1.1. For application in television, one interesting example might be the so-called "Hawk-Eye technology" introduced to the live television on some important tennis events. As shown in Fig. 1.1(a), several high-speed cameras are installed at different locations and angles around the tennis court to track the trajectory of the fast moving tennis whose highest speed could exceed 200 km/h. These video data are then fed into the back-end host system to compute and display a three-dimensional record of a statistically most likely moving path of the tennis and thus, to assist the referee to make more precise and impartial judgements. With respect to the industry area, high-speed cameras are widely applied in industrial inspection, high-speed surveillance, machine vision and so on. As one instance, they are adopted in the automobile crash test to record the whole procedure when a fast test car collides with a rigid wall from different directions, as shown in Fig. 1.1(b). By means of these experimental data, the engineers could find out and improve some potential flaws of their design and enhance the safety performance of the product. As another instance, high-speed cameras also distinguish themselves in modern large-scale production line. Fig. 1.1(c) demonstrates a beverage manufacturing line which could assemble up to thousands of units per minute. The high-speed cameras are employed to inspect the bottles are correctly filled, capped and labelled. In terms of scientific research, high-speed cameras are also powerful tools

equipped in lots of biomedical or physicochemical laboratories to analyse some fast physiological phenomena or chemical processes. For instance, by means of these cameras, a group of researchers successfully figure out the flight mechanics of the dipteran insects, such as blowfly shown in Fig. 1.1(d), which are widely recognized as the smallest and most agile flying animals and their wingbeat frequency could reach as high as 200 Hz [1].

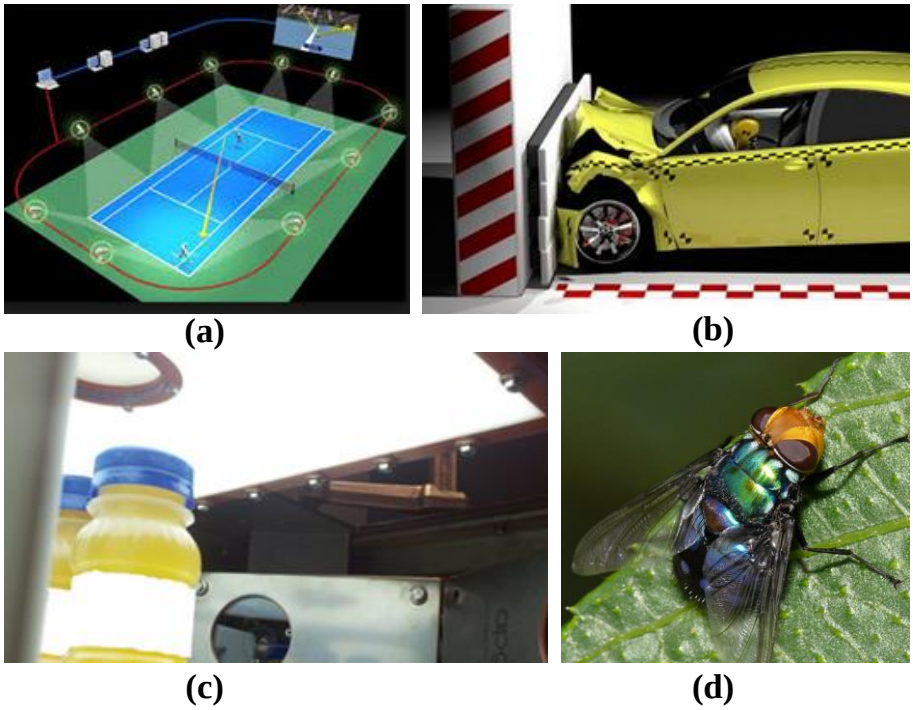
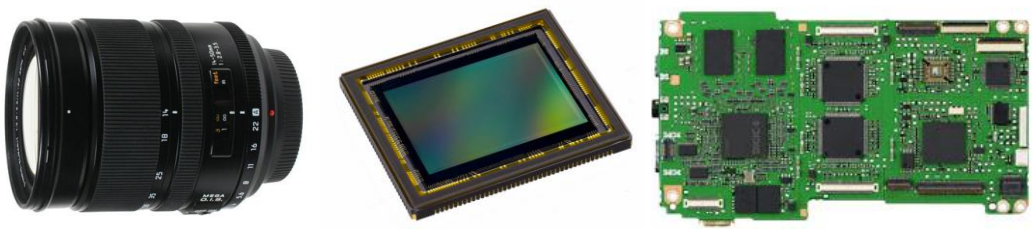


Fig. 1.1 Application examples of the high-speed camera.
(a) "Hawk-Eye technology" in live sports. (b) Automobile crash test. (c) Industry inspection in production line. (d) Animal locomotion.
(*All pictures come from Internet, corresponding sources are listed at the end of this Chapter)

As illustrated in Fig. 1.2, in general, the mainstream imaging system is composed of three basic parts: the front-end lens as the optical module to converge the ambient light, the image sensor as the photosensitive device to convert light into electrical signals and finally the back-end data acquisition system to store and process these data. Nowadays, as the heart of the whole system, the image sensor prevalently adopted in various application areas is the CMOS active pixel sensor (APS) which is driven by the continuous advancement of the CMOS fabrication process and improvements in circuit design methodology. The basic operation principle of the APS could be simply illustrated with two concepts as the exposure time and frame rate. Firstly, the pixel in the APS needs certain time, which is termed as the exposure time, to integrate the very

small photocurrent into a useful recognizable voltage signal. In addition, the APS captures a whole image (frame) at periodic time slices and the maximum number of frames generated by the sensor for one second is called frame rate, which usually depends on the resolution and readout speed of the sensor and also the data transfer rate of interface.



(a) **(b)** **(c)**

Fig. 1.2 Typical components of an imaging system.
(a) Lens. (b) Image sensor. (c) Data acquisition system.
(*All pictures come from Internet, corresponding sources are listed at the end of this Chapter)

With respect to application in high-speed cameras to detect instantaneous motions, it can be imagined that the APS has to run at very short exposure time and also high frame rate. Otherwise, it would either produce blurred images (long exposure time) or inadequate samples (low frame rate) which would be unqualified for further analysis. Both requirements, however, pose critical challenges to the design of the image senor itself and also the back-end data acquisition system. The short exposure time might degrade the corresponding image quality while the high frame rate characteristic burdens the image sensor with increased readout channel bandwidth and power consumption. From the perspective of back-end acquisition system, massive quantities of data are produced by the sensor, often in the order of GB/s. However, these video data usually contain a high level of redundancy. Large amounts of unimportant data have to be stored and processed before the features of interest are extracted. This means mass memory device and high-end processor should be integrated at the back-end. Both of these facts, on the one hand, increase the total cost of the whole system to a great extent. On the other hand, they make the high-speed camera very difficult to be used in some real-time applications because of the huge back-end signal processing requirement. To be more specific, the commercialized high-speed camera Photron SA3 could be used to illustrate the aforementioned issues. This camera is a 1-Mega pixel, high-speed video imaging system embedded with a frame-based APS featuring

17 μm pixel pitch and 12 bit on-chip ADC. Its frame rate could reach maximum 2000 frames per second (fps) under full resolution and the throughput is as high as 3 GB/s in this case. At the back-end, the largest provided memory device option (8 GB) could only support 2.72 seconds video recording under this condition. In addition, in order to process these massive data, the dedicated software "Photron FASTCAM Viewer" is also developed for the system. According to these facts, the whole high-speed camera costs more than 10,000 US dollars with 100 W total power consumption [2].

Inspired by the working principles of human retina, another distinctive kind of image sensor works based on the concept of "event", instead of "frame", and thus it is called the event-based image sensor. This new device has been proven to be more superior in high-speed application than the frame-based counterpart. As a successful implementation, the dynamic vision sensor (DVS) [3] works in continuous-time and only concentrates on the temporal contrast in field of view (FOV). In other words, the DVS only reports the dynamic contents of the scene (useful information) and automatically filters the static ones (background). In DVS, each pixel continuously monitors local light intensity change and reports an "event" to outside as address-event if certain threshold is reached. In external, asynchronous auxiliary logic circuits process the pixel requests based on handshaking communication protocol. The output of the sensor is not continuous frames, but a stream of asynchronous address-events. The back-end processor reconstructs the dynamic image of the scene based on the received address-events and corresponding timing information. No longer limited by any traditional concept such as exposure time and frame rate, the speed of this kind of sensor is determined by the event latency which is defined as the minimum delay from the occurrence of temporal contrast in FOV to the actual readout of corresponding event. It is usually in the order of several to decade of microseconds, which ensures DVS to detect fast motion which is traditionally captured by expensive, high-speed APS running at thousands of frames per second, but with significantly reduced amount of data. This kind of device has been successfully applied in some real-time machine vision algorithms such as high-speed motion detection, object tracking and shape recognition.

The binary address-events output of DVS can describe the contour of a motion because of the corresponding temporal contrast. However, the absence of pixel intensity information still severely restricts its applications since no grey-scale data

and also the background information are provided. In lots of situations, it is very difficult for the humans or computer algorithms to recognize the meaning of a motion by just using the contour information. This flaw has led to the development of two subsequent works, namely asynchronous time-based image sensor (ATIS) [4] and dynamic and active pixel vision sensor (DAVIS) [5]. Both designs render the conventional DVS capable of achieving the absolute light intensity information but based on different schemes. The ATIS adopts the pulse width modulation (PWM) scheme while the DAVIS employs active pixel sensor (APS) scheme. Such functionality combination addresses the requirement for both high speed object detection and tracking (using DVS output) and detailed object categorization (using intensity information). The event-driven intensity readout strategy also reduces the delivered data volume of the sensor. These advantages broaden the application of the DVS sensor to much wider areas which are usually undertaken by traditional high-speed frame-based APS. However, both systems still rely on the integration method to achieve intensity information and therefore, the illumination time resolution is limited by the required exposure time which could be as long as tens of milliseconds. This leads to certain temporal mismatch between the fast address-events output and the relatively slow intensity readout, which could either cause strong motion artifacts for ATIS or obvious image lag between the dynamic and static contents of the scene for DAVIS.

As an attempt to cope with aforementioned issue, this thesis presents a new DVS sensor that directly takes the voltage of the pixel front-end photo receptor as its intensity information. Since this voltage responds to light intensity in continuous-time, no additional exposure time is required and thus it takes no effort to match the output of the high-speed address-event and the corresponding intensity information. In order to implement this function and also ensure the image quality, some circuit design improvements have been developed in pixel, column and system level. Besides the event-driven intensity readout mode, this sensor is also capable of outputting continuous frames by addition of an external global control signal. This signal could force the whole pixel array to be fired and thus to obtain a full-array on-demand reference picture. Since the output voltage of the photo detector is logarithmically related to light intensity, wide intra-scene dynamic range up to 120 dB can also be achieved for this sensor.

1.2 Thesis Organization

The rest of this thesis consists of four chapters. Chapter 2 illustrates the literature review on some related works. The fundamental knowledge, including pros and cons, of the frame-based APS is first given. It is followed by detailed descriptions and analyses of basic concept and working principle of the DVS. The final part of this chapter introduces and summarizes two subsequent designs, ATIS and DAVIS, which integrate the DVS with intensity information readout capability. Chapter 3 first illustrates the basic idea and architecture of the new DVS sensor. The following part mainly concentrates on design considerations along the intensity readout path (from front-end photo detector to final ADC) with lots of efforts being paid to improve the image quality of this sensor. Chapter 4 presents the detailed measurement results of this sensor with respect to intensity readout and DVS function in detecting temporal contrast. The performance comparison with some related works are also summarized in this chapter. The last chapter presents the conclusion of this design and some recommendations for the future works.

Note: Sources of pictures from Fig. 1.1 and Fig. 1.2

Fig. 1.1(a):
https://trainingwithjames.wordpress.com/research-papers/the-impact-of-the-hawk-eye-system-in-tennis/
Fig. 1.1(b):
http://www.bankrate.com/finance/auto/how-much-safer-is-car-that-scores-higher-on-crash-tests.aspx
Fig. 1.1(c):
http://www.vision-systems.com/articles/print/volume-20/issue-10/departments/technology-trends/industrial-inspection-smart-cameras-check-bottles-at-high-speed.html
Fig. 1.1(d):
http://www.aussiecreatures.net/304
Fig. 2.1(a):
http://www.cameralabs.com/reviews/Leica1450mm/
Fig. 2.2(b):
https://martinmoorephotography.wordpress.com/2012/04/04/how-does-a-dslr-work/
Fig. 2.2(c):
http://www.ielectron.ie/repair-services/camera-repairs/canon-camera-repairs

Chapter 2

Literature Review

2.1 Frame-Based Image Sensor

2.1.1 Basic Concepts

After invented by AT&T Bell Laboratories at 1969, the charge-coupled device (CCD) has become the mainstream solid-state image sensor technology for a long time [6]. However, with continuous advancement of the fabrication process of CMOS technology, considerable research efforts have been switched to the design and implementation of CMOS image sensor (CIS). The early deficiencies of the CIS over the CCD device, such as large pixel area, large temporal noise and fixed-pattern noise has been mitigated significantly by means of both improvements in circuit design and fabrication process. In addition, the adoption of CMOS technology also enriches the functionalities of image sensor by integrating some previous back-end analog and digital processing units down to the sensor level. Based on these facts, more and more popular CISs not only snatch lots of existing markets originally undertaken by the CCD device but also continuously exploit new application prospects of the image sensor. With respect to different architectures of CIS, the active pixel sensor (APS) demonstrates better tradeoff among some key specifications of the imager, such as pixel size, readout noise, readout speed and so on, compared with the passive pixel sensor (PPS) and digital pixel sensor (DPS). Therefore, it has dominated both the academic and industrial community and almost become the standard architecture for image sensor design. Since the output data format is usually continuous frames, this kind of sensor is also called frame-based APS which would be elaborated in following paragraphs.

The simplified block diagram of the frame-based APS is shown in Fig. 2.1 [6]. The basic building blocks include the pixel array, the row and column access circuitry and the readout circuitry. The pixel array, as the core of an image sensor, consists of a

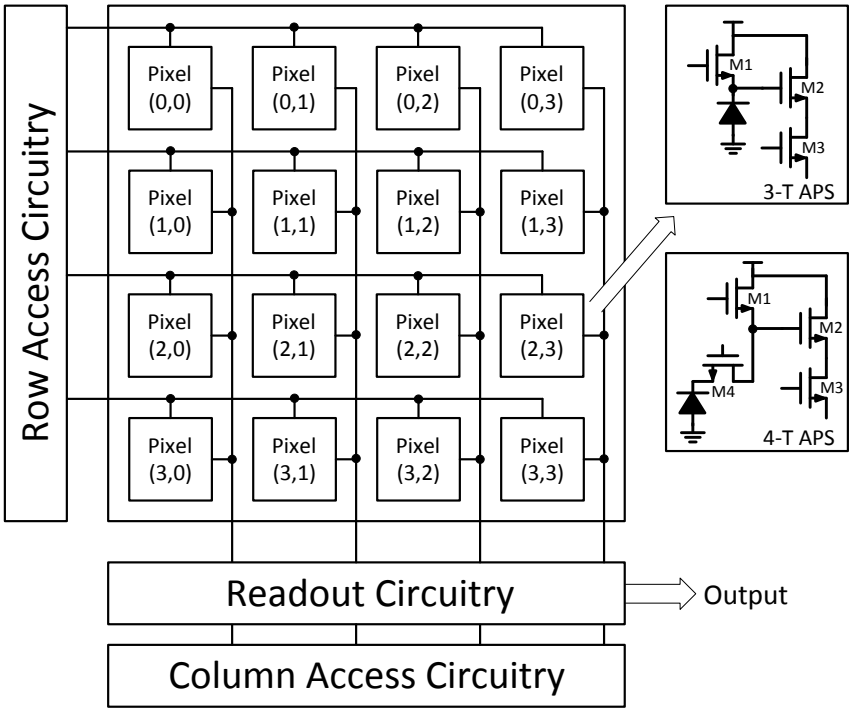


Fig. 2.1 Architecture of the frame-based APS.

large number of pixels distributed uniformly in two dimensions in space (only 4×4 pixels are shown for convenience). Each pixel first converts the incident light into a corresponding proportional photocurrent through a photodiode and then integrates this current into a useful voltage signal which could be further amplified or quantized by column readout circuitry and reported off chip. The most common pixel architecture is the 3-T and 4-T APS, which are also shown in Fig. 2.1. It should be pointed out that, throughout the whole thesis, the body connections of all NMOS and PMOS transistors are set to ground and power supply respectively, and are not shown in following schematics for convenience. With respect to the 3-T APS, the pixel is comprised of a photodiode and three extra transistors to implement such basic functions as reset ($M1$), buffering ($M2$) and accessing ($M3$). In contrast, the 4-T APS makes use of a pinned photodiode and adds a transfer gate transistor ($M4$) to the original 3-T APS. This brings about some advantages as high sensitivity, small dark current and also the possibility of implementation of the global shutter (instead of rolling shutter of 3-T APS) to the 4-T APS at the expense of larger pixel area or smaller fill factor. The peripheral row and column access circuitry are incorporated to control each pixel to output its individual information by means of the shared readout resources one after another and therefore, to avoid any potential readout collisions.

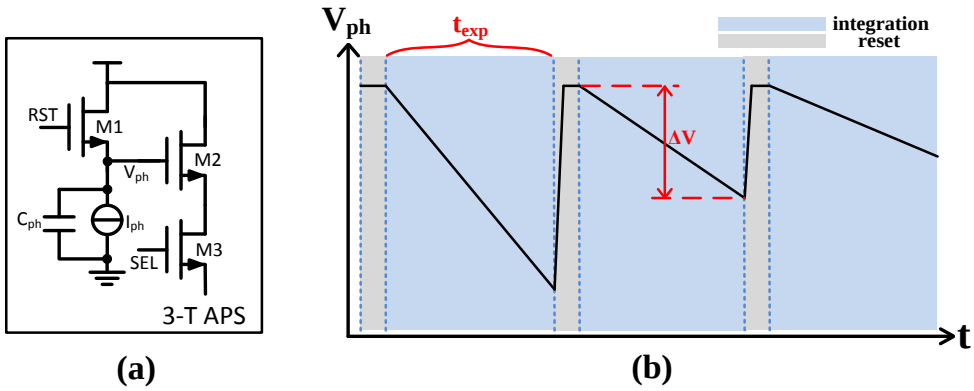


Fig. 2.2 Operation principle of the 3-T APS. (a) Schematic. (b) Timing diagram.

The operation of the frame-based APS, in principle, is quite straightforward and transparent. More specifically, the simpler 3-T APS is used to explain such procedure, which is shown in Fig. 2.2. The pixel works alternatively between the reset and integration mode. During reset mode, $M1$ turns on and V_{ph} is pulled up to a high reset level. Afterwards, $M1$ turns off and the pixel is switched into the integration mode during which V_{ph} drops as the photocurrent (I_{ph}) discharges the capacitance of this node (C_{ph}). The total time spent on the integration mode is called exposure time (t_{exp}). At the end of this mode, the integration voltage ΔV, as expressed in equation (2.1), is readout as the representation of light intensity sensed by this pixel. The Q_t in (2.1) represents the total charges generated by photodiode during integration mode and is the integral of photocurrent $I_{ph}(t)$ over the whole exposure time t_{exp}. It should be noted that these operation procedures of pixel are controlled by external row and column access circuitry.

$$\Delta V = \frac{Q_t}{C_{ph}} = \frac{\int_0^{t_{exp}} I_{ph}(t) \cdot dt}{C_{ph}} \tag{2.1}$$

In order to get a full picture, the image sensor scans the whole pixel array once according to a particular order. As the most common implementation, progressive scanning reads the information of pixel array from the first until the last pixel row by row. As for the image sensor with 4×4 pixel array shown in Fig. 2.1, the row access circuitry selects the first row (*Row0*) at the beginning. The column access circuitry selects the first column (*Column0*) in the meanwhile and thus, the integration voltage of *Pixel (0,0)* is achieved by means of the readout circuitry. After that, the pixel in the next column, *Pixel (0,1)*, is selected and its integration voltage is scanned out as well.

The same operation is executed repeatedly in this row until the last pixel in this row is readout, in this case *Pixel (0,3)*. The row access circuitry, then, switches to the next row (*Row1*) and the integration voltages from *Pixel (1,0)* to *Pixel (1,3)* are scanned out in sequence. Based on such progressive scanning mechanism, a full image which is called a frame could be achieved after the last pixel in the array, *Pixel (3,3)*, is read out. In general, the scanning of the next frame could be initiated immediately after the completion of the previous frame which results in continuous frames output of the image sensor. The maximum achievable number of frames is called the image sensor's frame rate, which is a crucial specification to evaluate its performance in high-speed application.

The simple pixel architecture leads to a very compact design, which is favourable for integration in very large-scale [7]. In addition, the well-understood progressive scanning mechanism simplifies the implementation of the peripheral control blocks (usually simple shift-register based scanners). Benefiting from these advantages and continuous advancement of the CMOS technology, the developments of frame-based APS towards larger resolution, higher frame rate and better image quality have enjoyed dramatic progress over the last few decades. Actually, it has been the mainstream image sensor design and dominated both the industrial and academic community for a long time.

2.1.2 Limitations in High-Speed Application

In-depth investigating the operation principle of frame-based APS, however, reveals its two main limitations especially in high-speed application. Firstly, as mentioned above, the pixel in APS works alternatively between reset and integration mode. The integration voltage ΔV is readout at the end of the integration mode as a representation of the illumination sensed by pixel. Since this voltage is determined by the total charges generated by photodiode during integration mode, the intensity information finally achieved is only related to the averaged illumination within the exposure time. This issue could be readily illustrated with equation (2.2), whose front part is same to equation (2.1).

$$\Delta V = \frac{Q_t}{C_{ph}} = \frac{\int_0^{t_{exp}} I_{ph}(t) \cdot dt}{C_{ph}} = \frac{t_{exp} \cdot I_{ph\_ave}}{C_{ph}} \tag{2.2}$$

By applying the mean value theorem of integrals and with fixed exposure time t_{exp}, ΔV is only proportional to the average photocurrent (within the exposure time) I_{ph_ave}, which corresponds to the average light intensity sensed by this pixel. In other words, only the average light intensity information within the exposure time could be attained by the frame-based APS while the transient characteristics are ignored autonomously.

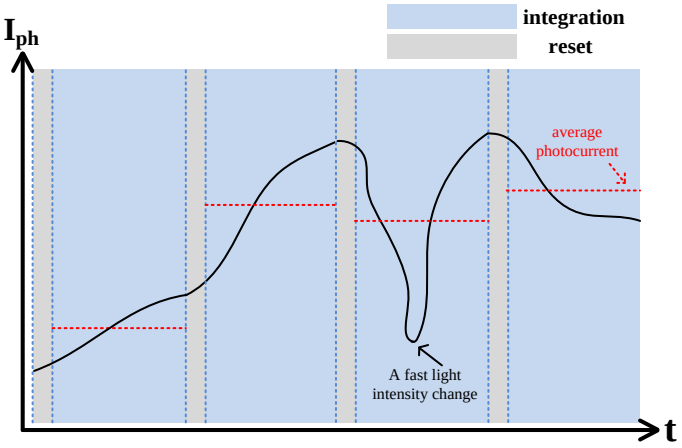


Fig. 2.3 Illustration of the artificial averaging effect of APS.

A direct consequence of this kind of artificial averaging effect is that any fast brightness change which might be caused by a fast motion lasting for very short time with respect to the total exposure time is difficult to be detected by the imager. Fig. 2.3 illustrates this issue by showing the photocurrent of a single pixel. It can be seen that a fast light intensity change happens within the third frame, but the information we finally achieved is only related to the average brightness (represented by average photocurrent shown in red dash line). When it comes to the images, this would cause obvious motion blur as shown in Fig. 2.4 (a). In this case, the sensor captures a fast rotating fan with long exposure time as $1/30\ s$ and the edge of the fan has been mixed with the background white ceiling and could not be clearly recognized. This problem could only be relieved by shortening the exposure time. After reduced to $1/400\ s$, the fan is in sharp focus and its edge is highly recognizable as shown in Fig. 2.4(b). Actually, the required exposure time to detect high-speed motion is determined by the relative speed of the motion within FOV and very short exposure time is required to figure out very fast motion. This makes sense in principle but poses threat in real application. It should be remembered that the useful integration voltage swing of the pixel is still proportional to exposure time, as illustrated in (2.2). The short exposure

(a) **(b)**

Fig. 2.4 Sample images of APS with different exposure time (a) $1/30\ s$ and (b) $1/400\ s$ [8].

time would lead to a small voltage swing. For example, with 10 pA I_{ph_ave} and 10 fF C_{ph}, the integration voltage ΔV is 1 V for 1 ms exposure time and would reduce to only 100 mV for 100 μs exposure. In contrast, other nonidealities of the APS such as pixel nonuniformity and temporal readout noise still remain constant. This indicates every 20 dB drop of signal-to-noise ratio (SNR) and tenfold rise of fixed-pattern-noise (FPN) of the APS with per decade shortening of the exposure time. As a result, the reconstructed image quality would be deteriorated to some extent with the same effect when an image is taken at very dark environment. This reveals the tradeoff of degraded image quality for high-speed camera.

Another important issue of CMOS APS is large readout data redundancy, which is caused by the frame-based readout strategy. As mentioned in previous paragraph, short exposure time ensures the APS to generate sharp image with respect to fast motion. However, this is only one prerequisite condition and may not adequate for real applications since a large number of sample images should also be provided in the meanwhile. The latter requirement is met by the APS with high frame rate. Take the imaging of the flight procedure of dipteran insects as example, as mentioned in previous chapter, the wingbeat frequency of such kind of animal could reach as high as 200 Hz. For a 200 fps APS, it could only output one sharp image (if short exposure time is guaranteed) for one complete wingbeat process. In contrast, the sample images would increase to 20 (at different phase) if the frame rate is promoted to 4000 fps. In other words, more images could be produced with higher frame rate, which in turn ensures more continuous recording and observation of fast motion. In general, the frame rate of high-speed APS is usually higher than 500 fps. On the other hand, based

on the aforementioned analysis, the pixel in APS is a little “passive” because it could not self-report its information outside without the control of row and column access circuitry. From the perspective of the sensor, since it is unable to know which pixel contains the useful information (temporal contrast), it has no choice but to scan the whole pixel array in order to achieve an updated image, regardless of whether this information has changed or not compared with the previous one. This is to say that some unchanged background contents, even though contain no useful information, are still readout mechanically and periodically. Combined with the high frame rate of the APS, it is easy to imagine that hundreds or thousands of frames produced (by the APS) per second must contain huge amount of unnecessary data. Acquisition and processing these massive data, on the one hand, burden the image sensor itself with increased readout channel bandwidth and larger power dissipation. As for the back-end signal processing, they also require mass storage device and powerful post-processor. Both of these facts increase the cost of the whole system markedly.

2.2 Event-Based Image Sensor

2.2.1 Human Retina

The two main limitations associated with the frame-based APS stem from the acquisition of visual information as a series of “snapshots” at periodic discrete instants. The relevant progressive scanning readout strategy, as mentioned above, is mechanical and inefficient. In contrast, there exists another kind of device which works quite differently with the current mainstream frame-based image sensor. The working principle of this type of image sensors is based on ‘event’, instead of ‘frame’, and accordingly they are called event-based image sensor. They are inspired and developed by the distinctive operation principle of human retina and proven to be more efficient in high-speed application than the frame-based counterpart.

The human eye, as the product of natural evolution for millions of years, is a very complicated biological imaging system and believed to be more intelligent and powerful to receive and process the ambient visual information than the artificial system. For the most basic components, both systems are similar to each other with lens, photosensitive device and back-end signal processing unit. The photosensitive

device in artificial camera is the image sensor while in human eye it refers to the retina. Although implement the same function, these two components operate quite differently. As indicated earlier in this chapter, the mainstream CMOS APS works synchronously and captures continuous frames at a given speed. In contrast, the operation principle of the retina is considered to be more advanced with its self-timing and data-driven characteristic. Investigating the detailed structure and working principle of retina could stimulate the inspirations for development of "smart" image sensors which might be free from the aforementioned drawbacks of frame-based imagers in detecting high-speed motions. Actually, the concrete operation principles of human retina are very complicated and only partially understood until now. Based on current knowledges, however, two different kinds of ganglion cell, namely "Magno" and "Parvo" cell, of the retina are widely recognized to play important roles in sensing the ambient visual information [9]. These two kinds of cell are responsible for converting the dynamic and static information of the visual scene respectively. The "Magno" cell is sensitive to the transient change of the scene and its corresponding Magno-cellular pathway has short latency and fast response rate. The "Parvo" cell, on the other hand, is more concentrated on the sustained information and has much slower cellular pathway. Its main function is to extract more details, such as the texture and colour, of the visual image.

The differentiation and collaboration of both types of ganglion cell lead to a distinctive operation principle of human retina, which is quite different from that of the frame-based APS. In observing a scene, the "Magno" cells firstly monitor whether there exist light intensity changes in FOV and if detected, lock the relevant spatial information rapidly. They filter the unrelated region and narrow down the interested area which needs to be further observed to recognize more details, while the latter task is accomplished by the "Parvo" cells. It should be noted that the "Magno" and "Parvo" cells work continuously and asynchronously in processing the visual information and therefore, there is no such concept as exposure time and frame. In other words, the human retina could respond to light intensity change immediately. It is immune from the undesired averaging effect resulted from the artificial integration procedure of the frame-based imagers. On the other hand, since only the dynamic region detected by the "Margo" cells needs to be further recognized by the "Parvo" cells, the delivered data volume and post-processing workload is also smaller for human retina. Last but not the least, it should be pointed out even with such advanced mechanism, the human

eye is still not capable of recognizing high-speed motions due to the well-known "persistence of vision". Actually, a high-speed camera equipped with an APS with frame rate of hundreds of fps has already outperformed the human eye in detecting fast motions. This is because that the transmission of the neural signals in human being is implemented by means of the ions' motion in solution between different cells. Unfortunately, this speed is only about $1/10^7$ that of the mobility of electrons in silicon.

2.2.2 Dynamic Vision Sensor (DVS)

With respect to the implementation of human retina with integrated circuits, actually all the frame-based image sensors are emulations of the "Parvo" cells in terms of functionality because they could only report pixel intensity information. In order to attain full advantages of human retina, the appropriate modeling and emulation of the "Magno" cells are indispensable. Although developed long time ago, most of the early such biomimetic devices were only built for demonstrations or verifications of certain basic neurobiological theories [10-13]. Some intrinsic deficiencies of these conceptual devices, such as large silicon area, large pixel mismatch and so on, make them unqualified for real applications. The first truly usable device to simulate the "Magno" cell is usually regarded as the so-called dynamic vision sensor (DVS) which is introduced by Tobi's group in 2008 [3]. The reported vision sensor works asynchronously and only concentrates on temporal contrast in FOV. Focused on improvements in circuit design, the sensor demonstrates good characteristics with wide dynamic range, low latency, low power consumption and small pixel mismatch. The circuit topology, especially the pixel design, has almost become the standard implementation in such kind of sensor and is widely cited in subsequent designs.

The basic architecture of the DVS sensor with a simplified 4×4 pixel array is shown in Fig. 2.5. All pixels are connected with each other to constitute a two-dimension network through row and column asynchronous handshaking buses, where RR/RA refer to row request/acknowledgement and CR/CA column request/acknowledgement. Other building blocks also include the external handshaking logic, address encoder and arbiter tree in row and column direction respectively. These peripheral modules,

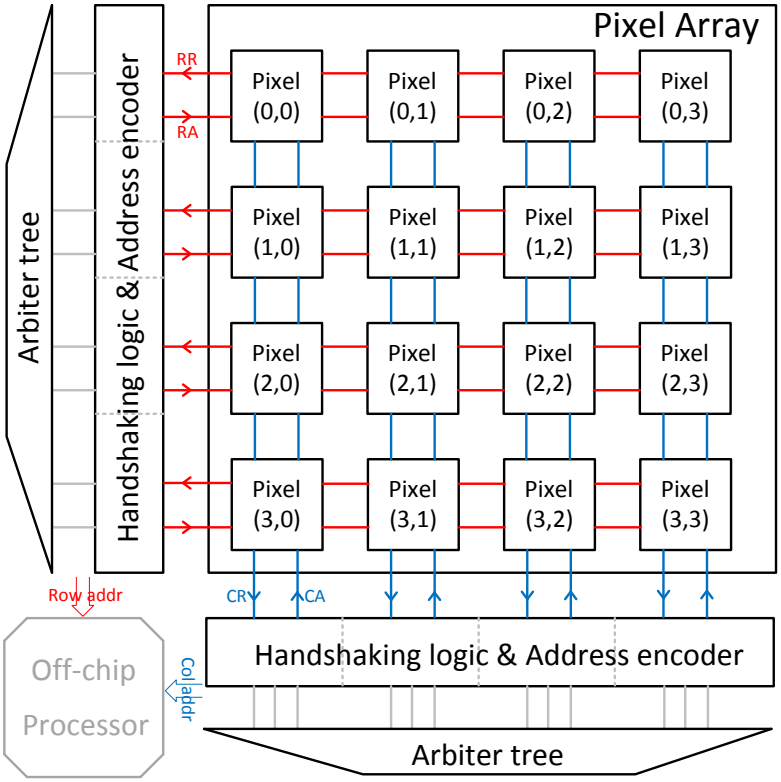


Fig. 2.5 Architecture of DVS sensor.

which correspond to the row and column access circuitry in frame-based APS, are designed based on the asynchronous 4-phase handshaking protocol to process pixels. In addition, the row and column arbiter tree are incorporated at the top of external handshaking logic to avoid collisions when a large number of fired pixels send request signals at the same time. These pixels would be served sequentially based on the arbitration results from the arbiter tree. In order to implement the function of "Magno" cell which is only sensitive to brightness change of the visual scene, the pixel design is very unique and it is also the very beginning point to understand the operation principle of this kind of sensor.

Different from the frame-based APS, each pixel in DVS continuously monitors local light intensity change and issues asynchronous spike to outside when such change exceeds the given limit. Fig. 2.6 shows the schematic of the DVS pixel which contains four separate parts: logarithmic receptor, difference amplifier, comparators and interface logic [9]. The front-end logarithmic receptor converts the photocurrent into a logarithmic output voltage (V_{O1}) through an active feedback network formed by an inverting amplifier $A1$ and a single transistor M_{fb}. The corresponding bandwidth is thus enhanced significantly compared with the passive logarithmic receptor which

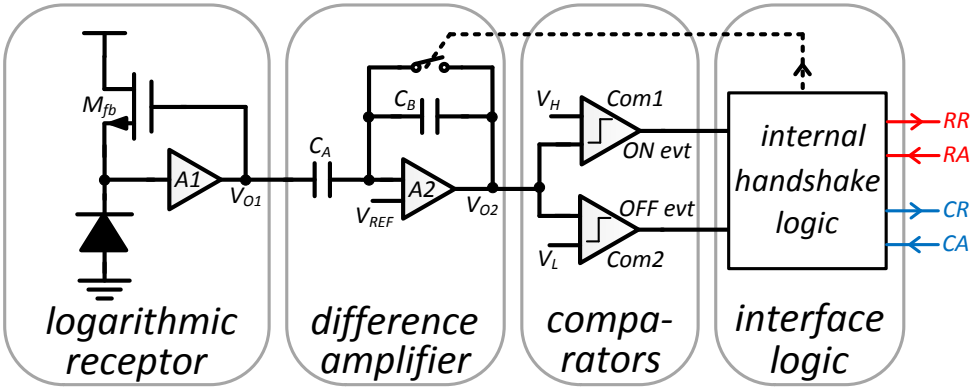


Fig. 2.6 Abstracted schematic of DVS pixel.

directly feeds the photocurrent into a diode connected MOSFET. The subsequent difference amplifier provides a large voltage gain (C_A/C_B around 20) to increase the temporal contrast sensitivity of the sensor. Two comparators $Com1$ and $Com2$ set the contrast thresholds that could be detected by the pixel. An "ON/OFF event" is detected when the threshold V_H/V_L is reached by V_{O2}. The pixel is then called to be fired and the internal interface logic would be triggered by issuing the request signals to outside peripheral modules to ask for being served.

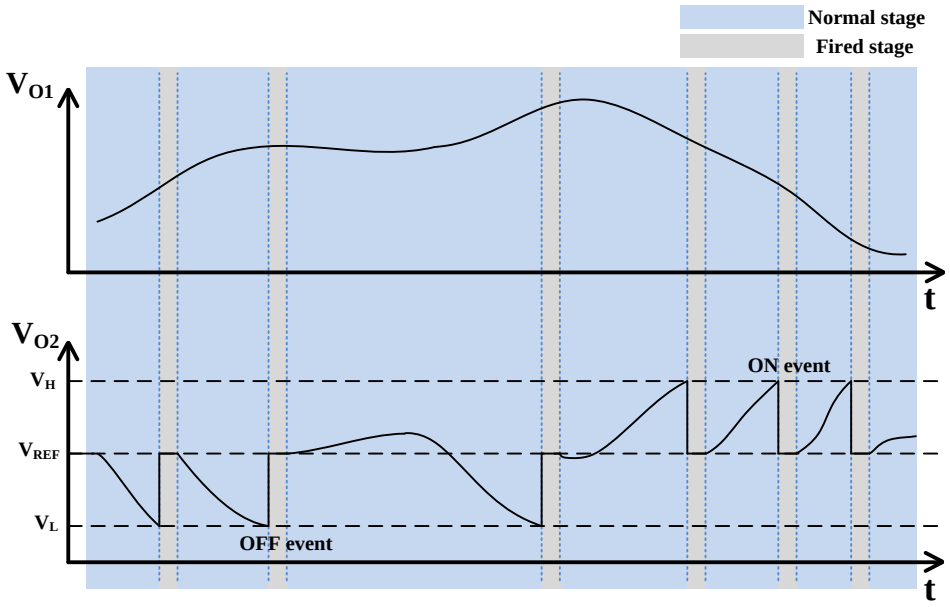


Fig. 2.7 Operation principle of DVS sensor.

Fig. 2.7 illustrates the working principle of DVS in pixel level by showing the waveforms of two key signals V_{O1} and V_{O2}. The output of logarithmic receptor, V_{O1},

continuously responds to local incident light intensity in logarithmic form. This voltage is amplified by the subsequent difference amplifier to V_{O2} and then fed into two continuous-time comparators which detect light intensity change towards two different directions. When V_{O2} exceeds the threshold window ($V_L\ to\ V_H$), the pixel would enter into the fired stage (shown as dark region) during which the handshaking communication with the peripheral modules is initiated by this pixel. In the meanwhile, the in-pixel logic circuit also resets the difference amplifier (V_{O2} is pulled back to V_{REF}) to avoid this fired pixel to generate persistent events. The processing of fired pixel is according to the 4-phase handshaking protocol between in-pixel and external handshaking logic circuits. As illustrated in Fig. 2.8, the fired pixel first issues RR (Row Request) and when this row (of the fired pixel) is selected by the row arbiter tree, the relevant RA (Row Acknowledgement) would be granted. The pixel then issues CR (Column Request) in column direction and once CA (Column Acknowledgement) is granted by the column arbiter tree, the pixel is being processed and its address information would be reported to the off-chip processor as address-events (the temporal information is automatically recorded when the event is received by the processor). The handshaking communication would then be terminated by revoking the RR (also RA) and CR. Afterwards, the pixel would return back to the normal stage during which it would sense the light intensity change again. Such operation principle leads to an imager which only focuses on temporal contrast of the visual scene, and this is exactly the function of the "Magno" cells in human retina.

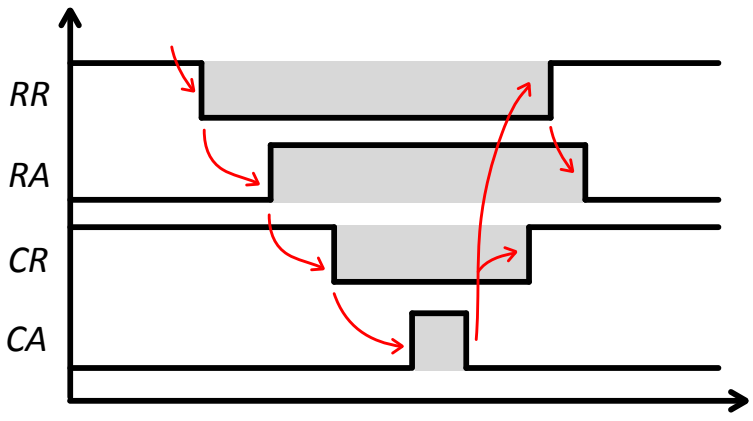


Fig. 2.8 Four-phase handshaking protocol to process the fired pixel.

From system level, the operation principle of DVS could be explained with a simple example shown in Fig. 2.9. In this scenario, an object flies across the FOV of the sensor with the corresponding track indicated as dash line in this figure. When the object is out of FOV, no temporal contrast could be recognized by the sensor. The DVS is thus called in idle state since it reports nothing to the back-end. Once the object flies into FOV and the corresponding temporal contrast is detected, the DVS

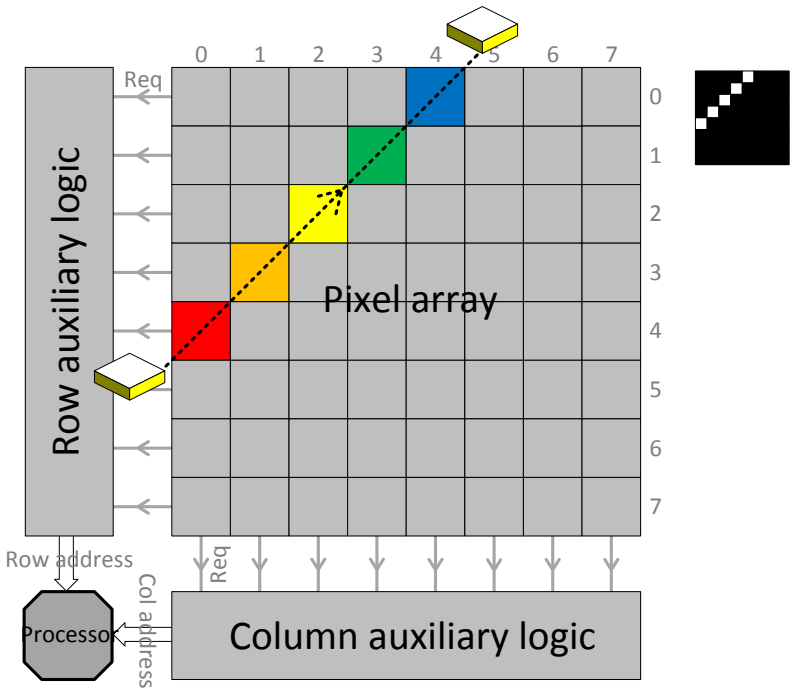


Fig. 2.9 Address-event representation of the DVS in detecting moving object (reconstructed image is also shown at top right corner).

would enter into active state and reports a stream of address-events which could be utilized to reconstruct the object's motion track. More specifically, when the object flies across the sensing area of first pixel (red), the light intensity sensed by this pixel would change accordingly. Once the temporal contrast exceeds the given threshold, this pixel would be fired and then processed according to aforementioned handshaking protocol. In the meanwhile, its row and column address would be reported out to the off-chip processor as (*Row4, Col0*). Afterwards, when the object is moving into the sensing area of the next pixel (orange), this pixel would also be fired and its corresponding addresses is delivered out as (*Row3, Col1*). With this procedure going on, the following pixels whose corresponding sensing area is covered by the moving object would report an event respectively in sequence. In this case, the yellow, green

and blue pixel as shown in Fig. 2.9 would report an address-event as *(Row2, Col2)*, (*Row1, Col3*) and (*Row0, Col4*) in succession. Afterwards, the object flies off the FOV again and therefore, the DVS returns back to the idle state. In terms of the back-end processor, it would receive only 5 address-events in total. Besides, the according timing information of each event is also recorded at back-end when it is reported by DVS. Based on these temporal and spatial information, the track of the moving object could be reconstructed as a monochrome picture which is also demonstrated at the top right corner in Fig. 2.9. In this picture, a white pixel represents an occurrence of corresponding event within the time period of observation. Otherwise, the pixel would be marked as black. Such reconstruction method of the dynamic content of a scene is termed as address-event-representation (AER).

(a) **(b)**

Fig. 2.10 Sample images of the DVS. (a) Faces. (b) Driving scene [3].
(gray- no event/white-"on" event/black-"off" event).

With the concept of AER, some important characteristics of the DVS could be figured out as following. Firstly, the reconstructed images are black-and-white images which only describe the occurrence of temporal contrast with relevant spatial and temporal information. This indicates that only the contour of moving objects could be detected and displayed. There is another scenario that the brightness contrast of the static background could also be detected if the sensor itself moves back and forth. The sample images of the DVS with respect to the latter situation are shown in Fig. 2.10. Actually, two different kinds of event are distinguished in DVS. The "ON" event is generated by gradually brighter contrast and marked as white in the reconstructed image while the "OFF" event is produced by gradually darker contrast and marked as

black. This differentiation provides certain sense of depth for the image and also facilitates some back-end image processing algorithms. However, this also reveals one significant disadvantage of the DVS that no intensity information is provided by the sensor due to the lack of corresponding readout circuit. The application of the DVS is thus restricted to high-speed motion detection, object tracking and so on. This flaw of DVS has been partially settled with the development of two subsequent works which integrate the intensity information readout circuit to the original design. The more detailed introduction of these two works would be given latter in this chapter. Apart from this, the DVS demonstrates superior performance over the framed-based APS in high-speed application. On the one hand, the pixel in DVS works in continuous-time, which means that it could respond to fast light intensity change rapidly (as long as the bandwidth is large enough). The required exposure time of APS is no longer needed here and the speed of DVS is only determined by the event latency, which could be as short as several to decade of microseconds. This ensures the DVS to detect and record very fast motions with precise timing information. On the other hand, compared with the 3-T APS, the pixel in DVS is more "active" since it could work autonomously. As indicated earlier in this chapter, the pixel of the 3-T APS is a little "passive" because it could only rely on the external row and column access circuitry to address it and report its own information. Consequently, the APS must scan all the pixels periodically in order to achieve the updated visual information. Large amount of useless data are readout repeatedly which results in huge data redundancy especially for APS with high frame rate. In contrast, the DVS pixel could self-report its information to outside when an event is detected, which indicates only useful information is delivered out (the useless data are filtered at pixel level). The peripheral control logic, in this case, does not need to scan all the pixels mechanically but only to respond and serve the fired pixels instead. Such readout mechanism improves the data communication efficiency of DVS to a great extent and therefore relieves the pressure for the back-end data acquisition and processing system.

Focused on improvements in circuit design, the reported DVS sensor with 128×128 pixels achieves wide dynamic range (>120 dB), low pixel mismatch (2.1 %) and low power dissipation (24 mW). The 15 μs event latency is also short, which enables the sensor to perceive precise timing information of events in observing high-speed motions. The output data volume, depends on the dynamic content of scene, is typically orders of magnitude lower than that of the frame-based counterpart.

Combined with lower power consumption, the sensor has been successfully used in such applications as high-speed robotic target tracking, traffic data acquisition and so on.

2.3 Event-Based Image Sensor with Intensity Readout

The DVS sensor, as a successful emulation of the "Magno" ganglion cell, is only sensitive to temporal contrast of a scene. The absence of absolute light intensity information, unfortunately, restricts its applications in some image processing algorithms, such as object recognition and classification, which require the static content of a scene. In other words, the functionality of the "Magno" and "Parvo" cells should be integrated monolithically in order to fully make use of the advantages of the human retina. This leads to the subsequent one trend to enable the DVS with intensity information acquisition function (the other trend is to improve the specifications of the pure DVS, such as contrast sensitivity, event latency and so on [14-16]). Two pioneering works implement such functionality integration by combining the DVS with extra in-pixel light intensity measurement circuits based on pulse width modulation (PWM) and active pixel sensor (APS) scheme respectively. These two designs are the asynchronous time-based image sensor (ATIS) and the dynamic and active pixel vision sensor (DAVIS).

2.3.1 Asynchronous Time-Based Image Sensor (ATIS)

The first attempt to enable the DVS to obtain pixel light intensity information is called the ATIS, which is introduced by Christoph Posch's group in 2011 [4]. On the basis of the DVS pixel, an extra PWM imaging circuitry is added to measure the absolute brightness of the visual scene. As shown in Fig. 2.11(a), the pixel in ATIS consists of two parts: change detector and conditional exposure measurement unit. The change detector shares the same circuit topology with aforementioned DVS and it is in charge of detecting relative intensity change (temporal contrast). The exposure measurement unit, which is triggered by the front-end change detector, is designed based on PWM scheme to achieve the absolute light intensity information. The PWM technique, as a

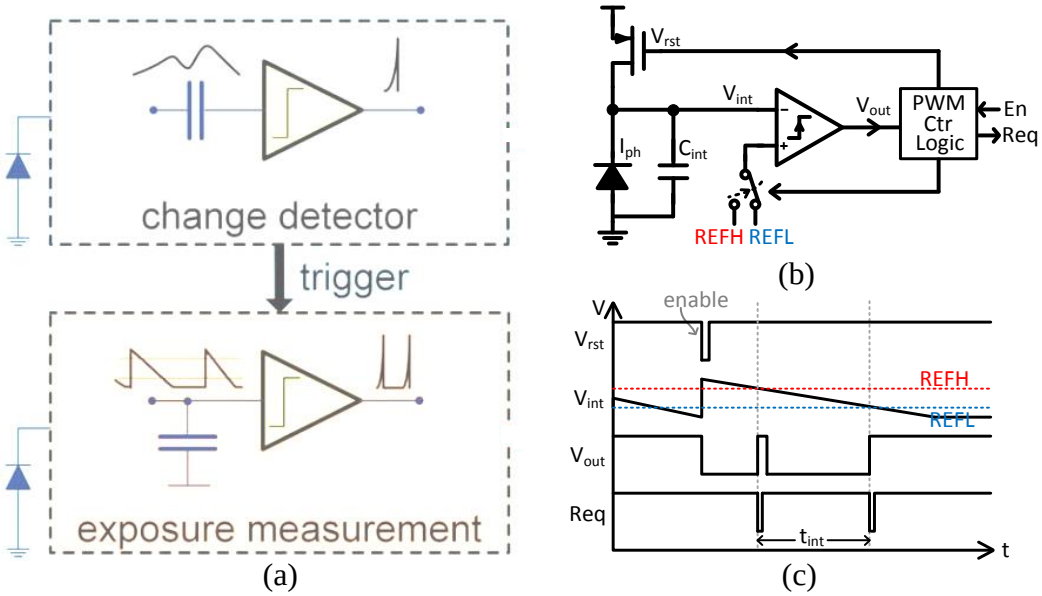


Fig. 2.11 (a) Block diagram of ATIS pixel [9]. (b) Schematic and (c) Operation principle of the exposure measurement circuit.

commonly-used current-measurement method, first converts the input current into a linearly varying voltage through an integrator. Afterwards, by means of a comparator, it measures the time instant when this voltage reaches a constant threshold and uses it to evaluate the input current. It is an effective method to quantize current and has been widely applied in time-based image sensors to broaden the dynamic range. It should be pointed out that two separate photodiodes are adopted in the ATIS pixel with one for change detector and the other for exposure measurement unit to implement their respective function.

Fig. 2.11(b) shows the detailed exposure measurement circuit of ATIS. Actually, it is an improvement to the classic PWM circuits with two threshold voltages ($REFH/REFL$) constituting a time-domain correlated double sampling (TCDS) circuit. The operation principle could be explained with Fig. 2.11(c). When the change detector monitors an event, it will enable the corresponding exposure measurement circuit once by imposing a short reset pulse (V_{rst}). The integration voltage V_{int} is first reset to VDD and then drops during the subsequent integration stage. At the beginning, the higher threshold ($REFH$) is selected by comparator. When V_{int} reaches $REFH$, the comparator flips and the PWM control logic issues the first request (Req) to the external handshaking logic. At the same time, the threshold of the comparator is switched to the lower one ($REFL$) and the integration continues until V_{int} finally reaches $REFL$. The PWM control logic sends another request (Req) at this moment

and the time difference (t_{int}) between these two requests is calculated off-chip to evaluate the light intensity of this pixel with their mutual inverse relationship. By employing such TCDS technique, the comparator offset, switching delay error and KTC noise could be eliminated and therefore, good image quality could be ensured by ATIS.

The PWM scheme ensures wide intra-scene dynamic range of ATIS while the proposed TCDS circuit enhances the image quality. In addition, the time-domain readout strategy of light intensity information through asynchronous request signals is fully compatible with that of the DVS sensor. This simplifies the peripheral logic design and the back-end signal processing since both the temporal contrast and the grayscale data are reported in terms of asynchronous events. Combined with in-pixel conditional exposure measurement triggered by front-end change detector, the ATIS could yield lossless high-quality video compression across the focal plane.

2.3.2 Dynamic and Active Pixel Vision Sensor (DAVIS)

Different from ATIS, the other implementation, DAVIS, makes use of the conventional APS scheme to obtain the absolute intensity information [5]. It mainly consists of a dual-mode compact pixel design which merges the 4-T APS with DVS sharing a single photodiode. This is possible based on the fact that the front-end photoreceptor in DVS only converts the photocurrent into a logarithmic voltage without destructing it. As a result, such intact photocurrent could be reused in APS to be integrated over time to generate a voltage signal.

Fig. 2.12 shows the schematic of the DAVIS pixel which contains two separate parts: the APS (red) and DVS (blue). In addition, another cascode transistor ($M5$) is also incorporated as an isolation of these two parts. Although combined in a single pixel, the APS and DVS could operate independently without interfering with each other. The APS is responsible for continuous frames output while the DVS for address-events. The 4-T APS ($M1 - M4$) is preferred in this design to implement the global shutter, instead of rolling shutter, during the intensity readout mode. The adoption of global shutter is crucial to avoid the appearance of undesired motion blurring in images since all pixels could be controlled synchronously to have the same exposure time. Another potential advantage of DAVIS is the possibility of

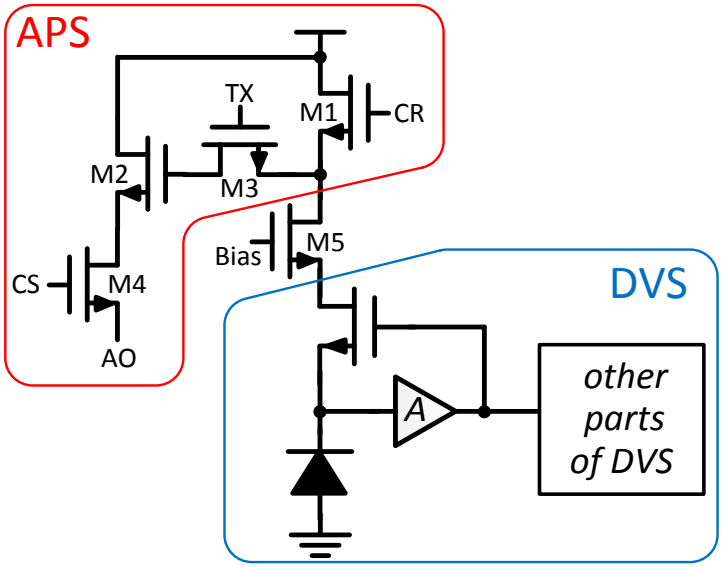


Fig. 2.12 Abstracted schematic of DAVIS pixel.

implementing region of interest (ROI) readout, which results from the separation of the exposure and readout phase.

Due to the compact pixel design sharing the same photodiode, the area of a single DAVIS pixel is only 18.5×18.5 μm^2, which is much smaller than 30×30 μm^2 of the previous ATIS. The DAVIS is capable of reporting asynchronous address-events and synchronous frames at the same time. In machine vision algorithm, the former could be made use of to track and segment fast moving object while the latter to further recognize more details in region of interest. According to the dynamic content of the scene, the address-events output might be utilized as an index to adjust the frame rate of the intensity output adaptively. This brings about significant reduction of the transported data volume and also total power consumption. Moreover, the synchronous frame output also makes the DAVIS compatible with mainstream image processing algorithms and applications.

2.3.3 Summary of ATIS and DAVIS

Although implemented differently, both ATIS and DAVIS render the DVS capable of achieving absolute light intensity information. The corresponding event-driven intensity readout mechanism (actually only ATIS, DAVIS still outputs full frame) alleviates such problems as large data redundancy and power consumption which inflict the conventional frame-based imagers for a long time. In addition, such

functionality integration also broadens the application prospects of DVS from the early high-speed motion detection and object tracking to the more generalized algorithms such as object recognition, classification and so on.

On the other hand, however, both designs still suffer from some limitations which might hinder their popularity to some extent. With respect to ATIS, one noticeable problem is the considerably large pixel area, which is 30×30 μm^2 even use 0.18 μm CMOS technology. The adoption of two photodiodes, together with the complicated intensity readout circuitry, occupies lots of valuable pixel area. This poses great challenge for further implementation with much larger resolution, which is usually one basic requirement for some industrial applications. What's more, the time-based acquisition of intensity information also triples the asynchronous event data volume. As for DAVIS, although the output frame rate of intensity readout might be adjusted according to the dynamic contents of visual scene (represented by detected address-events), certain degree of data redundancy is still existed because a full frame, instead of just the region of interest, is exported as a whole. What's more, the APS intensity readout method also restricts the intra-scene dynamic range to only 51 dB, much lower than that of the natural scene (>120 dB) [9].

In addition to the issues mentioned above, both designs still rely on integration strategy to achieve the brightness information. Though the output is in the form of asynchronous time spike, the PWM technique adopted by ATIS, in essence, is still a variety of integration method but converts the integration voltage into time-domain information. In both cases, such integration procedure still requires some exposure time which in turn retards the achievement of the brightness information. This leads to undesired discordance between the very fast address-events output and the postponed intensity readout, which neutralizes the potential advantages of the event-based system. In fact, the address-event latency is in the order of several microseconds for both designs, while the time resolution for intensity readout could be as long as tens of milliseconds. Such kind of mismatch leads to some problems when adopts these sensors in real high-speed applications. In terms of ATIS, for example, an object moving back and forth rapidly in FOV can become invisible due to the postponed event-triggered intensity readout mechanism [5]. The corresponding intensity measurement would never be finished as new events are generated continuously and

quickly. With respect to DAVIS, even though it could output full frame, there still exist significant image lag between the dynamic and static information of the scene.

This page is intentionally left blank.

Chapter 3

Design of a DVS with Direct Logarithmic Output and Full-Frame Picture-on-Demand

Based on the analyses of ATIS and DAVIS, a new method to achieve the pixel's intensity information for DVS sensor is presented here as an attempt to cope with some drawbacks of the previous works. Instead of integrating the small photocurrent, this sensor directly takes the voltage of front-end logarithmic receptor as the intensity information of pixel. Since this voltage responds to incident light intensity continuously, no additional exposure time is needed and thus it takes no effort to match the fast address-event output with the corresponding intensity information in both spatial and temporal domain. An area-efficient column-parallel SAR ADC array is also adopted to further accelerate the achievement of the brightness information. The adoption of continuous-time logarithmic readout strategy also brings about some other advantages such as small pixel area and wide intra-scene dynamic range. What's more, an external control signal is also added which can force all the pixels to produce request signals and in turn a full-array on-demand reference picture can be obtained. This enables the sensor to output continuous frames besides the event-driven intensity readout mode. The following section first presents the basic idea of this design and also the sensor architecture. Afterwards, detailed design considerations of the whole intensity readout path from in-pixel logarithmic receptor to ADC are elaborated. Finally, the SNR and FPN of the sensor are figured out and followed by some other system design considerations.

3.1 Basic Idea and Sensor Architecture

Fig. 3.1 shows the basic implementation of the new DVS sensor. As mentioned in last chapter, the front-end logarithmic receptor of DVS pixel converts incident light into a logarithmic output voltage (V_{O1}) continuously. This means that V_{O1} has already contains the intensity information of pixel. With respect to the ATIS and DAVIS, this

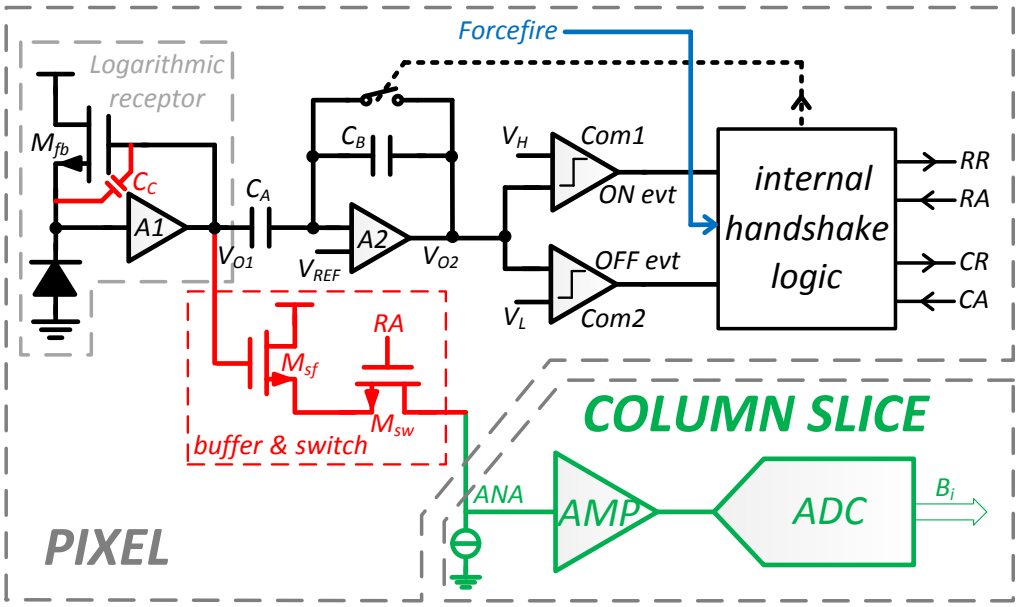


Fig. 3.1 Basic implementation of the new DVS sensor (pixel and column slice).

voltage is only made use of in the DVS part to detect temporal contrast while the intensity information is achieved by means of extra exposure measurement circuit. In contrast, this design directly employs this voltage and reports it outside pixel as its intensity information. The logarithmic receptor, in this way, is not only used in the DVS part to detect relative intensity change but also as a front-end I-V converter to report the absolute intensity information. Since V_{O1} responds to light intensity continuously, the time-consuming exposure procedure adopted in ATIS and DAVIS is no longer required and therefore, the sensor is capable of reporting fast address-events and their corresponding intensity information simultaneously. The minimum address-event latency and time resolution for intensity readout are both in the magnitude of microsecond. The aforementioned undesired discordance between the fast address-events and the postponed intensity output of the ATIS and DAVIS could be settled accordingly. In order to integrate such continuous-time logarithmic-response intensity readout method with DVS, some design modifications are developed at pixel, column and system level. By means of these circuit design improvements, this sensor also possesses some other interesting characteristics such as small pixel area, high intra-scene dynamic range and flexible output configuration.

As to pixel, on the basis of previous DVS design, only two extra transistors M_{sf} and M_{sw} (shown in red in Fig. 3.1) are added to report the logarithmic output voltage V_{O1} from pixel to the column readout bus ANA. M_{sf} works as a source follower to drive

the large capacitance along the readout bus. M_{sw} is an access switch controlled by the row acknowledgement signal (RA). It allows the pixels in the single selected row ($RA = 1$) to report their logarithmic output voltages to the shared column buses and thus avoids readout collisions. The pixel design in this sensor is very compact (corresponding to small pixel area) due to the following two facts. (1) The logarithmic receptor is reused to detect relative intensity change and report the absolute intensity information. (2) In order to implement the latter function, on the one hand, only two transistors (M_{sf} and M_{sw}) are added to the previous DVS design. On the other hand, no extra control signals are needed and thus the bus resources are saved. In addition, since the front-end photoreceptor responds to light intensity in logarithmic manner, wide intra-scene dynamic range (up to 120 dB) could also be achieved for this sensor. Last but not the least, an extra compensation capacitance C_C is also inserted in the logarithmic receptor (as shown in Fig. 3.1) to improve its stability and reduce the output temporal noise. This part will be illustrated later in this chapter.

In column level, as shown in Fig. 3.1, the output voltage from pixel is first conditioned by a switched-capacitor (SC) amplifier and then quantized by the subsequent ADC. One well-known drawback associated with the logarithmic-response image sensor is poor sensitivity at low brightness contrast, which corresponds to a small output voltage swing. This small-swing voltage is amplified to increase the useful input range for ADC. What' more, the SC amplifier also adjusts the common-mode level of pixel output voltage to accommodate the ADC's quantization range. The following SAR ADC is employed to quantize the amplified output voltage into digital codes due to its relatively simple architecture and low power dissipation compared with other ADC architectures. Both the SC amplifier and SAR ADC are implemented in column-parallel architecture, which is compatible with DVS because the fired pixels are also processed on the basis of an entire row. In order to place the SAR ADC into narrow column slice, an area-efficient architecture with significantly simplified capacitance-DAC is adopted. What's more, the column-parallel readout circuits also work in event-based mode, which means that only the output voltages from fired pixels in the selected row are amplified and quantized. As for the normal pixels in this row, the corresponding column readout circuits take no action. This is beneficial to reduce the total power consumption of the readout system.

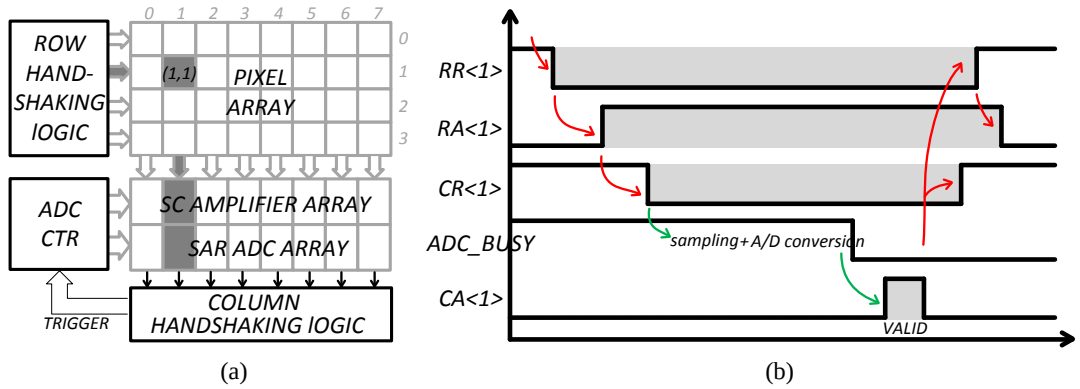


Fig. 3.2 Modified handshaking communication protocol.
(a) Simplified system diagram. (b) Timing diagram to process the fired pixel.

From system level, in order to achieve the address-event and corresponding intensity information simultaneously, the original 4-phase handshaking scheme (of DVS) has been modified a little. Fig. 3.2(a) shows a simplified system diagram with 4×8 pixels, among which the *Pixel (1,1)* is fired and being processed based on the modified handshaking communication protocol shown in Fig. 3.2(b). The fired pixel first sends *RR<1>* (active low) to the external row handshaking logic and after being selected by the row arbiter tree, it will receive *RA<1>* (active high). As shown in Fig. 3.1, the V_{O1} of this fired pixel (actually all pixels in this row) would be transferred to the column readout bus. In the meanwhile, the pixel would issue *CR<1>* in column direction. As for the column handshaking logic circuit in this case, it would hold on the handshaking communication and not immediately acknowledge this request signal. Instead, it would spare some time for the column readout circuit to quantize the logarithmic output voltage from *Pixel (1,1)*. In order to implement this, an additional ADC control logic was incorporated and a control signal ADC_BUSY is introduced and first set to 1 to block the normal handshaking communication in column direction. As demonstrated in Fig. 3.2(b), the falling edge of *CR<1>* will trigger the ADC control logic to generate a series of control signals for the column readout circuits to perform a complete A/D conversion. After that, ADC_BUSY would be pulled down to 0, which recovers the previously blocked acknowledgement to column request. The corresponding *CA<1>* will be issued to process the fired pixel. Since the sensor is self-acknowledged, it would also send a $VALID$ signal to the off-chip processor which would sample the address-event and the corresponding intensity information simultaneously when $VALID$ is high. For the sake of easier interfacing, clock signal is also introduced into the system. It should be pointed out that, in reality, the fired pixels

are processed on the basis of an entire row. After *RA* is issued, all the fired pixels in this row would be processed in sequence with each pixel occupying one clock cycle. The column readout circuits also work in parallel to quantize the output voltages from all the fired pixels of this row at the same time. Since pixel output voltage responds to light intensity continuously, the integration procedure is no longer needed. The delay of this sensor to achieve the brightness information is only the time spent on A/D conversion. Such procedure usually takes sub-μs, which brings about negligible effect to the final address-event latency (usually in the order of several μs). In summary, the event-driven continuous-time intensity readout strategy of this sensor leads to high-speed lossless video compression across focal plane. In addition, for the purpose of further expanding this sensor's functionalities, an extra control signal $Forcefire$ (as shown in Fig. 3.1) is also introduced as a global control signal which can force the whole pixel array to be fired (when $Forcefire$ is activated by the off-chip processor) and therefore to generate a whole image. The addition of this control signal makes the sensor capable of outputting continuous frames, which enriches the applications of this sensor because of the compatibility with current mainstream image processing algorithms.

Based on the idea mentioned above, a prototype DVS sensor with 160×192 pixels was implemented with AMS 0.35 μm CMOS technology and the system architecture is shown in Fig. 3.3. The pixels are connected through row and column handshaking buses and also the column readout bus. The external row control circuits include row handshaking logic (includes row address encoder) and the following row arbiter tree. They are responsible for managing the handshaking communications with pixels in row direction (receiving row request and issuing row acknowledgement). The same set of circuits are also placed in column direction as the column control logic which in charge of the handshaking communication in column direction. In addition, the column-parallel readout circuits including the amplifier and ADC are inserted to quantize the logarithmic output voltages from fired pixels. The relevant control logic is placed at the top of the column arbiter tree to generate the required control signals when a new row is selected. With respect to the output, the sensor could report 8-bit row/column address, 9-bit greyscale data and a valid signal to off-chip processor. The sensor could be configured into three operation modes, namely the analog, digital and progressive mode. It would output continuous frames (analog mode), pure binary

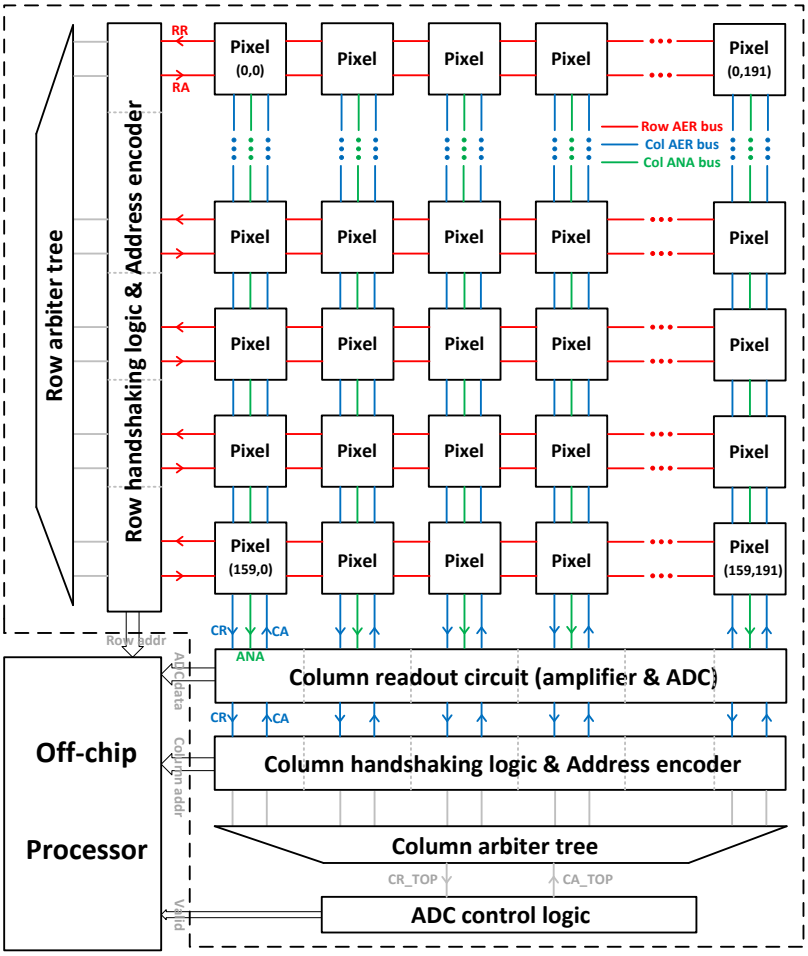


Fig. 3.3 Block diagram of the sensor architecture.

address-events (digital mode) and the combination of address-events and corresponding intensity information (progressive mode) under these operation modes.

Apart from the aforementioned advantages, the continuous-time logarithmic-response image sensor, however, suffers from some intrinsic drawbacks which limit the achieved image quality. This is also the reason why this kind of sensor is not quite popular compared with the integration-mode counterpart [17-19]. As mentioned above, the pixel is not very sensitive to incident light intensity due to the logarithmic compression. In other words, the output signal swing is small for a scene with low brightness contrast. However, the noise performance of the logarithmic receptor is in fact deteriorated because of the increased bandwidth induced by the active feedback loop. As a result, the achievable signal-to-noise ratio (SNR) is limited to a low-level which might be unacceptable for the subsequent image processing algorithms. Another important issue associated with this kind of sensor is the relatively large fixed-pattern-noise (FPN) compared with the 3-T and 4-T APS. The pixel works continuously and

thus the absence of another reset state makes the sensor unable to apply the common correlated double sampling (CDS) technique to mitigate FPN. Therefore, the logarithmic-response sensor is very sensitive to the device parameter variations introduced by the fabrication process. Both of these issues might reduce the attractiveness of adoption of such readout method in DVS and deserve careful considerations. In the following analyses, most of the attentions have been paid to the design of the intensity readout path (front-end logarithmic receptor, column switched-capacitor amplifier and subsequent SAR ADC). They are followed by the analyses of SNR, FPN and other system design considerations. The main objective is to improve the image quality of this logarithmic-response sensor and thus makes it comparable with the integration-mode counterpart.

3.2 Intensity Readout Path Design

3.2.1 Overview

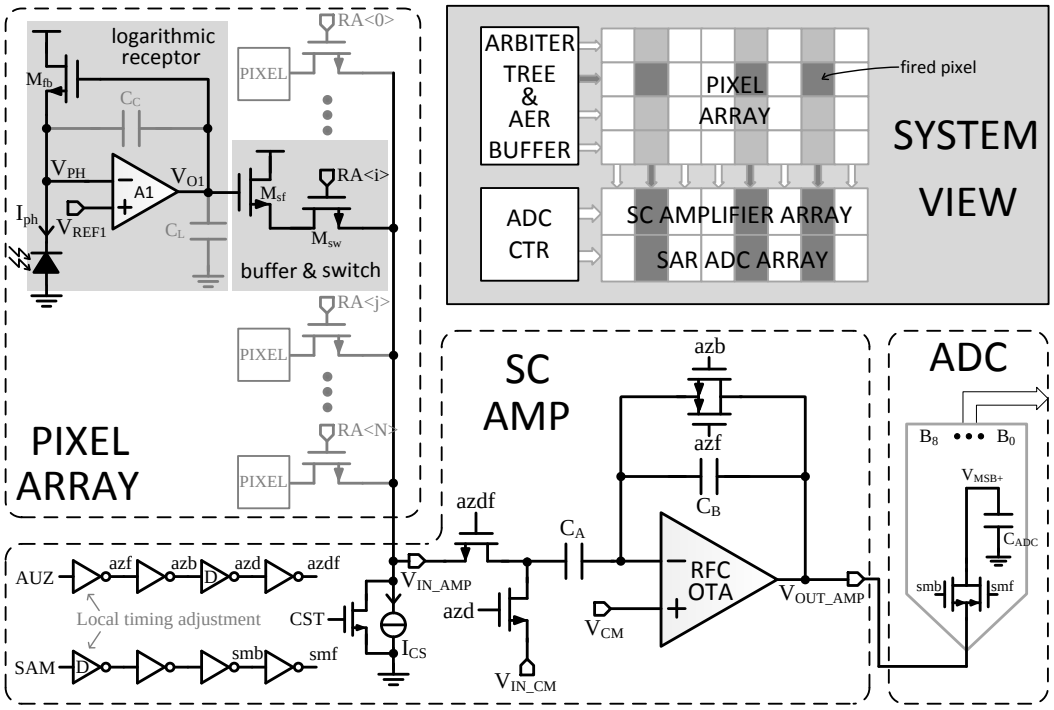


Fig. 3.4 Schematic of a complete intensity readout path.

Fig. 3.4 shows a complete intensity readout path from pixel to ADC. For simplicity, only the front-end logarithmic receptor together with the buffer and switch transistor

(M_{sf} and M_{sw}) are shown in pixel. The remaining parts only achieve the DVS function and have nothing to do with the intensity readout. The switched-capacitor (SC) amplifier and subsequent ADC are implemented in column-parallel architecture. They are responsible for conditioning and quantizing the output voltages from fired pixels in an entire row (as shown in system view in Fig. 3.4). As an interface between pixel and ADC, the SC amplifier has two main functions. On the one hand, since the swing of the pixel's output voltage is relatively small because of the logarithmic compression, it amplifies this voltage and therefore increases the useful input range for ADC. On the other hand, it adjusts the common-mode level of pixel output voltage to accommodate the quantization range of ADC. As the last part along this readout chain, the ADC is in charge of quantizing the amplified voltage into digital codes. In order to place it within a narrow column slice, an area-efficient SAR ADC with significantly simplified capacitance-DAC is adopted in this design. The detailed circuit design and operation principle of ADC will be introduced latter in this section.

The photocurrent I_{ph} changes according to the incident light intensity. As shown in Fig. 3.4, due to the negative feedback of inverting amplifier A1 and M_{fb}, V_{PH} is clamped at V_{REF1}. The feedback transistor M_{fb} works in subthreshold region since I_{ph} is very small (in the magnitude of pA to nA), as a result

$$I_{ph} = I_0 \cdot e^{(V_{GS\_fb} - V_{TH\_fb})/\xi V_T} \tag{3.1}$$

where V_{GS_fb} and V_{TH_fb} are the gate-to-source and threshold voltage of M_{fb}, I_0 is the drain current of M_{fb} when $V_{GS\_fb} = V_{TH\_fb}$, V_T ($= kT/q$) is the temperature voltage, and ξ is the subthreshold slope factor. V_T equals to 26 mV under room temperature and ξ is about 1.3 for the AMS 0.35 μm CMOS technology. Therefore

$$V_{GS\_fb} = V_{TH\_fb} + \xi \cdot V_T \cdot \ln(I_{ph}/I_0) \tag{3.2}$$

and V_{O1} is

$$V_{O1} = V_{REF1} + V_{GS\_fb} = V_{REF1} + V_{TH\_fb} + \xi \cdot V_T \cdot \ln(I_{ph}/I_0) \tag{3.3}$$

it can be seen that V_{O1} responds to I_{ph} in logarithmic manner. For I_{ph} changes from I_{ph1} to I_{ph2}, the swing of V_{O1} is

$$\Delta V_{O1} = V_{O1}|_{I_{ph}=I_{ph2}} - V_{O1}|_{I_{ph}=I_{ph1}} = \xi \cdot V_T \cdot \ln(I_{ph2}/I_{ph1}) \tag{3.4}$$

it only depends on the relative change of the photocurrent (also light intensity) and has nothing to do with V_{REF1} and V_{TH_fb}. For example, if the incident light intensity doubles (corresponds to $I_{ph2}/I_{ph1} = 2$) under room temperature

$$\Delta V_{O1} = \xi \cdot V_T \cdot \ln\left(I_{ph2}/I_{ph1}\right) = 1.3 \cdot 26 \cdot \ln 2 = 23.4\, mV \quad (3.5)$$

The readout of V_{O1} begins when a new row is selected by the row arbiter tree. If the i^{th} row is selected, *RA<i>* would change to 1 while other row acknowledgements still remain 0. V_{O1} of the pixels in this row would be buffered to the shared column readout bus through M_{sf}, whose bias current (I_{CS}) is also shared by one column. After that, the column SC amplifier switches from the reset stage to sampling (amplifying) stage. A local timing adjustment circuit is incorporated to guarantee that the CMOS reset switch across OTA is firstly cut off before the input sampling voltage switches from V_{IN_CM} to V_{IN_AMP}. This ensures that the charge injection of the CMOS switch only contributes a constant offset to V_{OUT_AMP} and could be easily removed by the back-end FPN-cancellation method. The output voltage of the column SC amplifier V_{OUT_AMP} is directly sampled on the capacitance-DAC of the following SAR ADC, which would then convert this voltage into digital codes. At the end of the amplifying stage, V_{OUT_AMP} is

$$V_{OUT\_AMP} = V_{CM} - \frac{C_A}{C_B}\left(V_{IN\_AMP} - V_{IN\_CM}\right) \quad (3.6)$$

where V_{IN_CM} and V_{CM} are the input and output reference-voltage for the SC amplifier. For the sake of a large output swing, V_{CM} is set to around $VDD/2$ (1.6 V). On the other hand, V_{IN_CM} should be adjusted to the middle of the range of V_{IN_AMP} to make sure that V_{OUT_AMP} also centers on V_{CM}. Eliminating the voltage drop across M_{sw}, V_{IN_AMP} is

$$V_{IN\_AMP} = V_{O1} - V_{GS\_sf}$$

$$= V_{REF1} + V_{TH\_fb} + \xi \cdot V_T \cdot \ln\left(I_{ph}/I_0\right) - V_{TH\_sf} - \sqrt{\frac{2I_{CS}}{\mu_n \cdot C_{ox} \cdot S_{sf}}} \quad (3.7)$$

where V_{TH_sf} and S_{sf} are the threshold voltage and aspect ratio (W/L) of M_{sf}, μ_n is the electron mobility and C_{ox} is the unit area gate-oxide capacitance of MOSFET. As a result, V_{OUT_AMP} could be rewritten as

$$V_{OUT\_AMP} = V_{CM}$$

$$-\frac{C_A}{C_B}\left[V_{REF1} + V_{TH\_fb} + \xi \cdot V_T \cdot \ln\left(I_{ph}/I_0\right) - V_{TH_{sf}} - \sqrt{\frac{2I_{CS}}{\mu_n \cdot C_{ox} \cdot S_{sf}}} - V_{IN\_CM}\right] \quad (3.8)$$

Based on this equation, several key points with respect to the intensity readout strategy of this sensor could be emphasized as following. (1) V_{OUT_AMP} is inversely

proportional to $\ln I_{ph}$. This means that brighter pixel (with stronger incident light and larger I_{ph}) leads to smaller V_{OUT_AMP} and the corresponding ADC output code. On the other hand, the image sensor achieves very high dynamic range due to the logarithmic compression of I_{ph} but at the expense of relatively low sensitivity. (2) The SC amplifier provides the gain C_A/C_B. It amplifies the input signal (V_{IN_AMP}) and maps it with the full-scale range (FSR) of the subsequent ADC. In simulation, for per decade increase of I_{ph} under room temperature, V_{O1} increases about 78 mV and V_{IN_AMP} about 65 mV (due to the gain loss of source follower). As a result, for I_{ph} ranges about 6 decades (120 dB dynamic range), the swing of V_{IN_AMP} is about 390 mV. By setting the gain C_A/C_B to 4, the range of V_{OUT_AMP} is a little smaller than 1.6 V, which is close to the FSR of ADC. (3) The DC level of V_{OUT_AMP} could be adjusted by V_{IN_CM}. As mentioned above, in general, V_{IN_CM} should be in the middle of the whole range of V_{IN_AMP}.

Although responds continuously to the light intensity, this kind of logarithmic-response sensor is usually blamed for higher output noise-level (lower SNR) and larger FPN compared with the integration-mode counterpart. As a result, besides the above-mentioned fundamental principle, other design considerations of this logarithmic readout mechanism should be carefully investigated. The following analyses first elaborate some design details of the three parts along the readout path: front-end logarithmic receptor, column SC amplifier and SAR ADC. Based on that, the total readout temporal noise (determines SNR) and voltage mismatches of the readout path (determines FPN) are analysed since these two specifications are crucial to determine the final image quality of the sensor.

3.2.2 Pixel Front-End Logarithmic Receptor

Stability

Because of its internal feedback mechanism, the stability of front-end logarithmic receptor circuit should be first considered. The desired better stability benefits not only the DVS function but also the intensity readout. As the starting point, the small-signal open-loop transfer function of the feedback circuit is first solved. With some reasonable simplifications, the poles and zeros could be extracted for some intuitional

pole-zero analyses. Finally, the simulation results are also provided to further illustrate this issue.

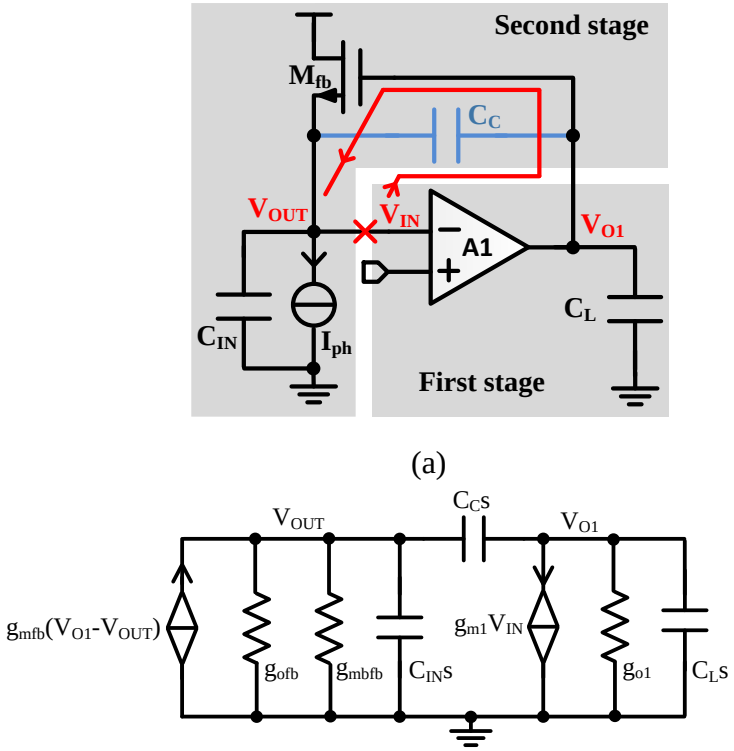


Fig. 3.5 (a) Schematic of logarithmic receptor for stability analysis. (b) Small-signal equivalent circuit.

The schematic of logarithmic receptor is shown in Fig. 3.5(a). The load of the subsequent source follower and all other parasitic capacitances associated with V_{O1} is modelled as a lumped capacitance C_L. In addition, the photodiode is now replaced as an ideal current source I_{ph} in parallel with a capacitance C_{IN}. Besides the equivalent capacitance of photodiode, C_{IN} also includes the input capacitance of $A1$ and the parasitic capacitance of M_{fb}. In addition, capacitance C_C is also shown in Fig. 3.5(a) to represent the total capacitances across the inverting amplifier $A1$. The following analyses and simulation results reveal that this capacitance is of great importance for the stability of the feedback loop. As shown in Fig. 3.5(a), break the loop at the negative input-end of $A1$ and apply an AC small-signal input voltage (V_{IN}) towards one direction, the open-loop transfer function could be calculated as V_{OUT}/V_{IN}. From a certain perspective, this logarithmic receptor circuit (after break the feedback loop) could be regarded as a two-stage amplifier. The first stage is $A1$ while the second stage

is a source follower (M_{fb}) biased by the photocurrent (I_{ph}). The circuit structure is quite simple but the potential issue lies in the second stage since I_{ph} is directly proportional to incident light intensity and thus could range over 6 decades. To keep the feedback circuit remain stable over such large range of I_{ph} should be deliberated. By the way, since all transistors in the logarithmic receptor work in subthreshold region, their respective transconductance is proportional to the corresponding bias current.

By means of the small-signal equivalent circuit shown in Fig. 3.5(b), the open-loop transfer function could be readily solved as following.

For V_{OUT}

$$\left[g_{ofb} + g_{mbfb} + (C_{IN} + C_C)s\right] \cdot V_{OUT} = g_{mfb} \cdot (V_{O1} - V_{OUT}) + C_C s \cdot V_{O1} \quad (3.9)$$

where g_{mfb}, g_{ofb} and g_{mbfb} are the transconductance, output conductance and body transconductance of M_{fb} and $g_{mbfb} = \eta \cdot g_{mfb}$, $\eta < 1$.

For V_{O1}

$$[g_{o1} + (C_L + C_C) \cdot s] \cdot V_{O1} = -g_{m1} \cdot V_{IN} + C_C s \cdot V_{OUT} \quad (3.10)$$

where g_{m1} and g_{o1} are the transconductance and output conductance of the inverting amplifier $A1$. Solve these two equations together and the open-loop transfer function is given by

$$H_{open\_loop}(s) = \frac{V_{OUT}(s)}{V_{IN}(s)} = \frac{-g_{m1}\left(g_{mfb} + C_C s\right)}{I + J \cdot s + K \cdot s^2} \quad (3.11)$$

where

$I = g_{o1}\left(g_{mfb} + g_{mbfb} + g_{ofb}\right)$

$J = g_{o1}C_{IN} + g_{o1}C_C + g_{mbfb}C_C + g_{ofb}C_C + g_{mfb}C_L + g_{mbfb}C_L + g_{ofb}C_L$

$K = C_{IN}C_C + C_{IN}C_L + C_C C_L$

According to this equation, the open-loop gain is

$$A = -\frac{g_{m1}}{g_{o1}} \cdot \frac{g_{mfb}}{g_{mfb} + g_{mbfb} + g_{ofb}} \quad (3.12)$$

in normal condition: $g_{mfb} \gg g_{mbfb}$, g_{ofb} and g_{o1}, so $I \approx g_{o1}g_{mfb}$ and $J \approx g_{mfb}C_L$. The denominator could be simplified as

$$D_{open\_loop}(s) = g_{o1}g_{mfb} + g_{mfb}C_L \cdot s + (C_{IN}C_C + C_{IN}C_L + C_C C_L) \cdot s^2$$
$$\approx g_{o1}g_{mfb}\left(1 - \frac{s}{p_d}\right)\left(1 - \frac{s}{p_{nd}}\right) \quad (3.13)$$

where $p_d = -g_{o1}/C_L$ and $p_{nd} = -g_{mfb}/(C_{IN} + C_C + C_{IN}C_C/C_L)$. This decomposition is only valid under medium and strong light condition as $g_{mfb} \gg g_{o1}$.

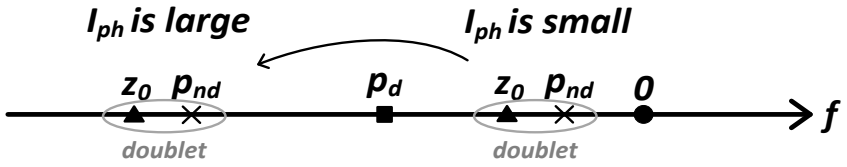


Fig. 3.6 Pole-zero diagram of the open-loop transfer function (z_0 and p_{nd} form a doublet).

After simplification, $H_{open\_loop}(s)$ now consists of two real poles (p_d and p_{nd}) and one zero ($z_0 = -g_{mfb}/C_C$), all located at left-hand-plane (LHP). p_d is the pole introduced by $A1$ and independent of I_{ph}. While p_{nd} and z_0 form a pole-zero doublet because they are both proportional to g_{mfb}, which is in turn proportional to I_{ph}. Such relationship is demonstrated in Fig. 3.6 as both p_{nd} and z_0 locate at low frequency with small I_{ph} and would be pushed to higher frequency when I_{ph} becomes larger. The relative position of z_0 and p_{nd} remains constant regardless of the incident light intensity as

$$\alpha = \frac{z_0}{p_{nd}} = 1 + \frac{C_{IN}}{C_C} + \frac{C_{IN}}{C_L} > 1 \tag{3.14}$$

α is defined as the ratio of z_0 to p_{nd} and used to characterize the closeness of this pole-zero doublet. As illustrated in (3.14), with constant C_{IN} and C_L, α is negatively correlated with C_C. A smaller C_C leads to a larger α which means z_0 separates from p_{nd} (a remote doublet), while a larger C_C results in a smaller α which means z_0 approximates to p_{nd} (a close doublet). In order to explain the influence of α on the stability of circuit, two extreme situations are considered here as (1) α is very large for $C_C \ll C_{IN}$; (2) α is very close to 1 for $C_C \gg C_{IN}$ (eliminate the last item C_{IN}/C_L). With respect to the first condition, z_0 locates at very high frequency and could be ignored accordingly. This leaves the system with two real poles p_d and p_{nd}, whose relative location would determine the stability of the feedback loop. Fig. 3.7 shows the positions of these two poles under this condition where GBW is the gain-bandwidth product calculated as $GBW = |A \cdot p_d| \approx g_{m1}/C_L$. For open-loop analysis, the phase margin (PM) is a commonly used criterion to evaluate the stability issue. It is defined as the phase difference between -180° and the phase-response of open-loop transfer

function at GBW. In general, the PM should be at least 45°, or more strictly 60°, to ensure the stability of the feedback loop. As for this scenario with only two poles, PM is 45° when p_{nd} equals GBW while below that the system would be unstable ($PM < 45^{\circ}$ when $p_{nd} < GBW$). By means of the aforementioned deductions, this boundary condition corresponds to I_{ph} equals $(C_{IN}/C_L) \cdot I_1$, where I_1 is the bias current of input transistor of $A1$ and usually in the magnitude of several decades of nA. Unfortunately, even under strong light condition, this requirement is still difficult to satisfy since I_{ph} is at most several nA. This means that the front-end logarithmic receptor is unstable when C_C is very small. With respect to the other condition, α approximates to 1 when $C_C \gg C_{IN}$. This leads to a very closed doublet of z_0 and p_{nd}. Their effects would always counteract with each other regardless of I_{ph}, which results in an approximate single-pole system (only p_d is left). As a result, this feedback circuit is unconditionally stable under this situation.

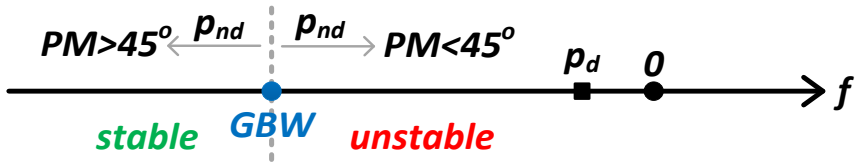


Fig. 3.7 Pole-zero diagram for stability analysis when α is very large.

Based on the above analysis, C_C is crucial to the stability of the front-end logarithmic receptor circuit. This conclusion is also verified by circuit simulations (Cadence Spectre) with the relevant results illustrated in Fig. 3.8. The simulation setup is as following: C_C is set to several constant values (0/5/10/15/20/50 fF), while I_{ph} is scanned from 3 pA to 5 nA under stability simulations. The PM is plotted to evaluate the stability issue under different conditions. Other parameters used in the simulations are $C_{IN} = 30\ fF$, $C_L = 60\ fF$, $I_1 = 20\ nA$. In addition, the size of feedback transistor M_{fb} is 1.5 $\mu m/1.5\ \mu m$. It should be pointed out that when I_{ph} is very small (several pA), the above analysis is no longer applicable since the decomposition of equation (3.13) is invalid. Under this condition, the dominant pole is determined by I_{ph} which locates at very low frequency. Therefore, the influences of other relatively high-frequency poles and zeros could be negligible with respect to the low-frequency GBW and the circuit is stable regardless of C_C. With medium and larger I_{ph}, it is evident in

Fig. 3.8 that the feedback loop is more stable with larger C_C, which is in accordance with the previous analysis. To be more specific, as for 50 fF C_C, the feedback loop enjoys excellent stability as PM remains higher than 75° for the whole range of I_{ph}. In contrast, when C_C is set to 0 (actually 0.5 fF due to the parasitic capacitances of M_{fb} and $A1$), the circuit is highly unstable as PM is less than 40°. In order to keep PM always larger than 45° for the whole range of I_{ph}, C_C should be at least 5 fF.

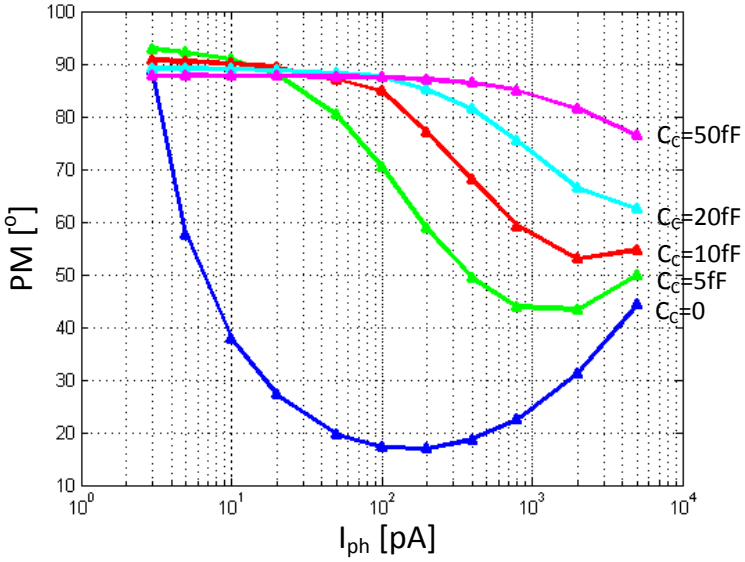


Fig. 3.8 Simulation results of PM with respect to different C_C.

The good stability of the front-end logarithmic circuit over whole range of I_{ph} is of great importance for this design. On the one hand, it ensures smooth first-order close-loop frequency response of the receptor without obvious peaking at certain frequency point. As a result, all the noise components are attenuated at high frequency, instead of being amplified at the peaking frequency of unstable circuit. This helps to reduce the total output noise of the logarithmic receptor. On the other hand, the good stability also indicates good immunity of the pixel from external interferences. Such interferences could be either the coupling from power supply, substrate or other digital handshaking buses across the pixels. The feedback loop could attenuate the induced interference spikes and return back to steady state gradually. The good immunity allows the logarithmic receptor to directly drive the subsequent difference amplifier (for DVS function) and source follower (for intensity readout) without incorporating another buffer circuit as adopted in [3]. This simplifies the pixel design and saves the

corresponding power consumption. With respect to selection of C_C, the desired stability of the feedback loop requires larger C_C. However, this could reduce the bandwidth of the logarithmic receptor as analysed in the next part. Therefore, the selection of C_C should be based on overall considerations of the stability and bandwidth (also the noise performance).

Bandwidth

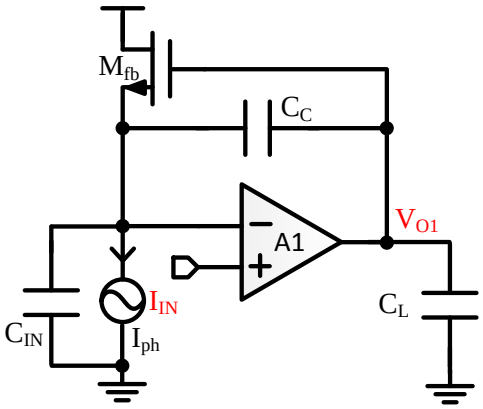


Fig. 3.9 Schematic of logarithmic receptor for bandwidth analysis.

This subsection analyses the small-signal bandwidth (BW) of the logarithmic receptor. In general, larger BW is desired to enable the DVS to detect fast motion with short latency. Fig. 3.9 shows the corresponding schematic to analyse BW. In this case, the input is an AC current (I_{IN}) from photodiode and the transfer function from I_{IN} to V_{O1} is

$$H(s) = \frac{V_{O1}(s)}{I_{IN}(s)} = \frac{g_{m1} - C_C s}{D_{close\_loop}(s)} \tag{3.15}$$

$D_{close\_loop}(s)$ is the denominator of the close-loop transfer function

$$D_{close\_loop}(s) = L + M \cdot s + N \cdot s^2 \tag{3.16}$$

where

$L = g_{m1}g_{mfb} + g_{o1}(g_{mfb} + g_{mbfb} + g_{ofb})$

$M = g_{o1}C_{IN} + g_{o1}C_C + g_{m1}C_C + g_{mbfb}C_C + g_{ofb}C_C + g_{mfb}C_L + g_{mbfb}C_L + g_{ofb}C_L$

$N = C_{IN}C_C + C_I C_L + C_C C_L$

in normal condition, $g_{o1}, g_{ofb} \ll g_{m1}, g_{mfb}, g_{mbfb}$ and $g_{mfb} \ll g_{m1}$, $D_{close\_loop}(s)$ could be simplified as

$$D_{close\_loop}(s) = g_{m1}g_{mfb} + g_{m1}C_C \cdot s + (C_{IN}C_C + C_IC_L + C_CC_L) \cdot s^2$$
$$\approx g_{m1}g_{mfb}(1 - s/p_d)(1 - s/p_{nd}) \quad (3.17)$$

where $p_d = -g_{mfb}/C_C$ and $p_{nd} = -g_{m1}/(C_{IN} + C_L + C_{IN}C_L/C_C)$. In addition, $H(s)$ also has one right-hand-plane (RHP) zero as $z_0 = g_{m1}/C_C$. Since both p_{nd} and z_0 are proportional to g_{m1} and therefore located at high frequency, the BW of the logarithmic receptor circuit approximately equals to p_d as

$$BW \approx |p_d| = g_{mfb}/C_C = I_{ph}/(\xi \cdot V_T \cdot C_C) \quad (3.18)$$

It could be seen that BW is inversely proportional to C_C, which is supported by the simulation results as shown in Fig. 3.10. All the simulation setups are same to the previous stability simulation. It is obvious that the logarithmic receptor enjoys largest bandwidth with smallest 5 fF C_C. As a result, in order to broaden the bandwidth, C_C should be smaller. However, this not only deteriorates the stability of the feedback loop but also introduces more output temporal noises as described in the next subsection.

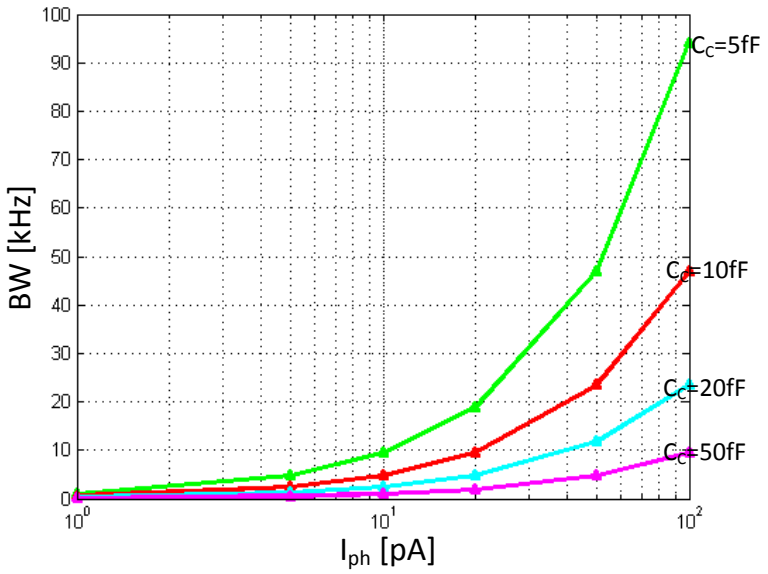


Fig. 3.10 Simulation results of bandwidth with respect to different C_C.

Output Noise

The noises of front-end logarithmic receptor account for a large proportion of the total output temporal noise of the intensity readout path. They should be kept as low as possible to enhance some crucial specifications of the sensor such as SNR and

minimum detectable temporal contrast. Fig. 3.11 shows the schematic for noise analysis which considers two independent noise sources: the input-referred noise of inverting amplifier $A1$ (V_{n1}) and noise from the feedback transistor M_{fb} (V_{n2}). Most of the output noises are contributed by thermal noise and therefore, the analysis only focuses on this kind of noise and neglects the others (such as flicker noise and so on).

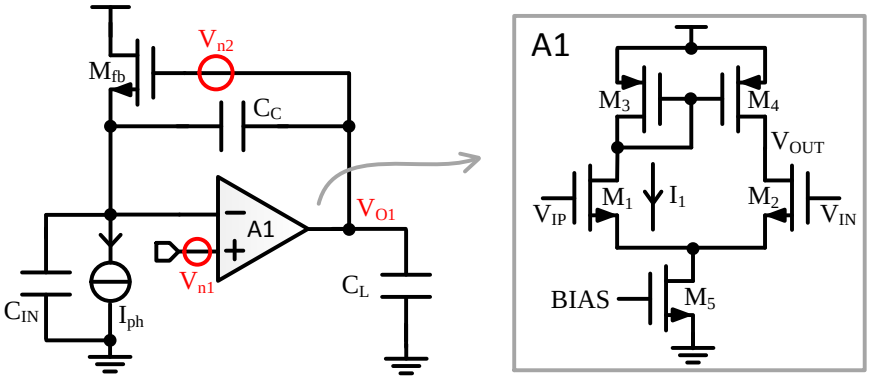


Fig. 3.11 Schematic of logarithmic receptor for noise analysis.

Based on the classic noise analysis method, the transfer function from V_{n1} and V_{n2} to the output V_{O1} is first solved as

$$H_1(s) = \frac{V_{O1}(s)}{V_{n1}(s)} = \frac{g_{m1}\left[g_{mfb} + g_{mbfb} + g_{ofb} + (C_{IN} + C_C)s\right]}{D_{close\_loop}(s)} \tag{3.19}$$

$$H_2(s) = \frac{V_{O1}(s)}{V_{n2}(s)} = \frac{g_{mfb}(g_{m1} - C_C s)}{D_{close\_loop}(s)} \tag{3.20}$$

as analysed in the previous subsection, $D_{close\_loop}(s) = g_{m1}g_{mfb}(1 - s/p_d)(1 - s/p_{nd})$ where $p_d = -g_{mfb}/C_C$ and $p_{nd} = -g_{m1}/(C_{IN} + C_L + C_{IN}C_L/C_C)$.

Firstly, $H_1(s)$ could be rewritten as

$$H_1(s) = G_1 \cdot \frac{(1 + s/z_0)}{(1 + s/p_d)(1 + s/p_{nd})} \tag{3.21}$$

where

$$G_1 = \frac{g_{m1}(g_{mfb} + g_{mbfb} + g_{ofb})}{g_{m1}g_{mfb} + g_{o1}(g_{mfb} + g_{mbfb} + g_{ofb})} \approx 1 + \frac{g_{mbfb} + g_{ofb}}{g_{mfb}}$$

and

$$z_0 = -\frac{g_{mfb} + g_{mbfb} + g_{ofb}}{C_{IN} + C_C}$$

It is evident that $z_0 < p_d \ll p_{nd}$. z_0 and p_d constitute a low-frequency pole-zero doublet while p_{nd} is a high-frequency non-dominant pole. The amplitude-

frequency response of $H_1(s)$ is shown in Fig. 3.12. At very low frequency ($f < z_0$), $|H_1(s)|$ equals to G_1. At medium-frequency ($p_d < f < p_{nd}$), it is pulled up by the low-frequency pole-zero doublet (z_0 and p_d) to G_M and begins to drop by $-20dB/decade$ at frequency higher than p_{nd}. The medium-frequency gain G_M is given by

$$G_M \approx G_1 \cdot \frac{p_d}{z_0} = \frac{g_{mfb} + g_{mbfb} + g_{ofb}}{g_{mfb}} \cdot \frac{g_{mfb}}{C_C} \cdot \frac{C_{IN} + C_C}{g_{mfb} + g_{mbfb} + g_{ofb}}$$
$$= 1 + C_{IN}/C_C \quad (3.22)$$

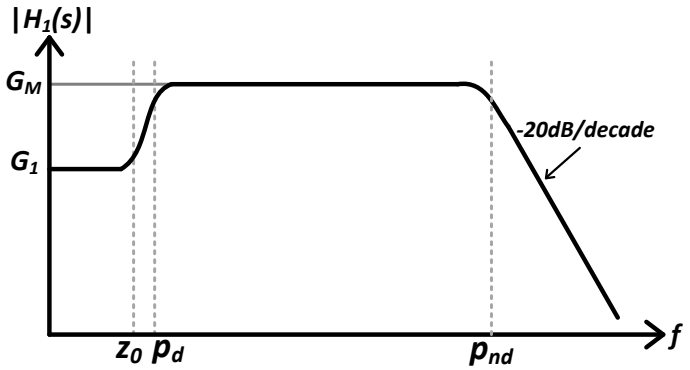


Fig. 3.12 Amplitude-frequency response of $H_1(s)$.

In order to compute output noise, the integral of $|H_1(s)|^2$ over the entire frequency range should be first calculated. Unfortunately, even for simple $H_1(s)$ with a low-frequency doublet and a high-frequency pole, such computation is still too complicated to be finished by hand. Using the MATLAB to accomplish this task is easy but ignores some insights and intuitionistic knowledges. In order to meet both requirements simultaneously, $H_1(s)$ has been simplified as a first-order low-pass system with low-frequency gain G_M and a single-pole p_{nd}. This simplification is valid because the pole-zero doublet locates at frequency much lower than p_{nd}. The calculation result is a little conservative (because G_M is larger than G_1 at frequency below the doublet) but with more conciseness and insight.

As for V_{n1}, the simple 5-transistor differential amplifier is used as shown in Fig. 3.11 and thus

$$\overline{V_{n1}^2} = 2\left(\overline{V_{M1}^2} + \overline{V_{M3}^2}\right) = 8\,qI_1/g_{m1}^2 = 8qI_1/(I_1/\xi V_T)^2 = 8q\xi^2 V_T^2/I_1 \quad (3.23)$$

the root-mean-square (RMS) noise of V_{o1} induced by V_{n1} is given by

$$V_{o1\_rms\_vn1} = \sqrt{\int_0^\infty \overline{V_{n1}^2} \cdot |H_1(j2\pi f)|^2 df} = \sqrt{\overline{V_{n1}^2}} \cdot \sqrt{\int_0^\infty |H_1(j2\pi f)|^2\, df} \quad (3.24)$$

with the above-mentioned simplification, $H_1(s)$ is now simplified as a first-order low-pass system

$$H_1(s) = \frac{G_M}{(1 + s/p_{nd})} \tag{3.25}$$

then

$$\int_0^\infty |H_1(j2\pi f)|^2 \, df = G_M^2 \cdot (p_{nd}/4) \tag{3.26}$$

now, $V_{o1\_rms\_vn1}$ could be derived as

$$\begin{aligned} V_{o1\_rms\_vn1} &= \sqrt{\overline{V_{n1}^2}} \cdot \sqrt{G_M^2 \cdot (p_{nd}/4)} \\ &= \sqrt{2\xi} \cdot (1 + C_{IN}/C_C) \cdot \sqrt{kT/(C_{IN} + C_L + C_{IN}C_L/C_C)} \end{aligned} \tag{3.27}$$

Secondly, the RMS noise of V_{o1} induced by the feedback transistor M_{fb} could also be analysed based on the same method. $H_2(s)$ could be decomposed as

$$H_2(s) = G_2 \cdot \frac{(1 + s/z_0)}{(1 + s/p_d)(1 + s/p_{nd})} \tag{3.28}$$

where

$$G_2 = \frac{g_{m1}g_{mfb}}{g_{m1}g_{mfb} + g_{o1}(g_{mfb} + g_{mbfb} + g_{ofb})} \approx 1$$

and

$$z_0 = \frac{g_{m1}}{C_C}$$

In this case the zero locates at high-frequency as $p_d \ll p_{nd} < z_0$. p_{nd} and z_0 constitute a high-frequency pole-zero doublet while p_d is a low-frequency dominant pole. For the convenience of calculation, $H_2(s)$ could also be simplified as a first-order low-pass system with a single-pole located at p_d.

M_{fb} works in subthreshold region and is biased by I_{ph}, so

$$\overline{V_{n2}^2} = 2qI_{ph}/g_{mfb}^2 = 2q\xi^2 V_T^2/I_{ph} \tag{3.29}$$

As a result, the RMS noise of V_{o1} induced by V_{n2} is

$$V_{o1\_rms\_vn2} = \sqrt{\int_0^\infty \overline{V_{n2}^2} \cdot |H_2(j2\pi f)|^2 df} = \sqrt{\overline{V_{n2}^2} \cdot} \sqrt{\frac{p_d}{4}} = \sqrt{\xi/2} \cdot \sqrt{kT/C_C} \tag{3.30}$$

Since V_{n1} and V_{n2} are independent with each other, the total RMS noise of V_{o1} is

$$V_{o1\_rms} = \sqrt{V_{o1\_rms\_vn1}^2 + V_{o1\_rms\_vn2}^2} \tag{3.31}$$

with $V_{o1\_rms\_vn1}$ and $V_{o1\_rms\_vn2}$ are shown in (3.27) and (3.30) respectively.

According to above analysis, the output noise of logarithmic receptor only depends on the capacitances C_{IN}, C_C and C_L. With respect to $V_{o1\_rms\_vn1}$, its equivalent noise capacitance C_{EQ} is

$$C_{EQ} = C_{IN} + C_L + C_{IN}C_L/C_C \tag{3.32}$$

with constant C_{IN} and C_L, this capacitance become smaller with larger C_C. However, the total noise performance is improved in this case since it also provides a smaller noise gain $1 + C_{IN}/C_C$. On the other hand, $V_{o1\_rms\_vn2}$ is directly inversely proportional to $\sqrt{C_C}$ and therefore becomes smaller for a larger C_C. As a result, the total noise performance of the front-end logarithmic receptor will be improved by enlarging C_C. This conclusion is verified by the simulation results which are shown in Fig. 3.13. Some basic parameters are $C_{IN} = 30\ fF$, $C_L = 60\ fF$, $I_1 = 20\ nA$. Under the smallest I_{ph} (3 pA), the simulated V_{o1_rms} is 3.90 mV_{rms} ($C_C = 0$), 1.38 mV_{rms} ($C_C = 5\ fF$), 1.03 mV_{rms} ($C_C = 10\ fF$), 0.78 mV_{rms} ($C_C = 20\ fF$), 0.57 mV_{rms} ($C_C = 50\ fF$). When C_C is 0, the logarithmic receptor shows very poor noise performance, which would be improved as C_C becomes larger. In addition, V_{o1_rms} is almost independent with photocurrent I_{ph} (except $C_C = 0$), which is in agreement with the above analysis. Besides the simulation results, the calculated V_{o1_rms} according to (3.31) are also shown as grey lines in Fig. 3.13. They are very close to the simulation results with only around 5 % errors and this demonstrates the reasonability of the previous deductions.

According to the aforementioned analyses of stability, bandwidth and noise performance of the front-end logarithmic receptor, C_C is of great importance and chosen as 10 fF in this design to provide a better tradeoff. According to Fig. 3.8, the 10 fF C_C ensures PM always larger than 55° for the whole range of I_{ph}, which indicates good stability. In addition, the BW of the logarithmic receptor is about 24 kHz under 50 pA I_{ph} and the total output temporal noise is about 1.03 mV_{rms}.

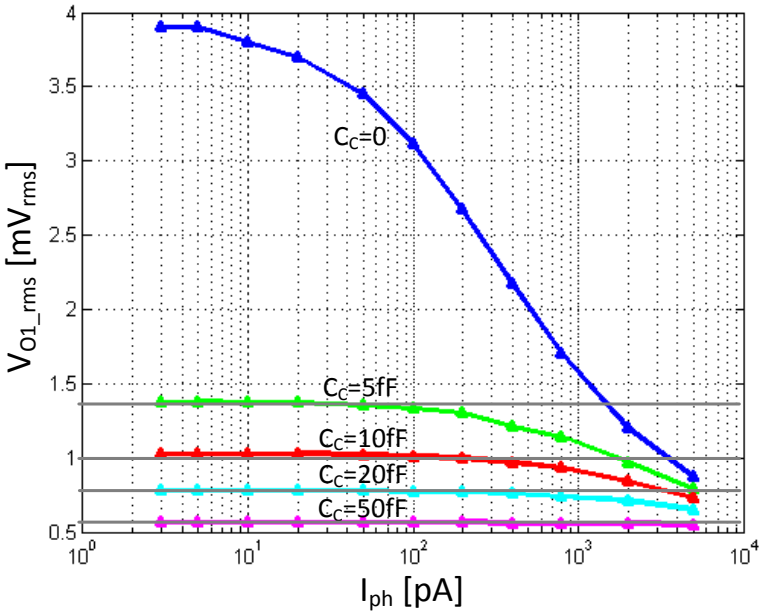


Fig. 3.13 Simulation results of V_{o1_rms} with respect to different C_C.

3.2.3 Column SC Amplifier

As mentioned before, the column SC amplifier is incorporated as the interface between pixel and subsequent ADC. It not only broadens the range of pixel's output voltage but also adjusts the corresponding DC level. Fig. 3.14 shows the schematic and timing diagram of the column SC amplifier. It works alternatively between the reset and amplifying mode. During reset mode ($SAM = 0/AUZ = 1$), the CMOS reset switch across OTA turns on and clamps V_{OUT_AMP} at V_{CM}. At the input end, the SC amplifier samples V_{IN_CM} through $M2$. The control signal SAM and AUZ are generated by global ADC control logic above the column arbiter tree. When a new row is selected, the column requests from the fired pixels in this row would be transferred to the top of column arbiter tree. Afterwards, the ADC control logic would trigger a complete A/D conversion process by setting SAM to 1 and AUZ to 0. The SC amplifier (only in columns contain fired pixels) then enters into the amplifying mode by first turning off the CMOS reset switch and then switching the input sampling voltage from V_{IN_CM} to V_{IN_AMP}. A local timing adjustment circuit is incorporated to make sure such sequence requirement. After V_{OUT_AMP} settles down as shown in equation (3.33), it is directly sampled on the capacitance-DAC of the subsequent SAR ADC which then converts this voltage into digital codes.

$$V_{OUT\_AMP} = V_{CM} - \frac{C_A}{C_B}\left(V_{IN\_AMP} - V_{IN\_CM}\right) \tag{3.33}$$

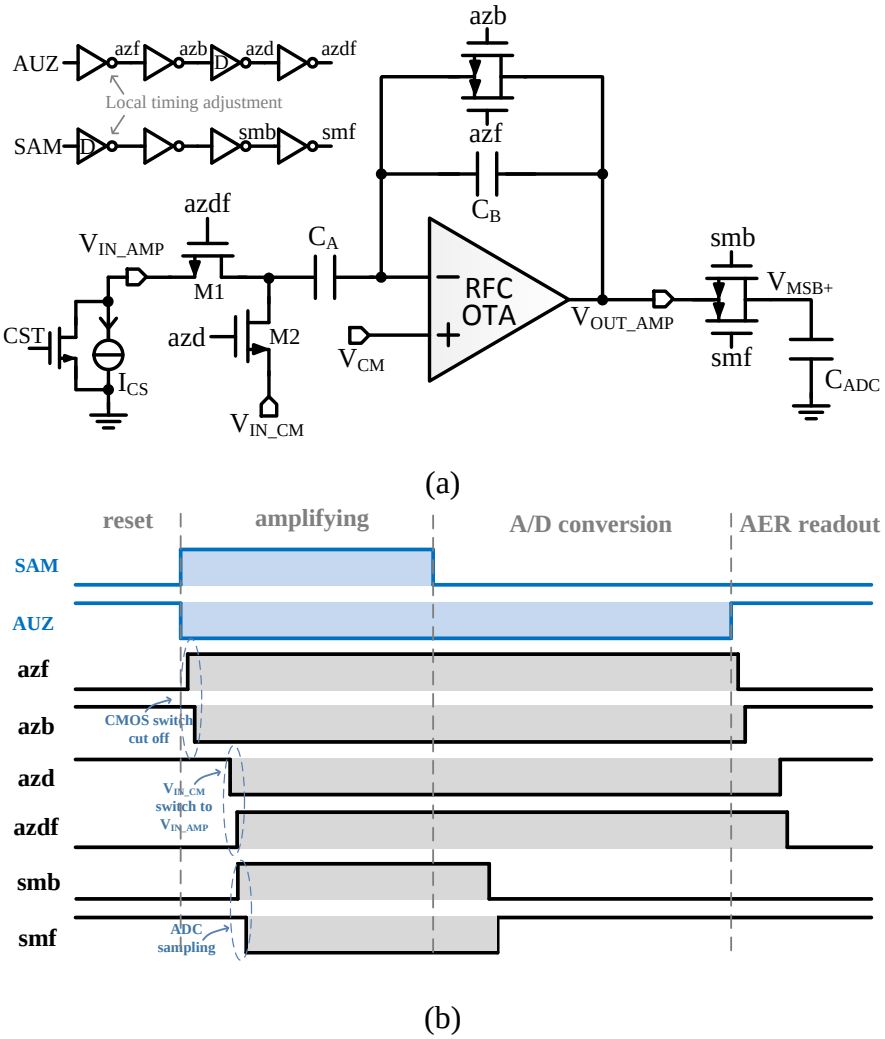


Fig. 3.14 (a) Schematic of the column SC amplifier. (b) Timing diagram.

As the requirement of following SAR ADC, V_{CM} is set to around $VDD/2$ (1.6 V in this design). The input reference-voltage V_{IN_CM} should be adjusted to the middle of the range of V_{IN_AMP} according to different light condition. By doing this, V_{OUT_AMP} would be distributed uniformly across the full scale range of ADC (instead of overflow). The gain provided by SC amplifier is C_A/C_B and chosen as 4 in this design ($C_A = 400\ fF$, $C_B = 100\ fF$). The column current source (I_{CS}) to provide bias current for in-pixel source follower is chosen as 3 μA. In order to accelerate the amplifying procedure, an extra switch controlled by CST is added in parallel with I_{CS}. It provides a low-resistive discharging path for the large parasitic capacitance along the column readout bus between switching of two different rows. It resets V_{IN_AMP} to GND and thus shortens

the amplifying procedure since the speed of in-pixel source follower to charge this capacitance is much higher than discharge it. Another important specification of the SC amplifier is bandwidth, which determines the required time spent for the amplifying stage. It could be expressed as

$$BW = \frac{C_B}{C_A + C_B} \cdot \frac{g_m}{C_{ADC} + \frac{C_A C_B}{C_A + C_B}} = \frac{g_m}{C_{EQ}} \tag{3.34}$$

and

$$C_{EQ} = C_A + C_{ADC} + \frac{C_A C_{ADC}}{C_B} \tag{3.35}$$

where g_m is the transconductance of OTA and C_{ADC} the load capacitance of subsequent ADC. With $C_A = 400\,fF$, $C_B = 100\,fF$ and $C_{ADC} = 400\,fF$, the equivalent capacitance C_{EQ} driven by OTA during amplifying stage is about 2.4 pF.

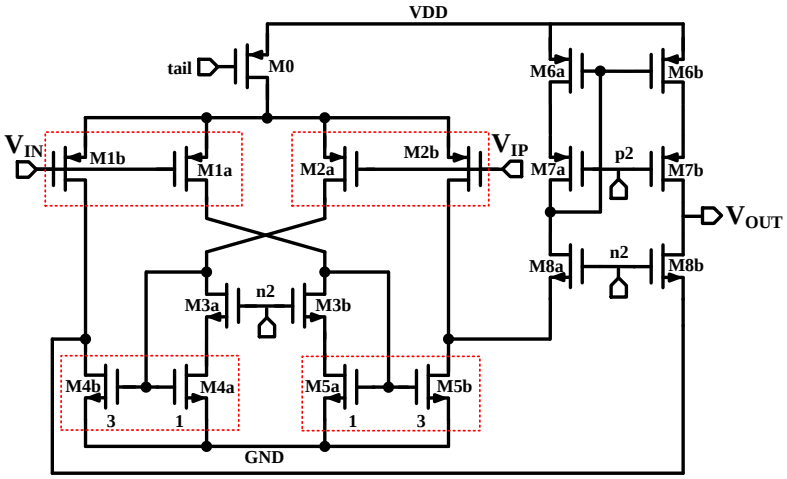


Fig. 3.15 Schematic of the RFC OTA.

As the core of SC amplifier, the OTA adopted in this design is the recycling folded cascode (RFC) OTA [20], which is an improvement to the conventional folded cascode (FC) OTA. The schematic of RFC OTA is shown in Fig. 3.15. With the same power consumption and chip area, RFC OTA enjoys larger DC gain, GBW and slew rate compared with the traditional FC OTA. The main modifications to the conventional topology lie in two aspects. Firstly, each transistor of the original input differential pair is split into two identical smaller transistors ($M1a/M1b$ and $M2a/M2b$). Secondly, the original tail transistors to provide bias current for the amplifier are also split into two current amplifiers ($M4a/M4b$ and $M5a/M5b$) with typical gain

of 3. The advantages of RFC OTA over the conventional topology mainly come from its enhancement of the equivalent input transconductance. Fig. 3.16 shows the comparison of the input stage of FC and RFC OTA. For the FC topology, as shown in Fig. 3.16(a), the transconductance of input transistor $M1$ and $M2$ is g_m. Applying a differential input voltage ($V_{IP} = -V_{IN} = 0.5V_{diff}$), the small-signal AC currents generated from these two transistors ($0.5g_mV_{diff}$) are directly added at the output node. As a result, the equivalent input transconductance of FC OTA is g_m. In contrast, the RFC OTA divides the original input transistor into two identical smaller transistors ($M1a/M1b$ and $M2a/M2b$). Each transistor has $0.5g_m$ tranconductance and generates $0.25g_mV_{diff}$ small-signal current in differential mode. The currents generated by $M1b$ and $M2a$ are added at node A and then transferred to the output node. Due to the multiplication effect of current amplifier ($M4a/M4b$), the small-signal current transferred to the output node by this branch is g_mV_{diff} ($0.25g_mV_{diff}$ from $M1b$ and $0.75g_mV_{diff}$ from $M2a$ after amplified by 3). Similarly, the other branch consists of $M2b$ and $M1a$ also produces the same small-signal current g_mV_{diff}. Consequently, the equivalent input transconductance of RFC amplifier is $2g_m$, which is double of that of the FC amplifier. This doubled equivalent input transconductance of RFC OTA results in larger DC gain and GBW. For the same GBW, the RFC amplifier consumes only half DC current of that of the FC amplifier. Moreover, the large-signal slew rate is also enhanced to twice. However, the tradeoff is that the adoption of current amplifier brings about a lower frequency non-dominant pole along

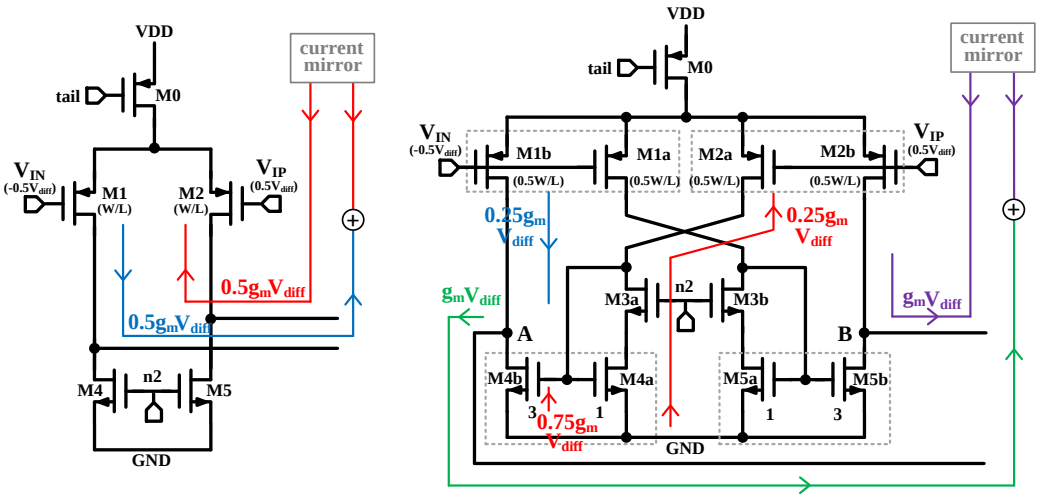


Fig. 3.16 Input stage of (a) Conventional FC OTA and (b) RFC OTA.

the small-signal path, which would degrade the PM to some extent. Fortunately, this is fully acceptable in this application with GBW around 10 MHz. In addition, the input-referred offset voltage and noise power also increase a little compared with that of the FC OTA. The simulation results of the RFC OTA are summarized in TABLE 3.1. It consumes 20 μA DC current and achieves 83 dB low-frequency gain. With 2.4 pF load capacitance, the GBW and PM of the OTA are 10.7 MHz and 79.4° respectively. The output swing is larger than 2 V (from 0.5 V to 2.7 V), which is enough to cover the FSR of the subsequent ADC.

TABLE 3.1 Performance summary of RFC OTA

DC gain	83 dB
Load capacitance (equivalent)	2.4 pF
GBW	10.7 MHz
PM	79.4°
Slew rate	10.5 $V/\mu s$
Output swing*	0.5-2.7 V
Input-referred offset voltage (σ)	5.5 mV
Input-referred noise power	33 $nV/\sqrt{Hz}$
1 % Setting time (10 mV step)	58.3 ns
1 % Setting time (1 V step)	148.2 ns
DC current	20 μA

*Referred to DC gain larger than 70 dB

3.2.4 Column-Parallel SAR ADC

Introduction and Basic Idea

Besides the continuous-time logarithmic readout strategy of pixel intensity information, this design also adopts the column-parallel ADC array to further accelerate the achievement of brightness information. The employment of a large number of ADCs working together to quantize the integration voltage of pixels for an entire row is a common practice to increase the speed of frame-based APS. The column-parallel ADC architecture is also compatible with DVS because the events are also processed on the

basis of an entire row. This is favourable for insertion of an ADC array to execute A/D conversion once for all the fired pixels of the selected row before their relevant addresses are reported out. With respect to the back-end processor, it could obtain the address-event and its corresponding brightness information concurrently.

Since a large number of ADCs are used in parallel, the hardware overhead and power consumption of an individual ADC are very important for the whole readout system. Several ADC architectures which are fit for column-parallel applications include the single-slope, cyclic and successive approximation register (SAR) ADC [21]. The single-slope ADC is hardware efficient but at the expense of long conversion time. The cyclic ADC works faster by means of a switched-capacitor circuit to implement the efficient binary search algorithm. However, the requirement for a high-performance OTA usually burns lots of power consumption. Relatively speaking, SAR ADC is a better choice with its combination of such advantages as relatively high conversion speed and low power dissipation. Instead of using switched-capacitor amplifier, mainstream SAR ADC usually exploits capacitance-DAC (cap-DAC) followed by comparator to implement the binary search algorithm. Lots of SAR ADC designs which have little requirement to chip area prefer to use the conventional binary cap-DAC for its better linearity characteristic [22-24]. They rely on a set of binary-weighted branch capacitors to adjust the input sampling voltage successively with a series of binary-scaled voltage steps. For an ADC with 9-bit resolution, the binary cap-DAC requires 512 (2^9) unit capacitances in total. This number would be even doubled if the differential architecture is adopted to suppress the common-mode noises from power supply and substrate. As for column-parallel implementation, however, it is very difficult to place and connect such large number of capacitances within a narrow column slice but still with good matching characteristic. As a result, some capacitance reduction methods should be employed to meet this strict layout requirement. Based on current knowledges, at least three such kinds of method could be found as elaborated below.

(1) Split cap-DAC technique: The original binary cap-DAC could be split into two smaller binary cap-DAC connected through a bridge capacitor. For example, a 8-bit split cap-DAC could be composed of two sub 4-bit binary cap-DACs connected through an unit bridge capacitor. The number of total capacitors is reduced to 33 ($2^4 + 2^4 + 1$), which is much smaller than that of the 8-bit binary cap-DAC (256).

The cost of such reduction, however, is that the split-cap DAC is more sensitive to the parasitic capacitances associated with the MSB and LSB node. In addition, the mismatches of the sub cap-DACs also have larger impact on the linearity of ADC [25].
(2) Top-plate sampling technique: Switching the classic bottom-plate sampling strategy to top-plate sampling could also bring about the reduction of total capacitors [26]. In this scheme, the input voltage is exactly sampled on the top plate of cap-DAC and therefore, the comparator could directly perform the first comparison which determines the most significant bit (MSB) without switching any capacitors. As a result, an *(N-1)*-bit binary cap-DAC is enough for the implementation of a *N*-bit SAR ADC, which halves the number of total capacitors accordingly. However, the top-plate sampling switch would introduce input-dependent distortion when turns off and therefore restricts the achievable resolution of ADC to 10-bit or less.
(3) Scaled reference-voltage technique: Most of the cap-DAC designs count on capacitors alone to accomplish a series of binary-scaled voltage steps for bit conversion. However, this objective could also be realized by a set of binary-scaled reference-voltages. For example, a desired voltage step $1/2 \cdot C \cdot \Delta V$ could be either comprehended as $(C/2) \cdot \Delta V$ or equivalently $(\Delta V/2) \cdot C$. The former could be implemented by a constant reference-voltage step (ΔV) applied to a binary-scaled capacitor $(C/2)$, while the latter by a binary-scaled reference-voltage step $(\Delta V/2)$ applied to a constant capacitor (C). In other words, the number of required binary-weighted capacitors could be reduced at the expense of additional reference- voltages [27] [28].

These capacitance reduction methods could be combined to further reduce the number of required capacitors. In order to fully exploit such potential, all of these three techniques are incorporated together in this design, which results in a complete simplification of the cap-DAC. Fig. 3.17 conceptually illustrates the evolution of adopted cap-DAC from the conventional 9-bit binary cap-DAC. Within each scheme, their respective reference-voltage switching sequence is also shown in the figure. For each capacitor, $C_i = 2^i \cdot C$, where C is the unit capacitor. Beginning from the 9-bit binary cap-DAC (A), it is first simplified as 8-bit binary cap-DAC (B) and then 8-bit split cap-DAC (C) after employing the top-plate sampling and split cap-DAC technique. By means of the scaled reference-voltage technique, it is finally simplified as a 4-bit split cap-DAC (D) as shown in Fig. 3.17. The final version requires only 7

unit capacitors but with an additional reference-voltage ($REF + H$) besides the original two ($REF -$ and $REF +$). Actually, another reference-voltage ($REF - H$) is still utilized in circuit design but not shown here for convenience. The voltage difference between $REF + H$ and $REF + (\Delta V2)$ is much smaller than that of $REF -$ and $REF + (\Delta V1)$. Therefore, the addition of $REF + H$ provides an extra fine reference-voltage switching scheme ($REF +$ to $REF + H$) in addition to the original one ($REF -$ to $REF +$), which now could be called as the coarse reference-voltage switching scheme. The application of both coarse and fine switching schemes to the

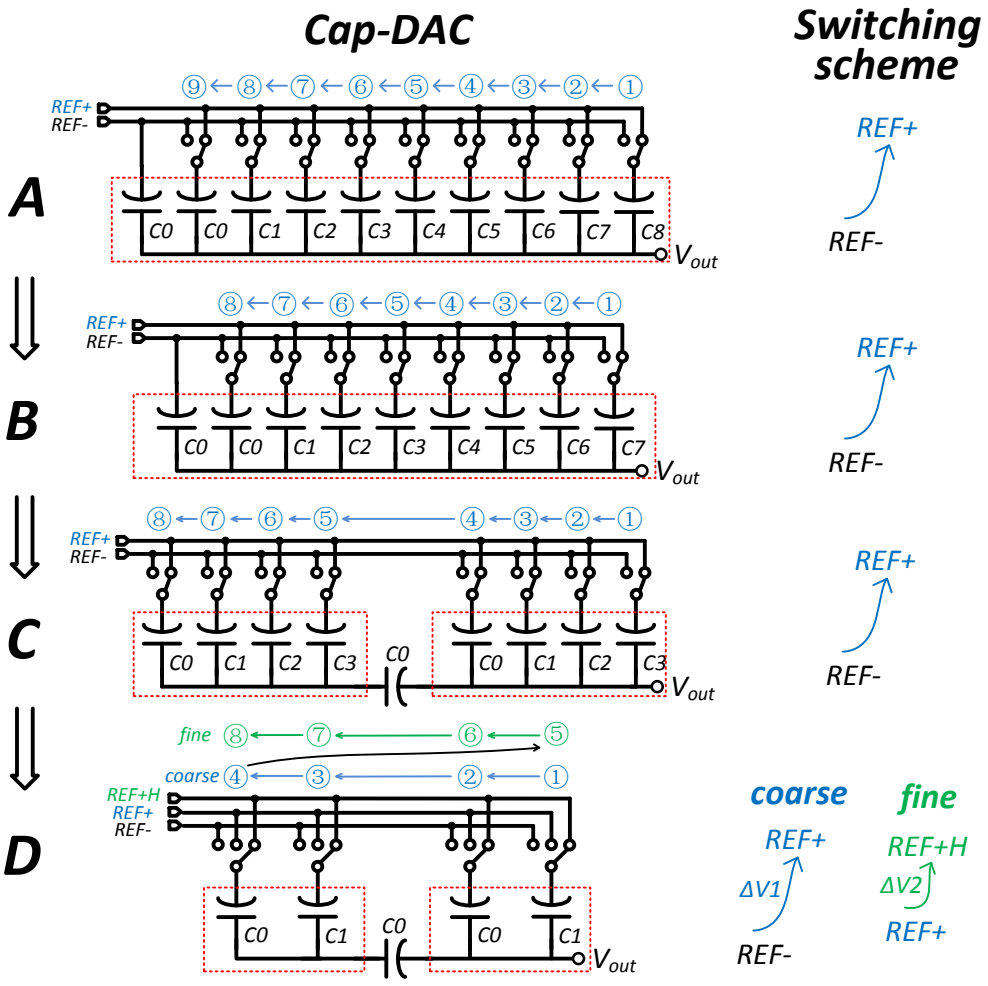


Fig. 3.17 Evolution of the adopted cap-DAC in this design. (A: 9-bit binary cap-DAC, B: 8-bit binary cap-DAC, C: 8-bit split cap-DAC, D: 4-bit split cap-DAC).

same cap-DAC is the key idea which leads to this simple and compact design. To be more specific, the MSB (B_8) is directly converted due to the top-plate sampling strategy. The conversion of remaining bits (B_7 to B_0) could be further divided into two steps (coarse and fine, as shown in D in Fig. 3.17). The first four binary-scaled voltage

steps of V_{out} for conversion of B_7 to B_4 are implemented by applying the coarse reference-voltage switching scheme ($REF-$ to $REF+$) to the 4-bit split cap-DAC. Afterwards, the conversion of the last 4 bits (B_3 to B_0) also makes use of the same split cap-DAC but with the fine reference-voltage switching scheme ($REF+$ to $REF+H$). It is noteworthy that in this design, the requirement $\Delta V1 = 16 \cdot \Delta V2$ should be met to ensure a continuous binary-search algorithm for the SAR ADC.

However, as every coin has two sides, the desired simplification of cap-DAC is not achieved for free. The application of these capacitance reduction techniques, as mentioned above, also introduces other potential design issues which might degrade the performance of ADC and thus deserves careful considerations. Fortunately, as for a medium-resolution (9-bit) ADC oriented for image sensor application, such design issues are usually acceptable due to the following reasons. Firstly, since the number of total capacitors has been cut down significantly, much larger unit capacitance could be chosen to mitigate the adverse effects caused by random mismatch and parasitic capacitance. Secondly, the input voltage of ADC, in this application, is an almost static voltage driven by the front-end switched-capacitor amplifier. Consequently, the sampling switch could be designed as a simple small-size CMOS switch which introduces negligible nonlinear distortion when cuts off (at the end of sampling phase). Finally, the reference-voltages are shared by all column-parallel ADCs and provided globally. The addition of extra reference-voltages doesn't consume too much resource. In the following paragraphs, the ADC architecture and its operation principle are described in detail. For integrity, the influences of these nonideal factors (capacitor mismatch, parasitic capacitance and reference-voltages mismatch) on the linearity of ADC are also investigated in depth through MATLAB system simulations. The according results are also presented as instructions for this ADC design.

Circuit Design and Operation Principle

Fig. 3.18 shows the architecture of adopted SAR ADC which implements two-step (coarse and fine) binary-search algorithm. It consists of four major parts: cap-DAC, reference-voltage MUX array, SAR control logic and comparator. The comparator is a classic sense amplifier [29] with two stage preamplifiers to reduce its offset voltage and kick-back noise. In order to suppress the common-mode noises from power supply

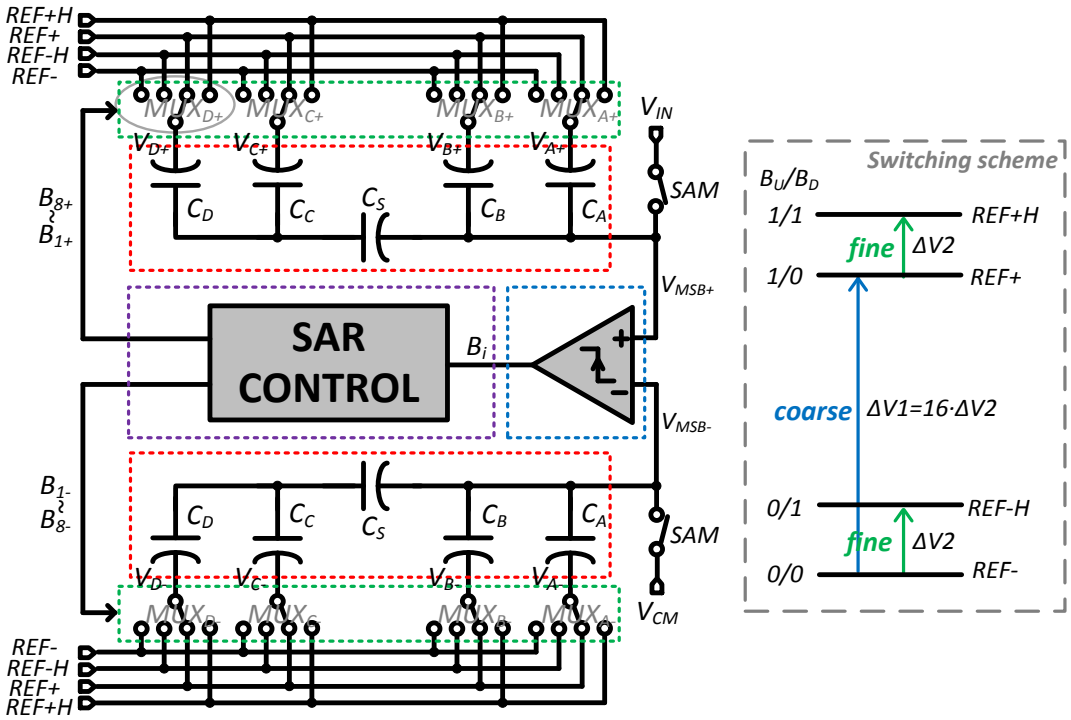


Fig. 3.18 Block diagram of the 9-bit two-step SAR ADC.

and substrate, pseudo-differential architecture is adopted with one cap-DAC sampling the input voltage V_{IN} while the other sampling the common-mode voltage V_{CM}. During A/D conversion, this design uses the monotonic capacitor switching scheme to save power consumption [26]. For this scheme, based on the previous bit converted, only the lower (or higher) input voltage (V_{MSB+} or V_{MSB-}) approximates to the other one (remains constant in this step) with the corresponding binary-scaled voltage step. Instead of making these two voltages approximate to V_{CM}, this monotonic scheme saves most of the switching energies consumed on capacitors and control logic circuit due to the reduced switching activity. The cap-DAC is the aforementioned 4-bit split cap-DAC which is composed of two 2-bit binary cap-DAC ($C_A = C_C = 2C$ and $C_B = C_D = C$) and an unit bridge capacitor ($C_S = C$). In order to implement the coarse and fine two-step reference-voltage switching scheme monotonically, four reference-voltages ($REF-$, $REF-H$, $REF+$, $REF+H$) are provided to the bottom plate of branch capacitors but only three upward switching schemes (one coarse scheme and two fine schemes as shown in Fig. 3.18) are allowable during bit conversion. The coarse switching scheme is from $REF-$ to $REF+$ with a large voltage difference $\Delta V1$. The fine schemes could be either from $REF-$ to $REF-H$ or from $REF+$ to $REF+H$ with a small voltage difference $\Delta V2$. As mentioned before, the requirement

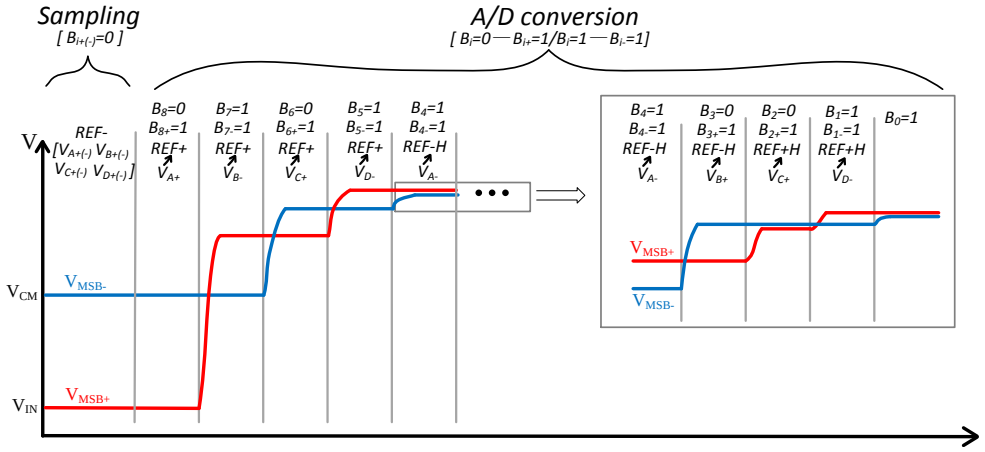
$\Delta V1 = 16 \cdot \Delta V2$ should be met in this design. The SAR control logic latches the comparison results (B_i) and accordingly, generates a series of control bits (B_{1+} to B_{8+}/B_{1-} to B_{8-}) in sequence to ensure the two-step monotonic switching procedure. It shares the same circuit topology with that of [26]. The 4-to-1 reference-voltage MUX determines the bottom-plate's voltage of the branch capacitor according to two control bits (B_U/B_D) allocated from SAR control logic, as shown in Fig. 3.18. The allocation of the control bits to different MUXs are summarized in TABLE 3.2.

TABLE 3.2 Control bits allocation for MUX array

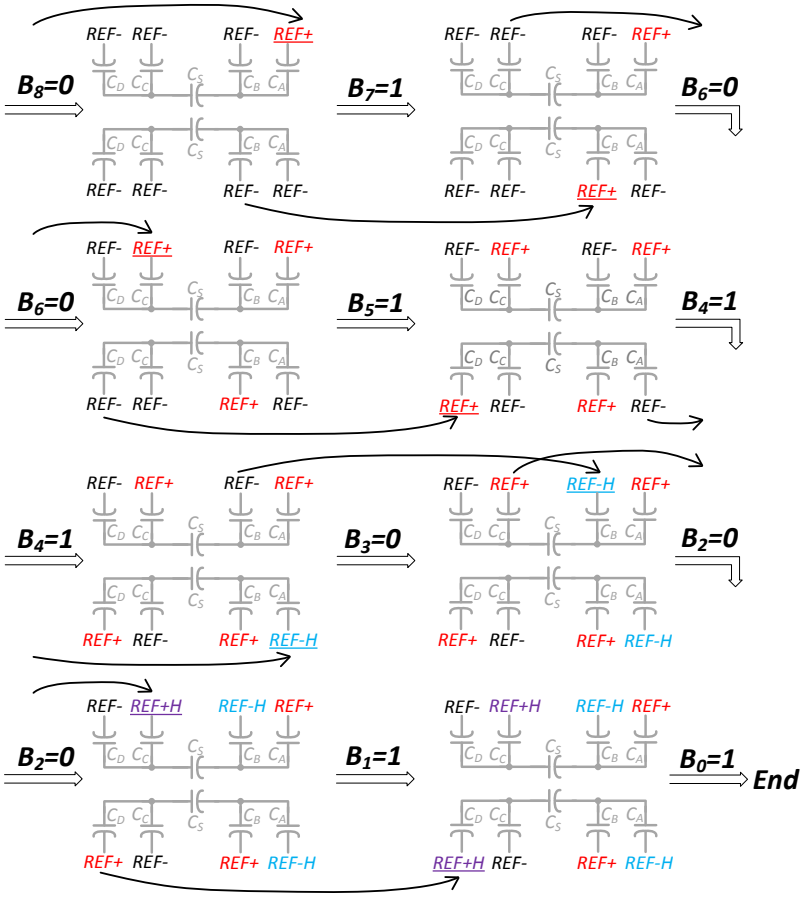
MUX	Control bits	
	B_U	B_D
$MUX_{A+(-)}$	$B_{8+(-)}$	$B_{4+(-)}$
$MUX_{B+(-)}$	$B_{7+(-)}$	$B_{3+(-)}$
$MUX_{C+(-)}$	$B_{6+(-)}$	$B_{2+(-)}$
$MUX_{D+(-)}$	$B_{5+(-)}$	$B_{1+(-)}$

The detailed working principle of the SAR ADC could be illustrated with Fig. 3.19. Fig. 3.19(a) shows an example of the transient waveforms of V_{MSB+} and V_{MSB-}. In addition, the corresponding reference-voltage switching scenarios are also shown in Fig. 3.19(b). The whole operation could be divided into two phases: sampling and A/D conversion. The SAR control logic resets all control bits (B_{1+} to B_{8+}/B_{1-} to B_{8-}) in sampling phase and thus the bottom plates of all branch capacitors are clamped at $REF-$ (both B_U and B_D of all MUXs are 0). In the meanwhile, the sampling switches turn on and the input voltages V_{IN} and V_{CM} are sampled on the top plates of the positive and negative cap-DAC as V_{MSB+} and V_{MSB-}. After that, A/D conversion phase begins and the comparator directly performs the first comparison with respect to these two voltages. The SAR control logic latches the first comparison result as MSB (B_8) and changes the corresponding control bits (B_{8+} and B_{8-}) to implement the first voltage step adjustment. In this case, V_{MSB+} is smaller than V_{MSB-} and B_8 is 0. The control logic accordingly adjusts B_{8+} to 1 while B_{8-} remains 0. At this time, V_{A+} jumps from $REF-$ to $REF+$ since the control bits B_U/B_D of MUX_{A+} change from 0/0 to 1/0 (Fig. 3.18). This results in an ideal $(8/15) \cdot \Delta V_1$ voltage step of V_{MSB+}. On the other hand, V_{MSB-} remains constant in this procedure since V_{A-} is still clamped at

$REF-$ (B_U/B_D of MUX_{A-} is still 0/0). As for the other situation, B_8 would be 1 if V_{MSB+} is larger than V_{MSB-}. B_{8-}, instead of B_{8+}, would change to 1 at this time which leads to the same voltage step of V_{MSB-}. Therefore, as to the monotonic switching scheme for each bit conversion, the larger input voltage remains constant but the



(a)



(b)

Fig. 3.19 Operation principle of the two-step SAR ADC.
(a) Transient waveforms of V_{MSB+} and V_{MSB-}. (b) Reference-voltage switching scenarios.

smaller one would be elevated the corresponding binary-scaled voltage step to follow the larger one. Back to this example, after V_{MSB+} settles down, the second comparison could be executed. The second bit B_7 is converted as 1 since V_{MSB+} is larger than V_{MSB-}. At this moment, B_{7-} would be pulled up to 1 which causes V_{B-} to step up to $REF+$. Since C_B is only half of C_A, the final voltage step of V_{MSB-} is also halved to $(4/15)\cdot\Delta V_1$. Based on such principle, the latter two bits B_6 and B_5 are determined as 0 and 1. The corresponding V_{C+} and V_{D-} switch from $REF-$ to $REF+$, which in turn lead to $(2/15)\cdot\Delta V_1$ and $(1/15)\cdot\Delta V_1$ voltage step of V_{MSB+} and V_{MSB-} respectively. After that, the next bit B_4 is converted as 1. Now, reviewing the bottom plate voltages of the total 8 branch capacitors, four ($V_{A+}/V_{B-}/V_{C+}/V_{D-}$) have been switched to $REF+$ while the other four ($V_{A-}/V_{B+}/V_{C-}/V_{D+}$) still remain at $REF-$. This marks the end of the coarse quantization stage (also the beginning of the fine quantization stage) since the conversion of the last 4 bits ($B_3/B_2/B_1/B_0$) is implemented by applying the fine reference-voltage switching scheme to the same cap-DAC on the basis of previous conversion results. As mentioned above, B_4 is 1 and consequently B_{4-} is 1 (B_{4+} remains 0). As for MUX_{A-}, its control bits B_U/B_D (B_{8-}/B_{4-}) now change from 0/0 to 0/1 which causes V_{A-} to jump from $REF-$ to $REF-H$. This results in an ideal $(8/15)\cdot\Delta V_2$ voltage step of V_{MSB-}. With the relationship $\Delta V1 = 16\cdot\Delta V2$, this voltage step also equals to $(1/30)\cdot\Delta V_1$, which is half of the previous $(1/15)\cdot\Delta V_1$. This guarantees the implementation of continuous binary-search algorithm of SAR ADC. With the same principle, B_3 is 0 and B_{3+} is pulled up to 1 by SAR control logic. V_{B+} switches from $REF-$ to $REF-H$ (since B_U/B_D of MUX_{B+} change from 0/0 to 0/1) and V_{MSB+} increases by $(1/60)\cdot\Delta V_1$. Also B_2 is determined as 0 which changes B_{2+} to 1. This makes V_{C+} jump to $REF+H$ because it has already been switched to $REF+$ in the coarse conversion step ($B_6 = 0$ and $B_{6+} = 1$). The according voltage step of V_{MSB+} is $(1/120)\cdot\Delta V_1$. This procedure is executed repeatedly until the least significant bit (LSB) B_0 is converted, which marks the end of a complete A/D conversion.

Nonideality Analyses

The basic idea, architecture and operation principle of the 9-bit two-step SAR ADC has been expatiated. The objective of this design is to simplify the cap-DAC as far as

possible in order to make the whole ADC suitable for application in column-parallel architecture. The combination of split cap-DAC, top-plate sampling and scaled reference-voltage techniques gives rise to an area-efficient two-step (coarse-fine) SAR ADC with monotonic switching procedure. It only requires 14 capacitors (7 capacitors for each single-end) together with 4 reference-voltages to accomplish the conversion of the required 9 bits. However, these techniques also bring about some adverse effects which might degrade the performance of ADC. The adoption of split cap-DAC makes the whole design more vulnerable to parasitic capacitances as well as mismatches of branch capacitors. In addition, more reference- voltages are introduced and their potential relative mismatches could also influence the ADC's linearity. These nonidealities should be investigated carefully in order to make this ADC design more robust. Therefore, three kinds of nonideality, namely the capacitor mismatch, parasitic capacitances and reference-voltage mismatch, are considered here. Regardless of sampling nonidealities, the linearity of SAR ADC is mainly determined by the accuracy of a series of binary-scaled voltage steps generated by cap-DAC during bit conversion. Since the MSB is directly determined after sampling, there are 8 such voltage steps (also called weights, W_1 to W_8) in total for conversion of the remaining bits. With the help of Fig. 3.20, their respective expression could be readily solved as summarized in TABLE 3.3 by applying the charge conservation principle to the MSB and LSB node respectively. For example, as illustrated in Fig 3.20, the largest weight

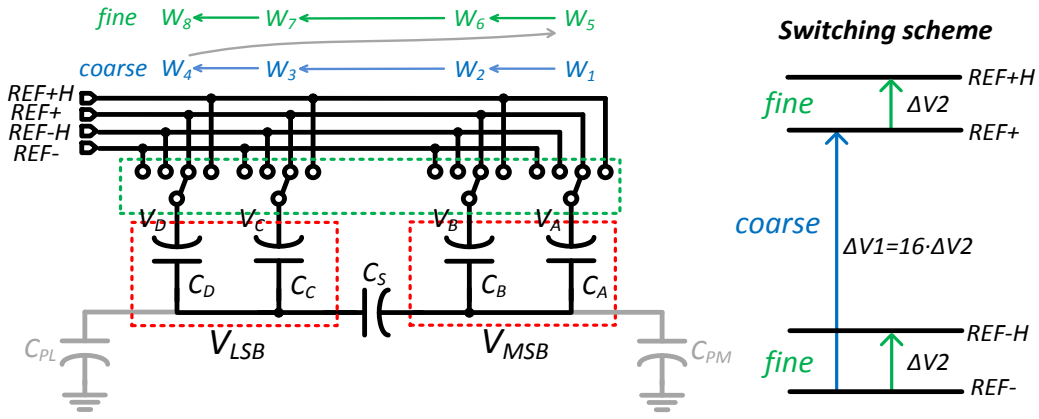


Fig. 3.20 Schematic of the cap-DAC to calculate W_1 to W_8
(C_{PM} and C_{PL} are the parasitic capacitances of the MSB and LSB node).

$W1$ is generated by switching the bottom plate of C_A from $REF-$ to $REF+$ (V_A jumps ΔV_1). The remaining weights could also be analysed according to the two-step

reference-voltage switching scheme shown in this figure. The C_{MSB} and C_{LSB} in TABLE 3.3 are the total capacitances associated with the MSB and LSB node as $C_{MSB} = C_A + C_B + C_S + C_{PM}$ and $C_{LSB} = C_C + C_D + C_S + C_{PL}$, where C_{PM} and C_{PL} are the parasitic capacitances of these two nodes.

TABLE 3.3 Summary of W_1 to W_8 of the SAR ADC

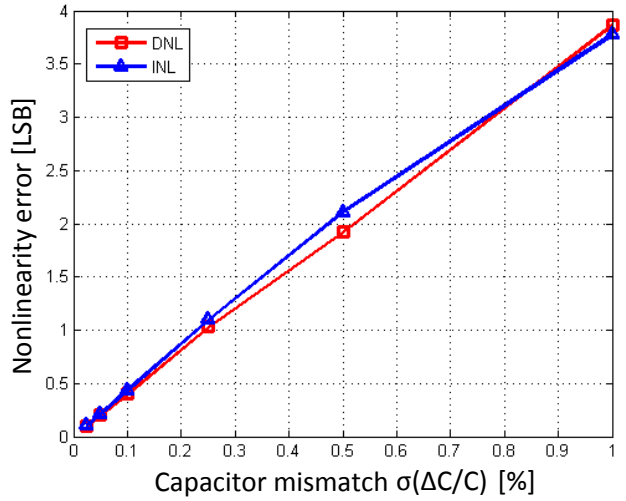
Weight	Expression	Ideal Value
W_1 (V_A jumps ΔV_1)	$\frac{C_A}{C_{MSB} - C_S^2/C_{LSB}} \cdot \Delta V_1$	$\frac{8}{15} \cdot \Delta V_1$
W_2 (V_B jumps ΔV_1)	$\frac{C_B}{C_{MSB} - C_S^2/C_{LSB}} \cdot \Delta V_1$	$\frac{4}{15} \cdot \Delta V_1$
W_3 (V_C jumps ΔV_1)	$\frac{C_C}{C_{MSB} - C_S^2/C_{LSB}} \cdot \frac{C_S}{C_{LSB}} \cdot \Delta V_1$	$\frac{2}{15} \cdot \Delta V_1$
W_4 (V_D jumps ΔV_1)	$\frac{C_D}{C_{MSB} - C_S^2/C_{LSB}} \cdot \frac{C_S}{C_{LSB}} \cdot \Delta V_1$	$\frac{1}{15} \cdot \Delta V_1$
W_5 (V_A jumps ΔV_2)	$\frac{C_A}{C_{MSB} - C_S^2/C_{LSB}} \cdot \Delta V_2$	$\frac{1}{30} \cdot \Delta V_1$
W_6 (V_B jumps ΔV_2)	$\frac{C_B}{C_{MSB} - C_S^2/C_{LSB}} \cdot \Delta V_2$	$\frac{1}{60} \cdot \Delta V_1$
W_7 (V_C jumps ΔV_2)	$\frac{C_C}{C_{MSB} - C_S^2/C_{LSB}} \cdot \frac{C_S}{C_{LSB}} \cdot \Delta V_2$	$\frac{1}{120} \cdot \Delta V_1$
W_8 (V_D jumps ΔV_2)	$\frac{C_D}{C_{MSB} - C_S^2/C_{LSB}} \cdot \frac{C_S}{C_{LSB}} \cdot \Delta V_2$	$\frac{1}{240} \cdot \Delta V_1$

The ideal value of each weight (W_1 to W_8) is also listed in TABLE 3.3 if all the above-mentioned nonidealities are ignored as following: (1) No capacitor mismatch, $C_A = C_C = 2C$ and $C_B = C_D = C_S = C$; (2) No parasitic capacitances, $C_{PM} = C_{PL} = 0$; (3) No reference-voltage mismatch, $\Delta V1 = 16 \cdot \Delta V2$. This leads to an ideal ADC without such nonlinearity errors as differential nonlinearity (DNL) and integral nonlinearity (INL). However, as described below, each kind of the nonideal factor could cause nonlinear distortion to ADC in different manner. (1) Capacitor mismatch: there always exists certain degree of mismatch between capacitors introduced by the fabrication process. Although the system mismatch could be solved by exquisite and symmetric layout, the random mismatch between capacitors is inevitable. As a result,

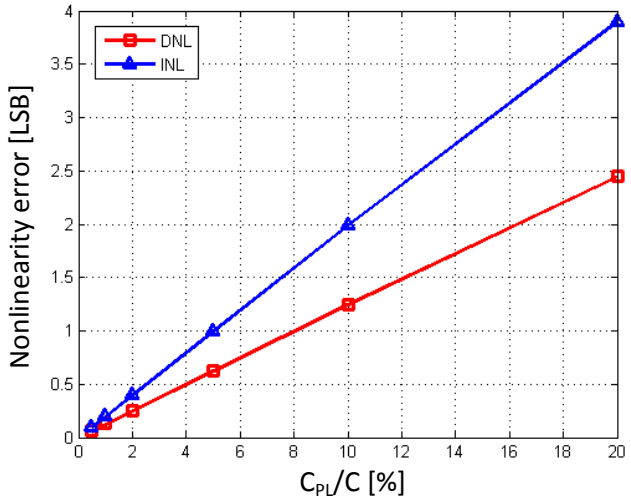
the practical value of the capacitance after being fabricated will be different from the nominal one. Directly dependent on these capacitances, the weights W_1 to W_8 would also deviate from their ideal value, which in turn introduces nonlinearity to ADC. (2) Parasitic capacitances: the parasitic capacitances associated with the MSB and LSB node (C_{PM} and C_{PL}) also influence the performance of ADC. The addition of these capacitances causes C_{MSB} and C_{LSB} larger than the nominal $4C$. Usually, the C_{PM} is acceptable as it only introduces constant gain errors for all weights. In contrast, the C_{PL} would be problematic since it would result in non-binary-scaled weights between W_2 and W_3, W_4 and W_5, W_6 and W_7. (3) Reference-voltage mismatch: as for the implementation of ideal binary-scaled weights, the requirement with respect to the reference-voltages $\Delta V1 = 16 \cdot \Delta V2$ should be satisfied in high-precision. However, just like the capacitors, there must exist certain degree of mismatches between these reference-voltages which would deteriorate such satisfaction. This in turn arouses the non-binary-scaled weights between $W_1/W_2/W_3/W_4$ and $W_5/W_6/W_7/W_8$.

As for the 9-bit SAR ADC, the length of each conversion code (000-1FF) is fully determined by the 8 weights (W_1 to W_8) shown in TABLE 3.3. The aforementioned nonidealities cause the actual weight to deviate from its ideal value and therefore, the conversion codes would have different lengths. Such nonuniformity could be characterized as DNL and INL to evaluate the static performance of ADC. In order to study the influences of these nonidealities on ADC's linearity deeply, a system model is established in MATLAB and used for simulation. Compared to the other method with manual deviations and computations, the MATLAB system simulation is not only much faster but also more convenient to be modified to analyse the influences of different factors. For each nonideality, the practical length of each conversion code is first calculated with the expressions of W_1 to W_8 summarized in TABLE 3.3 (middle column). Based on these data, the nonlinear errors (DNL and INL) of the ADC are figured out. In general, they should be kept below 1 LSB. By means of the following simulation results, some basic design parameters of the SAR ADC, such as the unit capacitance, could be determined in order to meet this requirement. The following paragraphs first illustrate the individual effect of each nonideal factor, which is followed by a performance summary of the SAR ADC.

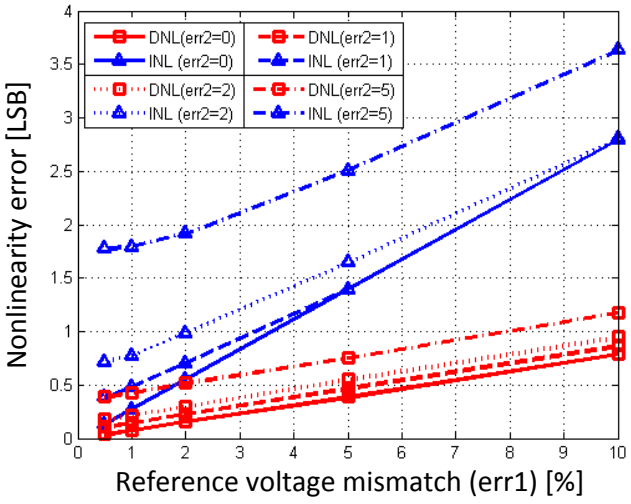
(1) Capacitor mismatch: in general, the standard deviation (σ) of the relative capacitor mismatch is commonly used to describe the capacitor matching characteristic. As for



(a)



(b)



(c)

Fig. 3.21 Simulation results of ADC's nonlinearity errors with respect to (a) Capacitor mismatch. (b) Parasitic capacitance. (c) Reference-voltage mismatch.

the PIP (poly-insulator-poly) capacitor provided by the AMS 0.35 μm CMOS technology, σ is given by

$$\sigma\left(\frac{\Delta C}{C}\right) = \frac{0.45}{\sqrt{W \cdot L}}\% \tag{3.36}$$

where $W \cdot L$ is the area of the capacitor in μm^2. It can be seen that the larger the capacitor is, the smaller the relative mismatch it could guarantee. By directly using this parameter in MATLAB program to generate a series of unmatched capacitors, the length of each conversion code of ADC and thus the corresponding DNL and INL could be solved easily. For each σ, firstly, 1000-times simulations are run and the maximum DNL/INL is picked out. Secondly, such simulation was repeated 10 times and the average value (of the DNL/INL selected in each individual simulation) is calculated as the final DNL/INL. Fig. 3.21(a) shows the ADC's nonlinearity error with respect to the relative capacitor mismatch ($\sigma(\Delta C/C)$). Two significant points should be figured out: σ equals to 0.25 % ($W \cdot L = 3.24\ \mu m^2$) for 1 LSB DNL/INL and 0.12 % ($W \cdot L = 14.1\ \mu m^2$) for 0.5 LSB DNL/INL.

(2) Parasitic capacitances: as analysed before, the C_{PM} doesn't introduce nonlinearity distortion to ADC and thus the C_{PL} is the only parasitic capacitance that matters. Fig. 3.21(b) shows the ADC's nonlinearity error with respect to the normalized C_{PL} by unit capacitance C. It can be seen that the DNL/INL decreases almost linearly with smaller C_{PL}/C. To ensure at worst 1 LSB INL, C_{PL}/C should less than 5 %. More strictly, it should less than 2.3 % for 0.5 LSB INL.

(3) Reference-voltage mismatch: the requirement $\Delta V1 = 16 \cdot \Delta V2$ should be met in high precision to ensure the linearity of ADC. In reality, however, the mismatches between these reference-voltages make this requirement only be partially satisfied with inevitable errors. Since $\Delta V2$ actually comes from two scenarios ($REF-$ to $REF-H$ or $REF+$ to $REF+H$), it could be further split as $\Delta V2_{-}$ and $\Delta V2_{+}$ for each situation. $\Delta V2$, now is used for the average value of $\Delta V2_{-}$ and $\Delta V2_{+}$. There are total two kinds of error ($err1$ and $err2$) recognized in this analysis. $err1$ is used to characterize the relative mismatch between $\Delta V1$ and $16 \cdot \Delta V2$ and defined as

$$err1 = \left|\frac{\Delta V1}{16 \cdot \Delta V2} - 1\right| \times 100 \tag{3.37}$$

while $err2$ is introduced to characterize the relative mismatch between $\Delta V2_{-}$ and $\Delta V2_{+}$ and defined as

$$err2 = \left| \frac{\Delta V2_+ - \Delta V2_-}{\Delta V2} - 1 \right| \times 100 \quad (3.38)$$

Fig. 3.21(c) shows the simulation results of DNL/INL for different combinations of $err1$ and $err2$. The $err1$ is scanned from 0.5 % to 10 % while $err2$ is fixed at four different levels (0, 1 %, 2 %, 5 %). Based on the MATLAB simulation results, some key design parameters of ADC could be chosen accordingly. The effects of random capacitor mismatch and parasitic capacitance determine the minimum unit capacitance. A 97 fF PIP capacitor with area of 6×18 μm^2 was selected as the unit capacitor. The adverse effect of random capacitor mismatch is alleviated to a great extent (0.16 LSB DNL/INL) in this case. On the other hand, the C_{PL}, which is extracted as 1.5 fF from post-layout simulation, only causes 0.3 LSB INL for the total ADC. The four reference-voltages ($REF-$, $REF-H$, $REF+$, $REF+H$) are provided by on-chip DACs whose control words are written from off-chip processor. Based on the simulation results from Fig. 3.21(c), they should be adjusted to guarantee at worst 1 % $err1$ and $err2$ to ensure 0.5 LSB INL for the ADC. Other specifications of the SAR ADC are summarized in TABLE 3.4.

TABLE 3.4 Performance summary of the SAR ADC

Technology	0.35 μm CMOS
Supply voltage	3.3 V
Full scale input range	1.8 V
Sampling rate	1.67 MS/s
Resolution	9 bit
Input referred noise	0.2 LSB_{rms}
DNL	-0.18/+0.22 LSB
INL	-0.34/+0.38 LSB
Power consumption	76 μW
Core area	30×700 μm^2

3.3 SNR and FPN of the Sensor

The design details of logarithmic receptor, SC amplifier and SAR ADC have been described in the above sections. Based on these analyses, the SNR and FPN of this

image sensor could be calculated accordingly. These two specifications are very important to determine the final image quality of the sensor.

3.3.1 SNR

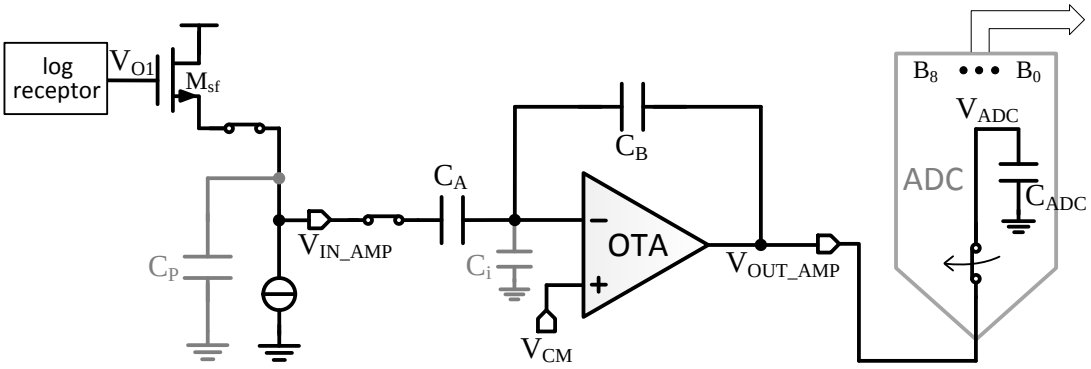


Fig. 3.22 Simplified schematic of the readout path for temporal noise analysis.

In order to calculate the SNR of image sensor, the total readout noise should be first figured out. Fig. 3.22 shows the simplified schematic of the readout path for temporal noise analysis. The in-pixel logarithmic receptor converts photocurrent into V_{O1}. It is buffered to the column readout bus through M_{sf} and then conditioned by the column SC amplifier to V_{OUT_AMP}. Finally, this voltage is sampled as V_{ADC} on the top plate of cap-DAC and then converted to digital codes by the 9-bit SAR ADC. The total parasitic capacitances along the readout bus are lumped as C_P and estimated as 1 pF in this design. The input parasitic capacitance of RFC OTA is also considered here as C_i which equals to 50 fF. Other design parameters are: $C_A = 400\ fF$, $C_B = 100\ fF$, $C_{ADC} = 400\ fF$, $g_{msf} = 90\ \mu S$, $\eta_{sf} = 0.18$ and $g_{mOTA} = 180\ \mu S$. According to these parameters, the transfer function of the readout path from V_{O1} to V_{ADC} could be solved as

$$H(s) = \frac{V_{ADC}(s)}{V_{O1}(s)} = G \cdot \frac{(1 - s/z_0)}{(1 - s/p_1)(1 - s/p_2)} \tag{3.39}$$

where

$$G = -\frac{1}{1 + \eta_{sf}} \cdot \frac{C_A}{C_B} = -3.4$$

and $z_0 = 2\pi \cdot 280M\ rad/s$, $p_1 = -2\pi \cdot 4.6M\ rad/s$ and $p_2 = -2\pi \cdot 17.2M\ rad/s$.

Four noise components which account for the total accumulated noise of V_{ADC} are recognized and analysed. The contributions from the noise induced by in-pixel logarithmic receptor, source follower (M_{sf}), column RFC OTA and ADC sampling switch are analysed to figure out the total readout noise of the signal path.

(1) According to previous analysis, the total output noise of pixel front-end logarithmic receptor is

$$V_{o1\_rms} = \sqrt{V_{o1\_rms\_vn1}^2 + V_{o1\_rms\_vn2}^2} \tag{3.40}$$

where $V_{o1\_rms\_vn1}$ and $V_{o1\_rms\_vn2}$ are two noise components corresponding to the in-pixel inverting amplifier and feedback transistor respectively. They are shown as

$$V_{o1\_rms\_vn1} = \sqrt{2\xi} \cdot (1 + C_{IN}/C_C) \cdot \sqrt{kT/(C_{IN} + C_L + C_{IN}C_L/C_C)} \tag{3.41}$$

$$V_{o1\_rms\_vn2} = \sqrt{\xi/2} \cdot \sqrt{kT/C_C} \tag{3.42}$$

In pixel design, $C_{IN} = 30\ fF$, $C_L = 60\ fF$ and $C_C = 10\ fF$. Since ξ equals to 1.3 for this technology, $V_{o1\_rms\_vn1}$ and $V_{o1\_rms\_vn2}$ are calculated as 810 μV and 520 μV respectively. The total noise V_{o1_rms} is about 960 μV. When converts this noise into the output-end (V_{ADC}), the classic noise calculation method should be adopted. From the perspective of frequency-domain analyses, however, it should be noted that the noise bandwidth of the in-pixel front-end logarithmic receptor is much smaller than that of the subsequent readout path. Since all transistors in pixel work in subthreshold region, the equivalent noise bandwidth for $V_{o1\_rms\_vn1}$ and $V_{o1\_rms\_vn2}$ are only in the order of several decades to hundreds kHz, which are much smaller than p_1 and p_2 (several MHz). This means that all the noise components of logarithmic receptor are amplified and accumulated to the output-end without attenuation. Therefore, the RMS noise of V_{ADC} induced by the in-pixel logarithmic receptor could be calculated as

$$V_{ADC\_rms\_lr} = |G| \cdot V_{o1\_rms} = 3.4 \cdot 0.96\ mV = 3.264\ mV \tag{3.43}$$

(2) The RMS noise of V_{ADC} induced by the in-pixel source follower M_{sf} is

$$V_{ADC\_rms\_sf} = \sqrt{\int_0^\infty \overline{V_{n\_sf}^2} \cdot |H(j2\pi f)|^2 df} \tag{3.44}$$

where $\overline{V_{n\_sf}^2} = 4kT\gamma/g_{msf}$ ($\gamma = 2/3$). It is calculated as 300 μV by means of MATLAB.

(3) The RMS noise of V_{ADC} induced by the RFC OTA during amplifying stage could be expressed as

$$V_{ADC\_rms\_OTA} = \frac{1}{2}\left(1 + \frac{C_A}{C_B}\right) \cdot \sqrt{\overline{V_{in}^2}} \cdot \sqrt{\frac{g_m C_B}{C_A C_B + C_A C_{ADC} + C_B C_{ADC}}} \quad (3.45)$$

Based on the specifications of RFC OTA provided by TABLE 3.1, $V_{ADC\_rms\_OTA}$ is calculated around 675 μV.

(4) The RMS noise of V_{ADC} induced by the ADC sampling switch is

$$V_{ADC\_rms\_sam} = \sqrt{kT/C_{ADC}} \quad (3.46)$$

C_{ADC} is 400 fF in this design and thus $V_{ADC\_rms\_sam}$ is 102 μV.

In summary, the total readout temporal noise V_{ADC_rms} is the accumulation of the four aforementioned factors. Based on equation (3.47), it is computed as 3.35 mV_{rms} and evidently, the in-pixel logarithmic receptor contributes the majority.

$$V_{ADC\_rms} = \sqrt{{V_{ADC\_rms\_lr}}^2 + {V_{ADC\_rms\_sf}}^2 + {V_{ADC\_rms\_OTA}}^2 + {V_{ADC\_rms\_sam}}^2} \quad (3.47)$$

The SNR could also be achieved accordingly as

$$SNR = 20\log\left(\frac{V_{sig}}{V_{ADC\_rms}}\right) \quad (3.48)$$

where V_{sig} is the useful signal range received by ADC. Based on the relative brightness contrast of the scene, the SNR is also different. For a thousandfold range of illumination, V_{sig} is about 795 mV, which corresponds to around 48 dB SNR. This indicates a relatively good image quality (as illustrated in the measurement results) and should be acceptable for applications in most of the back-end image processing algorithms.

3.3.2 FPN

Besides the SNR, another important issue associated with the logarithmic-response image sensor is larger FPN. The conventional integration-mode image sensor (3-T or 4-T APS) requires some exposure time for in-pixel parasitic capacitance to integrate the very small photocurrent into a relatively large recognizable output voltage. Thanks to the integration and reset operation principle, the FPN could be reduced significantly by means of the correlated double sampling (CDS) technique. The readout circuit samples the integration voltage twice: once immediately after reset and the other after integration. What concerned is only the difference between these two voltages. The pixel induced offset voltages are almost cancelled by this method and therefore, the

useful difference voltage is only proportional to the photocurrent. For the logarithmic-response image sensor, the pixel directly converts the photocurrent into a continuous-time voltage by means of a transistor working in subthreshold region. Therefore, the CDS method becomes unusable in this case due to the absence of another reset state. That's the reason why such kind of image sensor usually suffers from much larger FPN and poor image quality. Some on-chip FPN-cancellation methods have been proposed to mitigate this problem. For example, by introducing an additional false reset state in pixel, [17] enables the continuous–time logarithmic-response sensor to adopt the CDS method as well. Another analog self-calibration scheme is also proposed in [18] to reduce the larger FPN. However, both sensors require additional complex circuits to implement such function. They not only introduce lots of extra transistors in pixel but also require more complex timing control. In order not to make the pixel design too complicated, this sensor adopts the common back-end CDS method which implements such function with back-end signal processing. This method requires no modifications to the sensor design in circuit level but could also achieve good FPN cancellation effect at the expense of a little more back-end signal processing. The following part first analyses the causes of FPN in this design and then briefly introduces the back-end CDS method to reduce FPN.

The variation of CMOS fabrication process results in mismatches of certain parameters of the active and passive devices. In general, the transistors' threshold voltage variation is the most significant factor since it is usually much larger than that of other parameters, such as aspect ratio W/L and so on. So in the following analysis, only the offset of transistor's threshold voltage is considered. By means of the intensity readout path shown in Fig. 3.4 and the following equation of V_{OUT_AMP}, four threshold voltage offsets could be figured out that have influence on the final offset of V_{OUT_AMP}, which in turn determines the FPN of the image sensor.

$$V_{OUT\_AMP} = V_{CM} - \frac{C_A}{C_B}\left[V_{REF1} + V_{TH\_fb} + \xi \cdot V_T \cdot \ln\left(I_{ph}/I_0\right) - V_{TH\_sf} - \sqrt{\frac{2I_{CS}}{\mu_n \cdot C_{ox} \cdot S_{sf}}} - V_{IN\_CM}\right] \tag{3.49}$$

The four threshold voltage offsets are: (1) the input-referred offset voltage of inverting amplifier $A1$, denoted as V_{off_A1}; (2) the offset voltage of feedback transistor M_{fb}, denoted as V_{off_fb}; (3) the offset voltage of in-pixel source follower M_{sf}, denoted as V_{off_sf} and (4) the input-referred offset voltage of RFC OTA used in the column switched-capacitor amplifier, denoted as V_{off_OTA}. After taking all these offset voltages into consideration, V_{OUT_AMP} could be expressed as

$$V_{OUT\_AMP} = V_{OUT\_AMP\_ideal} + V_{OUT\_AMP\_off} \tag{3.50}$$

where V_{OUT_AMP} is now divided into two parts: $V_{OUT\_AMP\_ideal}$ is the ideal output voltage free from any internal offsets. It shares the same form with (3.49) and only depends on the photocurrent (I_{ph}) in logarithmic manner. On the other hand, $V_{OUT\_AMP\_off}$ is the total offset voltage of V_{OUT_AMP} shown as

$$V_{OUT\_AMP\_off} = V_{off\_OTA} + \frac{C_A}{C_B}\left(V_{off\_A1} + V_{off\_fb} + V_{off\_sf}\right) \tag{3.51}$$

since amplified by the gain C_A/C_B, the pixel-induced offset voltages (V_{off_A1}, V_{off_fb} and V_{off_sf}) have greater influence on $V_{OUT\_AMP\_off}$. Usually, the offset voltage is characterized by the standard deviation (σ), as a result

$$\sigma\left(V_{OUT\_AMP\_off}\right) = \sqrt{\sigma\left(V_{off\_OTA}\right)^2 + \left(\frac{C_A}{C_B}\right)^2\left[\sigma\left(V_{off\_A1}\right)^2 + \sigma\left(V_{off\_fb}\right)^2 + \sigma\left(V_{off\_sf}\right)^2\right]} \tag{3.52}$$

In order to determine $V_{OUT\_AMP\_off}$, the standard deviations of the four offset voltages are figured out by Monte-Carlo simulations in Cadence. The corresponding results are summarized in TABLE 3.5. According to these data, the calculated standard deviation of $V_{OUT\_AMP\_off}$ is about 52.3 mV. Similar to SNR analysis, the useful signal swing is about 795 mV for a thousandfold range of illumination. This means that the FPN is about 6.6 % under this condition, which is quite worse compared with that of the integration-mode image sensor. Such large FPN is far beyond the acceptable level and would degrade the image quality to a great extent.

With above calculation, the FPN of continuous-time logarithmic-response image sensor is very large, which means very poor image quality. In order to improve the FPN but not make the pixel design too complicated, the back-end FPN-cancellation method is adopted in this design. The fundamental of such method relies on that the offset voltage related to each pixel is time-invariant and thus, it could be cancelled by

TABLE 3.5 Summary of threshold voltage offsets along the signal path

Offset voltage	σ
V_{off_A1}	10.8 mV
V_{off_fb}	5.2 mV
V_{off_sf}	4.9 mV
V_{off_OTA}	5.8 mV

subtraction of two images under different light condition. The advantage of this method is that it only depends on the back-end signal processing to remove FPN and therefore requires no extra hardware overhead. An example of application of such method is shown in Fig. 3.23 where the number shown in each pixel is the corresponding intensity data quantized by ADC. The whole procedure is simple and mainly consists of three steps. Firstly, a whole reference image is taken by the sensor under uniform light condition (Fig. 3.23 (a)). Secondly, the FPN information is extracted by removing the average value of the whole reference image from each pixel (Fig. 3.23 (b)). Thirdly, under real-time application, the newly achieved image first subtracts the previously calculated FPN and then displays (Fig. 3.23 (c) and (d)). By the way, it should be pointed out that the effectiveness of this back-end CDS method

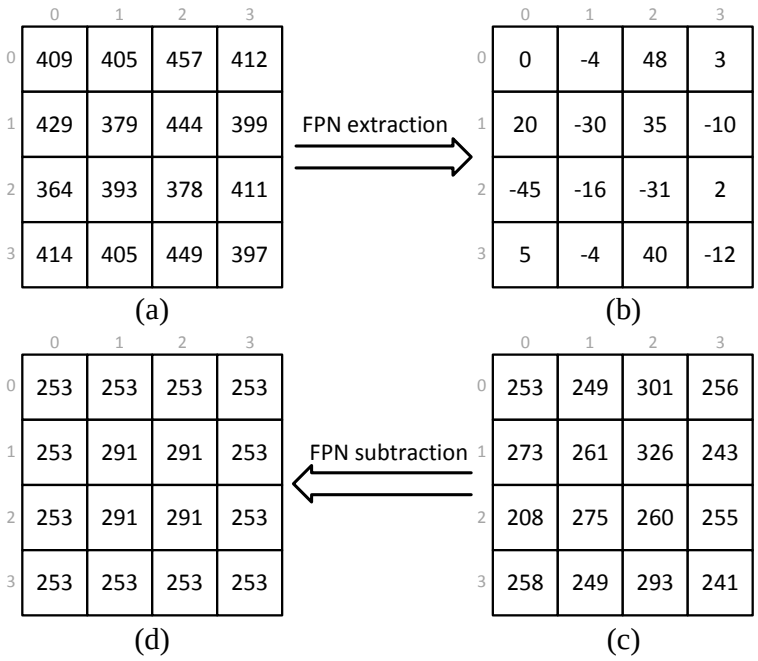


Fig. 3.23 Illustration of the back-end CDS method to cancel FPN.
(a) Reference image. (b) FPN extraction. (c) Real-time raw image. (d) Image after FPN subtraction.

to cancel FPN is influenced by the linearity of readout path. The threshold voltage offsets of transistors could be regarded as constant. As a result, if the read path is fully linear, the FPN could be cancelled completely. In reality, however, the readout path is certain to suffer from some nonlinearities, which would degrade the effectiveness of the FPN cancellation method. In this design, such nonlinearities mainly stem from the in-pixel source follower and the subsequent SAR ADC (nonlinearity error). Fortunately, since these nonlinearities are usually small, most of the FPN could still be cancelled by this method and the image quality is thus improved significantly.

3.4 System Considerations

The sensor was implemented using AMS 0.35 μm 2P4M CMOS process with an array of 160×192 pixels. The chip photograph is not available at present and alternatively, the layout of the whole image sensor is shown in Fig. 3.24. The total chip area is 7.3×7.4 mm^2. Thanks to the simple logarithmic intensity readout strategy, each pixel occupies 30×30 μm^2 with 14.5 % fill factor. The extra source follower and access switch only account for 5.5 % of the total pixel area. The analog and digital parts of pixel are separated with each other to mitigate the coupling from the noisy digital circuits to the sensitive analog parts. Besides, the power metals are also paved above the analog circuits to shield the potential coupling from the above digital handshaking buses. The whole pixel array is surrounded by a dedicated guard-ring circuit with large decoupling capacitances. This not only isolates external noises but also provides low-resistive power rails (VDD and GND) for pixels. The column readout circuit containing the SC amplifier, SAR ADC and AER handshaking logic was designed into a slice of 30×1020 μm^2. A detailed layout of the column slice is also shown at the bottom of Fig. 3.24. The two-step monotonic switching SAR ADC occupies about 30×700 μm^2 with 26 % area allocated to the cap-DAC. In addition, all the bias and reference voltages for the column slice are generated on chip. Specifically, four on-chip high-resolution DACs are adopted to generate the required reference-voltages for column-parallel SAR ADC array. They could be easily programmed by off-chip processor according to the standard Serial Peripheral Interface (SPI) protocol.

With respect to the peripheral AER logic, arbiter tree is implemented using a fair arbiter introduced in [30]. The effect of fixed priority tree arbiter element of AER

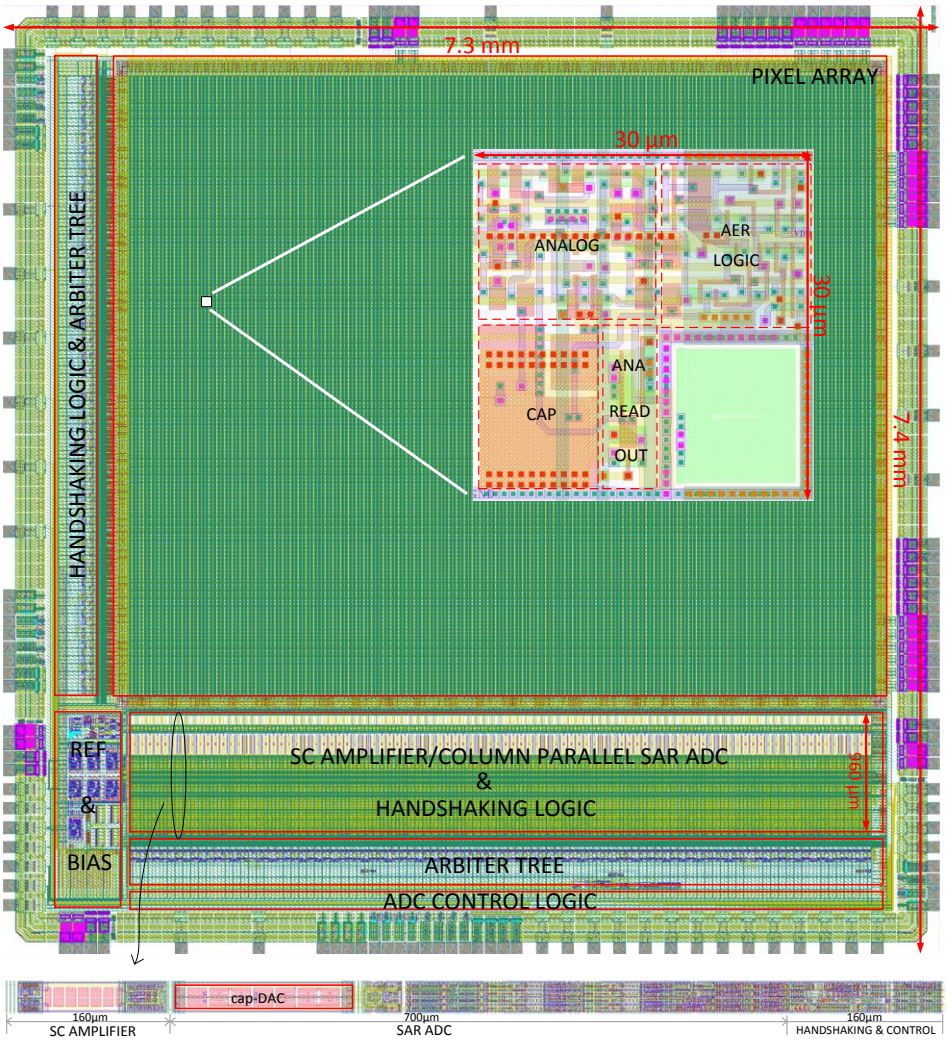


Fig. 3.24 System layout of the sensor
(detailed layout of pixel and column readout circuit are also shown).

sensor often results in unfair allocation of output bus to half of the pixels. An improved tree arbiter in this design has the capability to detect simultaneous requests and dynamically toggles the priority of requests. The featured eager propagation of request allows the request propagation to higher hierarchy during arbitration to minimize the delay time. Usually, the DVS sensor is asynchronous, based on a 4-phase handshaking scheme using request and acknowledge signaling. For the sake of easier FPGA interfacing, the clock signal is also introduced into the system. Once a row is selected, the fired pixels in this row will be processed in a burst mode. After the required conversion time of ADC, each pixel will take one clock cycle to be readout. Therefore, the sensor will push the data out to the FPGA (or other receiver) without being asked. Since the sensor is self-acknowledged, the output is only a *valid* signal combined with some data packages which include 8-bit row/column address and 9-bit

ADC data. The FPGA samples these data once for every clock cycle when $valid$ is high. Moreover, an external global control signal, namely $Forcefire$ is also added in this design which can force the pixel to generate a request unconditionally through the AER handshaking logic circuits. It ensures the capability of on-demand full-picture acquisition of this sensor. When $Forcefire$ is activated, all pixels of the sensor will be fired and after they are completely readout in sequence, a full frame can be produced accordingly. In summary, this sensor could work in three different modes: the analog, digital and progressive mode. For the analog mode, the sensor works as the conventional frame-based image sensor with periodic $Forcefire$ stimulus and it could generate maximum 300 frames per second. This mode makes the sensor compatible with mainstream image processing algorithms. With respect to the digital mode, the sensor works as a pure DVS sensor as it only reports the binary address information of events. The intensity readout path (column SC amplifier and SAR ADC), along with the reference and bias voltage generator, is disabled in this mode to reduce power consumption of the whole system. Finally, the sensor could also be configured into the progressive mode during which the address and intensity information of the fired pixels are reported simultaneously. Since the intensity readout is only trigged by corresponding events ($Forcefire$ disabled in this mode), the data volume needs to be communicated is reduced to a great extent according to the dynamic contents of the visual scene. Such event-driven intensity readout mechanism leads to lossless video compression on focal-plane which enables the sensor to detect high-speed moving object with precise time information.

This page is intentionally left blank.

Chapter 4

Measurement

4.1 Test Setup

In order to characterize the sensor, a FPGA-based testing platform was developed. The Opal Kelly FPGA board is employed and configured to provide input control signals, temporarily store data and communicate with a local host PC through USB link. The power of the sensor is provided by 3.3 V on-board voltage regulators. The sensor works well with maximum 40 MHz clock. Under test, the sensor could be configured into three different operation modes: analog, digital and progressive mode. The available data packets reported from the sensor consist of three elements: event address, illumination data and time stamp (automatically recorded by FPGA when the event is received). When working in analog mode, the sensor outputs continuous frames based on the global stimulus $Forcefire$. The output frame rate is equal to the frequency of stimulus and could reach maximum 300 fps. In digital mode, the sensor only reports address-event information while the column readout circuits are disabled to save power consumption. When configured in progressive mode, the address-event

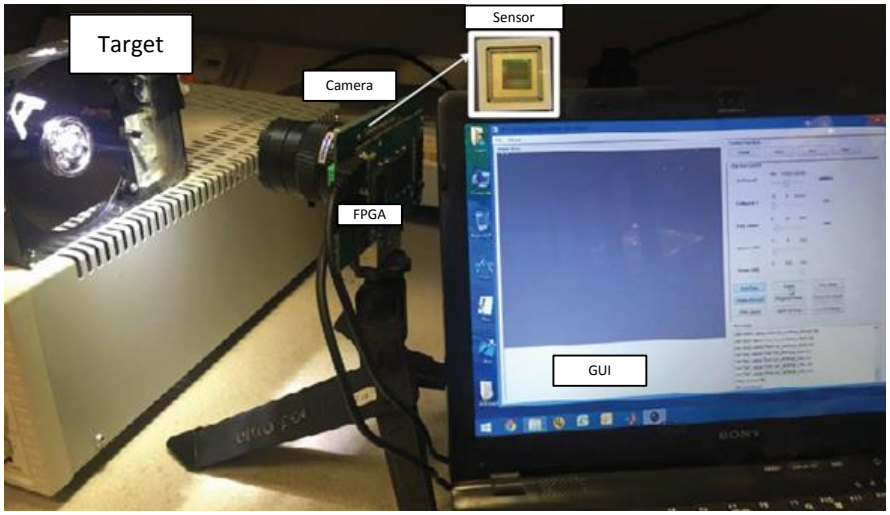


Fig. 4.1 Test setup for the sensor.

and its corresponding illumination data are provided by the sensor to generate lossless video compression across focal plane. A detailed diagram of the test platform is shown in Fig. 4.1. It mainly consists of the following several parts: the camera with mounted DVS sensor and FPGA board, the testing target and the graphical user interface (GUI).

4.2 Characterization Results

The total measurement works of the sensor have been divided into two parts. On the one hand, some important specifications with respect to the logarithmic-response intensity readout have been characterized. The static nonlinearity errors of column-parallel ADC are first measured and it is followed by the sensitivity, intra-scene dynamic range, SNR and FPN measurement of the sensor. On the other hand, some basic specifications of the sensor implementing the DVS function to detect temporal contrast are also tested as the contrast sensitivity, minimum event latency, illumination time resolution and dynamic range.

4.2.1 Characterization of Intensity Readout

ADC Nonlinearity Errors

The sensor is equipped with on-chip column-parallel ADC and thus, the pixel intensity information is reported out in the form of digital codes, instead of analog voltages. Since the measurements of such specifications of the image sensor as SNR and FPN rely on the analysis of these pixel intensity data, the functionality of ADC should be first verified. Consequently, the performance of on-chip column-parallel ADC array is tested with respect to the static nonlinearity errors DNL and INL. Actually, a dedicated ADC test mode is already incorporated in the sensor. During this mode, all other building blocks of the sensor are in reset mode except the column-parallel SAR ADCs and their corresponding reference and bias voltage generation circuits. The testing of the ADC array is also in parallel, which means that all the 192 ADCs sample and convert the same external input signal simultaneously and afterwards, the converted data are shifted out in sequence for further analysis. The external input signal is provided by an on-board high-resolution 14-bit DAC (AD5040) which is configured to

generate a slowly ramp voltage covering the full scale range of the ADC. Based on the standard code density measurement method, the average DNL/INL with respects to all ADCs are figured out as 0.33/0.89 LSB. This is a little worse than the simulation result as summarized in TABLE 3.4 but still acceptable for further measurements as SNR and FPN. It is recognized in our test that the most influential factor on the nonlinearity of ADC is the mismatch of the four reference-voltages ($REF-$, $REF-H$, $REF+$, $REF+H$) provided to the ADC array.

Input-Output Transfer Curve

The input-output transfer curve is measured to figure out the sensitivity and dynamic range of the sensor. This test focuses on the intensity output of a single pixel with respect to different illumination condition. Due to the limitation of test equipment, the maximum achievable range of illumination is 1 to 50k lux, which corresponds to 94 dB dynamic range. The testing results under 1/5/50/500/5k/50k lux light condition are shown as triangles in Fig. 4.2. It can be seen that the measurement results fit well with a line $y = 444.7 - 80.8 \cdot lg(x)$. This verifies the logarithmic-response of this sensor with respect to light intensity. The practical sensitivity of the sensor could also be recognized as 278.8 mV (80.8 LSB, 1 $LSB = 3.45\,mV$) for per decade change of incident light intensity. In addition, the intra-scene dynamic range of the sensor could also be estimated as high as 120 dB accordingly.

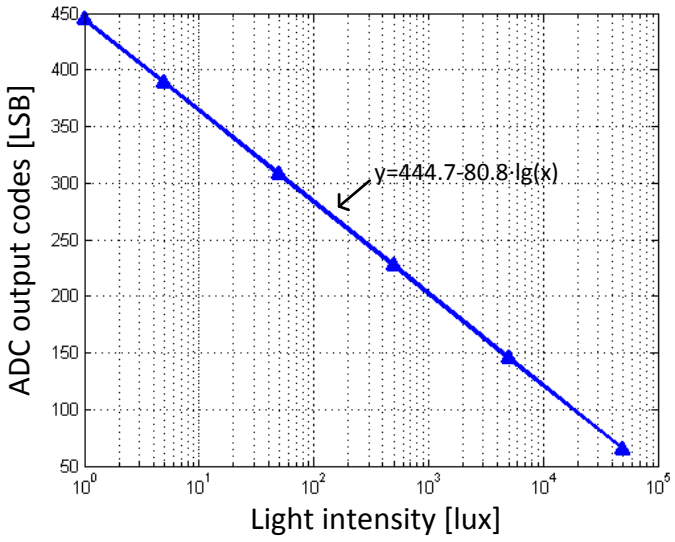


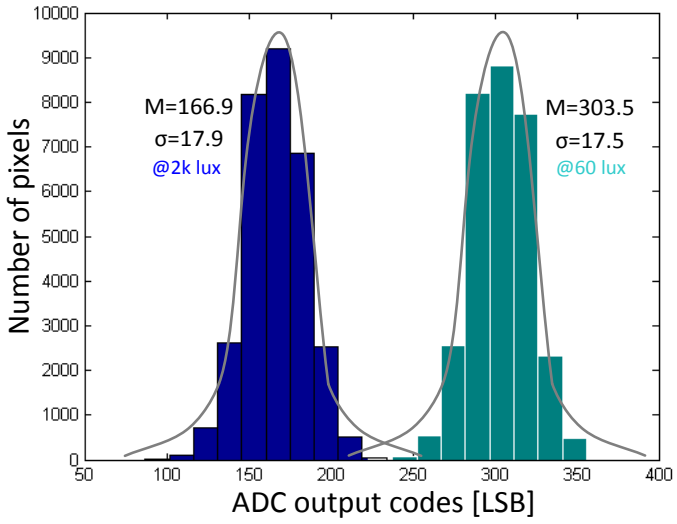
Fig. 4.2 Measurement of input-output transfer curve of the sensor.

SNR and FPN

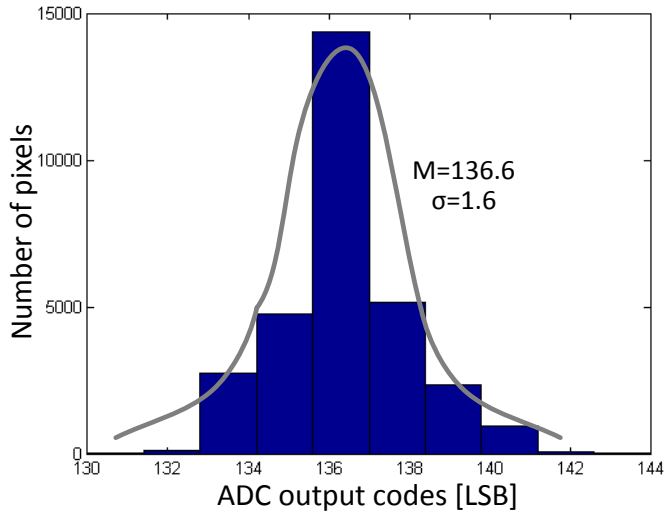
As two most influential factors on final image quality, the SNR and FPN are also measured for this sensor. In terms of SNR, since the sensitivity of sensor has already been characterized in the previous subsection, the remaining work is to figure out the readout temporal noise. As to the sensor in observing a static scene, the output of a single pixel is recorded repeatedly to calculate the readout noise. Ideally, the output should be always constant if no noise components are introduced along the readout path. In reality, the effect of the temporal noise components leads to fluctuations of output codes. In general, with enough samples, they should conform to the Gaussian distribution and the relevant standard deviation is recognized as the RMS readout noise of this sensor. Based on this method, the readout noise is worked out as 1.18 LSB_{rms} with 200 data samples recorded from the same pixel. Converted to voltage, this equals to 4.07 mV_{rms} since 1 LSB of the ADC is about 3.45 mV. Compared with the previous calculation result (3.35 mV_{rms}) in Chapter 3, the extra noises might be introduced by the photodiode shot noise and the ADC-induced readout noise during bit conversion (especially the temporal noise of comparator). Both of these two components are not considered in calculation as shown in equation (3.47). With this data, the SNR of the sensor could also be worked out. The sensitivity of the sensor is 278.8 mV for per decade change of light intensity and therefore, the corresponding SNR could be calculated as 36.7/42.7/46.3 dB for tenfold/ hundredfold/thousandfold brightness contrast of the scene.

The FPN of the sensor is characterized by taking a whole image under uniform light condition and working out the standard deviation of the output data from all pixels. Fig. 4.3(a) shows the histogram of the raw data from a whole image taken under 60 and 2k lux light condition. Each image contains 30720 (160×192) data samples. In fact, a large number of images are taken repeatedly and only the average number is used for FPN analysis. This is necessary to remove the adverse effect of temporal noise. It can be seen that the plotted histogram conforms to the Gaussian distribution. The mean value of the data is 303.5 LSB for 60 lux and 166.9 LSB for 2k lux. The corresponding standard deviation (also FPN) is 17.5 and 17.9 LSB respectively. On the other hand, in order to verify the effectiveness of the back-end FPN-cancellation method, the histogram of the difference of the raw data from these two images is also

plotted as illustrated in Fig. 4.3(b). It can be seen that these data samples enjoy much better uniformity compared with the raw data shown in Fig. 4.3(a). The standard deviation is significantly reduced to 1.6 LSB, only 9 % of the previous 17.7 LSB (average of 17.5 and 17.9 LSB). For a tenfold range of the illumination, this corresponds to about 2.1 % of the useful signal range. As a result, the back-end CDS method is very effective to cancel FPN for this design.



(a)



(b)

Fig. 4.3 Measurement of FPN.
(a) Histogram of raw data under 60 and 2k lux illumination. (b) Difference between the two sets of data.

Sample Image

The sample image of the sensor in analog mode under indoor laboratory environment is shown in Fig. 4.4. The intensity data is reported in 9-bit and for the sake of display, the LSB is abandoned automatically and thus the remaining 8-bit data are mapped into 256 grey levels. The raw image is displayed as Fig. 4.4(a), from which an obvious "rime fog" effect can be recognized. This is because the output intensity data are distributed over a small range around the mid-value, which is resulted from the low brightness contrast under indoor laboratory environment. Such "rime fog" effect could be easily removed with simple back-end signal processing by expanding the original small output data range. The right picture as shown in Fig. 4.4(b) is achieved based on such processing. It can be seen that the contrast is deepened and the "rime fog" effect is disappeared.

(a) (b)

Fig. 4.4 Sample image under indoor laboratory environment. (a) Raw image. (b) Image after processing.

4.2.2 Characterization of DVS Function

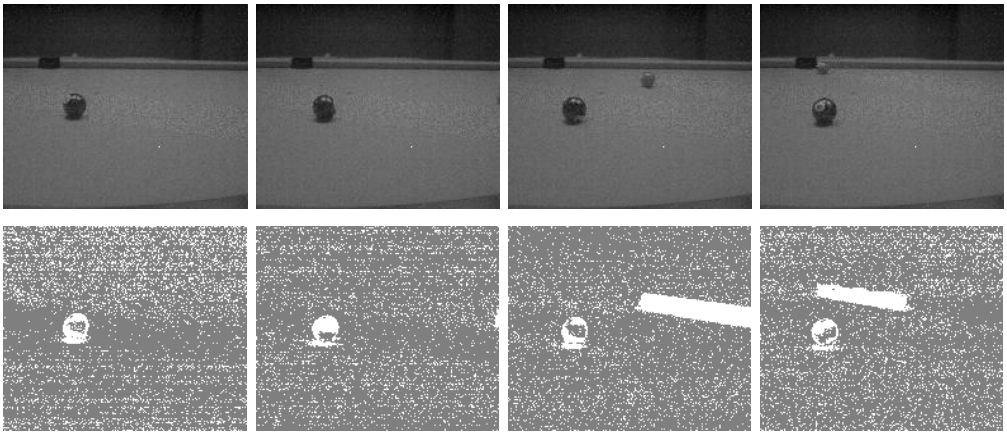
Basic Parameter

Based on the measurement methods illustrated in [3], some basic parameters with respect to the sensor in detecting temporal contrast are also figured out. Firstly, the contrast sensitivity which defines the minimum detectable contrast change is characterized with a moving gradient bar which would trigger the pixels to be fired continuously. The higher contrast sensitivity is desired in some applications to recognize more details of the moving objects at the expense of larger readout data

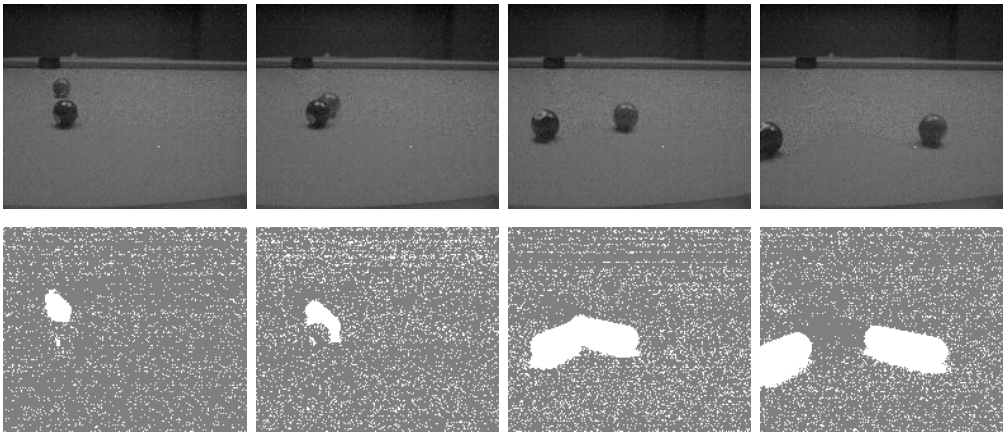
bandwidth and power consumption. As for this sensor, the contrast sensitivity is recognized as 30 % with around 150 mV threshold window. This sensitivity is lower than that of the ATIS (13 %) and DAVIS (11 %) because the gain of the in-pixel difference amplifier is set to only 7 to save the valuable pixel area. In addition, the minimum event latency and the illumination time resolution are also measured. By means of the adopted continuous-time pixel intensity readout strategy combined with the column-parallel ADC array, the sensor is capable of reporting the address-event and its corresponding intensity information simultaneously. Therefore, the minimum event latency and illumination time resolution are same of this sensor and characterized as 6 μs. Finally, the dynamic range of the sensor is also estimated as high as 120 dB since the sensor works well from very bright illumination (>10k lux) down to very dark condition (0.1 lux).

Sample Image

The sample images of the sensor working in digital mode are shown in Fig. 4.5 with the scenarios of fast moving billiards. In this digital mode, only the address-events and the corresponding time stamp are recorded by the FPGA. The sensor does not report the brightness information and thus the final image is a monochrome picture with a white pixel representing an occurrence of corresponding event within the time period of observation. With respect to Fig. 4.5(a), there is a rotating ball in near view while a high-speed moving ball in far view. As for Fig. 4.5(b), a high-speed moving ball is striking the other motionless ball. For clarity, the corresponding grey-scale images recorded by another sensor working in analog mode are also shown for reference. Due to the relatively long observation time, the fast moving billiard in Fig. 4.5(a) could not be figured out as its motion track has been merged into a long strip. The same situation also fits to Fig. 4.5(b). This is not because the sensor could not detect the high-speed moving object but for long reconstruction time for the image. Actually, a more vivid demonstration of this test could be illustrated better with an online video which could be found at https://www.youtube.com/watch?v=O7FZkV42rrM. The fast moving billiards could be easily recognized in this video. In addition, the stray white pixels in image are caused by background event noises, which are intrinsic for the DVS sensor [3, 14-16].



(a)



(b)

Fig. 4.5 Sample image in digital mode in observing moving billiards. Original image (above) and images reconstructed by address-events (below) in observing (a) Two moving billiards and (b) Billiards striking.

Finally, the sensor could generate lossless video compression across focal-plane in progressive mode. In this mode, the address-event and corresponding brightness information are reported simultaneously by the sensor. As to back-end, only the reported brightness information of the fired pixel is updated while others remain unchanged. A set of pictures are extracted in such progressive video and shown in Fig. 4.6. The testing setup is a toy car with the sensor bound on it approaching and separating the front toy car under indoor environment. The measured data volume is around 15 times less than that of the 30 fps frame-based APS.

Besides to the aforementioned characterization work, some other measurement results of this sensor could also be found in www.youtube.com/watch?v=VyXeeX21m5g. This video demonstrates the test of the sensor in detecting high-speed rotating hard disk, some real road applications and other scenarios.

Fig. 4.6 Sample image in progressive mode.

4.3 Performance Summary and Comparisons

In summary of aforementioned measurement results, the complete performance of the sensor is shown in TABLE 4.1. Some fundamental design parameters, together with the characterization results with respect to intensity readout and DVS function of the sensor are all summarized in this table. The most significant feature of this design is to integrate the continuous-time logarithmic intensity readout method to the conventional DVS sensor to accelerate the output of the address-event's corresponding intensity information.

Eliminating the DVS function to detect temporal contrast, this sensor could also be regarded as a logarithmic-response image sensor which reports the pixel intensity information in logarithmic manner. As a result, this sensor design has been firstly compared with other logarithmic-response image sensors as the results summarized in TABLE 4.2. It should be pointed out that for the sake of comparison, the sensitivity and temporal noise of each sensor have been converted with unity readout gain. In addition, the SNR and FPN are calculated with respect to one decade change of incident light intensity. [17] and [18] are relatively old designs featuring FPN-cancellation in pixel level while [19] is a dual-mode linear-logarithmic image sensor also with pixel-FPN cancellation. Since the sensor designed in this project also implements the function to detect temporal contrast, the pixel size is much larger (also corresponds to smaller fill factor). In addition, this design adopts the active logarithmic photoreceptor circuit which converts the photocurrent into a voltage through a feedback loop as illustrated in the previous chapter. Compared with the

TABLE 4.1 Performance summary of the sensor

BASIC PARAMETER	
Technology	0.35μm 2P4M CMOS
Supply voltage	3.3 V
Chip size	7.3 × 7.4 mm^2
Resolution	160 × 192
Pixel size	30 × 30 μm^2
Pixel components	55tr+3cap+1pd
Fill factor	14.5 %
Power consumption	85 mW (core)
INTENSITY READOUT	
Mode	Continuous-time logarithmic
Sensitivity	279 mV/dec
Temporal noise	4.07 mV_{rms}
SNR	46.3 dB (for 1k brightness range)
FPN	0.31 % (of the whole output range)
Intra-scene dynamic range	~120 dB
ADC	Column-parallel two-step SAR
Sample rates	1.67 MS/s
Resolution	9 bit
Area	0.021 mm^2
DNL/INL	0.33/0.89 LSB
Power Consumption	76 μW
DVS FUNCTION	
Contrast sensitivity	30 %
Minimum event latency	6 μs
Illumination time resolution	6 μs
Dynamic range	120 dB

passive photoreceptor circuit employed in [17]-[19], the active counterpart enjoys larger bandwidth which indicates faster response to light intensity change but also brings about more noises. Even for this, the SNR of this sensor is still comparable with [19] and better than [17] and [18]. This demonstrates the effectiveness of the extra compensation capacitance introduced in the front-end logarithmic receptor (as shown

in Fig. 3.1). What's more, the FPN is also in a preferable level due to the effective back-end CDS method. The larger power consumption of this sensor in comparison with [19] is aroused by the adopted in-pixel active amplifier, column-parallel ADC array and also much higher readout speed (300 fps over 50 fps).

TABLE 4.2 Performance comparison with other logarithmic-response image sensors

	[17]	[18]	[19]	This work
Technology ($\mu m\ CMOS$)	0.5	0.6	0.18	0.35
Supply voltage (V)	5	5	3.3	3.3
Resolution	525 × 525	384 × 288	100 × 100	160 × 192
Pixel size (μm^2)	7.5 × 10	24 × 24	6 × 6	30 × 30
Fill factor (%)	N/A	30	32.5	14.5
Sensitivity (mV/dec)	50	100	55	70
Temporal noise (mV_{rms})	2.28	2.3	0.746	1.02
SNR (dB/dec)	26.8	32.8	37.4	36.7
FPN ($\%/dec$)	4.6	3.8	1.96	2.1
Dynamic range (dB)	~120	~120	143	~120
Framerate (fps)	N/A	N/A	50	300
Power (mW)	300	150	1.88	85
ADC	Off-chip	On-chip	Off-chip	On-chip

In essence, this sensor design is an attempt to enable the DVS to achieve the pixel intensity information in a different manner with two pioneering works ATIS and DAVIS and therefore, the comparisons of these works are summarized in TABLE 4.3. Due to the simple logarithmic-response intensity readout method, the pixel design in this sensor is compact. It occupies 30×30 μm^2 with 0.35 μm CMOS technology and is estimated as 16×16 μm^2 if switches to 0.18 μm technology. This is smaller than that of the ATIS and DAVIS. In terms of the DVS function, the contrast sensitivity of the sensor (30 %) is lower because of the relatively smaller gain of the in-pixel difference amplifier. However, the illumination time resolution has been improved as fast as the minimum event latency (6 μs) by means of the continuous-time pixel intensity readout strategy and the adoption of the column-parallel ADC array. With respect to the image quality in intensity mode, the SNR and FPN of this design are also comparable with

that of the two previous works. Due to the logarithmic compression of the light intensity, the intra-scene dynamic range of this sensor is as high as the ATIS (120 dB).

TABLE 4.3 Performance comparison with other DVS designs

	ATIS [4]	DAVIS [5]	This work
BASIC PARAMETER			
Technology ($\mu m\ CMOS$)	0.18	0.18	0.35
Supply voltage (V)	3.3/1.8	3.3/1.8	3.3
Chip size (mm^2)	9.9 × 8.2	5 × 5	7.3 × 7.4
Resolution	304 × 240	240 × 180	160 × 192
Pixel size (μm^2)	30 × 30	18.5 × 18.5	30 × 30
Pixel components	77tr+3cap+2pd	47tr+3cap+1pd	55tr+3cap+1pd
Fill factor (%)	20/10	22	14.5
Power consumption (mW)	175	14	85
DVS FUNCTION			
Contrast sensitivity (%)	13	11	30
Minimum event latency (μs)	4	3	6
Illumination time resolution (μs)	20-30000	25000	6
Dynamic range (dB)	N.A	130	120
Readout bandwidth ($Meps$)	N.A	50	40
INTENSITY READOUT			
Scheme	PWM	APS	Logarithmic
Sensitivity	4.4 $\mu V/e$-	8.8 $\mu V/e$-	279 mV/dec
SNR (dB)	49 (average)	46(maximum)	46.3
FPN (%)	0.25	0.5	0.31
Intra-scene DR (dB)	125	51	120
ADC	Time-based	Off-chip ADC	Col-para SAR

Chapter 5

Conclusions and Recommendations for Future Work

5.1 Conclusions

This thesis demonstrates the research framework for the design of a dynamic vision sensor featuring direct logarithmic output and full-frame picture-on-demand. This sensor is not only capable of reporting the pixel intensity information but also matching it with the high-speed address-event output. Such characteristic is attained by directly taking the voltage of logarithmic photoreceptor as the pixel's intensity information. This voltage logarithmically responses to incident light intensity in continuous-time and thus the elimination of the conventional integration procedure requires no exposure time. By incorporating area-efficient column-parallel ADC array and modifying the 4-phase handshaking communication a little, the sensor could report the event address and the corresponding brightness information simultaneously. Besides, only two transistors are added to the classic DVS pixel which leads to a compact pixel design. The intra-scene dynamic range is also enhanced to up to 120 dB by the logarithmic compression of light intensity. An external control signal is added which can force all the pixels to produce a request and in turn a full-array on-demand reference picture can be obtained. The sensor could work in three different operation modes which make it suitable for different application areas. It could output continuous frames (analog mode), pure binary address-events (digital mode) and the combination of address-events and corresponding intensity information (progressive mode) under these operation modes.

Concentrated on the improvement of image quality in intensity mode, the design considerations of intensity readout path from in-pixel logarithmic receptor to column switched-capacitor amplifier and subsequent SAR ADC have been fully investigated. A compensation capacitance is introduced in the logarithmic receptor to improve its stability and reduce the output noise. The column switched-capacitor amplifier is

incorporated to amplify and accommodate the small-swing output voltage from the pixel to the ADC's full scale range. In order to place the SAR ADC into narrow column slice, an area-efficient architecture with significantly simplified cap-DAC is adopted. In addition, some nonideal factors which might influence the linearity of ADC are also investigated by means of the MATLAB system modeling and simulation.

Based on the aforementioned basic ideas and circuit design, a prototype DVS sensor with 160×192 pixels was implemented with AMS 0.35 μm 2P4M CMOS technology. The pixel pitch is 30 μm and the total chip area is 7.3×7.4 mm^2. The sensor works well with maximum 40 MHz clock with 85 mW power consumption. The performance of the sensor with respect to the intensity readout and DVS function was tested and summarized. The sensor could achieve 46.3 dB SNR and 0.31 % FPN, which make it comparable with the integration-mode APS. Most importantly, the illumination time resolution of the sensor has been reduced to only 6 μs and therefore, the undesired discordance between the very fast address-events output and the postponed intensity readout which inflicted the previous work of ATIS and DAVIS has been disappeared accordingly.

5.2 Recommendations for Future Work

This thesis provides a design with fast intensity readout strategy on the basis of DVS sensor. In view of the progress made at this juncture, the following points are recommended for future work to improve the current design.

(1) Column-parallel SAR ADC: The precision of the four reference-voltages provided to ADC plays an important role in determining its linearity. For the sake of more robustness, on-chip reference-voltage calibration circuitry might be incorporated.

(2) Arbiter: Although the adopted arbiter could dynamically change the priority of two requests, there still exist such condition in real test that the address-event output of the sensor would be locked into only two rows (when all rows continuously send request signals at the same time). As a result, the arbiter should be designed to be fairer to ensure each row has the same opportunity to be served.

(3) DVS: The contrast sensitivity of the DVS needs to be further improved with some modifications of pixel architecture. This could be achieved by either increasing the

gain of in-pixel difference amplifier or cascading more transistors of the front-end logarithmic receptor.

This page is intentionally left blank.

Author's Publications

(1) **Menghan Guo**, Ruoxi Ding and Shoushun Chen, "Live Demonstration: a Dynamic Vision Sensor with Direct Logarithmic Output and Full-Frame Picture-on-Demand," accepted at the *IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, Canada, May 2016.

(2) Heng Guo, Jing Huang, **Menghan Guo** and Shoushun Chen, "Dynamic Resolution Event-Based Temporal Contrast Vision Sensor," accepted at the *IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, Canada, May 2016.

This page is intentionally left blank.

References

[1] Walker SM, Schwyn DA, Mokso R, Wicklein M, Müller T, et al. (2014) "In Vivo Time-Resolved Microtomography Reveals the Mechanics of the Blowfly Flight Motor." *PLoS Biol* 12(3): e1001823. doi: 10.1371/journal.pbio.1001823

[2] [Online] Available: http://www.photron.com/datasheet/FASTCAM_SA3.pdf

[3] Lichtsteiner, P.; Posch, C.; Delbruck, T., "A 128×128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor," in *IEEE Journal of Solid-State Circuits*, vol.43, no.2, pp.566-576, Feb. 2008

[4] Posch, C.; Matolin, D.; Wohlgenannt, R., "A QVGA 143 dB Dynamic Range Frame-Free PWM Image Sensor With Lossless Pixel-Level Video Compression and Time-Domain CDS," in *IEEE Journal of Solid-State Circuits*, vol.46, no.1, pp.259-275, Jan. 2011

[5] Brandli, C.; Berner, R.; MinHao Yang; Shih-Chii Liu; Delbruck, T., "A 240×180 130 dB 3 μs Latency Global Shutter Spatiotemporal Vision Sensor," in *IEEE Journal of Solid-State Circuits*, vol.49, no.10, pp.2333-2341, Oct. 2014

[6] Ohta, Jun. *Smart CMOS image sensors and applications*. CRC Press, 2010

[7] El Gamal, A.; Eltoukhy, H., "CMOS image sensors," in *IEEE Circuits and Devices Magazine*, vol.21, no.3, pp.6-20, May-June 2005

[8] [Online]. Available: https://vimeo.com/blog/post/frame-rate-vs-shutter-speed-setting-the-record-str

[9] Posch, C.; Serrano-Gotarredona, T.; Linares-Barranco, B.; Delbruck, T., "Retinomorphic Event-Based Vision Sensors: Bioinspired Cameras With Spiking Output," in *IEEE Proceedings*, vol.102, no.10, pp.1470-1484, Oct. 2014

[10] M. Mahowald, *An Analog VLSI System for Stereoscopic Vision*. Boston, MA: Kluwer, 1994

[11] K. A. Zaghloul and K. Boahen, "Optic nerve signals in a neuromorphic chip: Part I and II," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 657–675, Apr. 2004

[12] E. Grenet, S. Gyger, P. Heim, F. Heitger, F. Kaess, P. Nussbaum, and P.-F. Ruedi, "High dynamic range vision sensor for automotive applications," in *Proc. SPIE*, 2005, vol. 5663, pp. 246–253

[13] E. Culurciello and A. G. Andreou, “CMOS image sensors for sensor networks,” *Analog Integr. Circuits Signal Process*., vol. 49, pp. 39–51, 2006

[14] Leñero-Bardallo, J.A.; Serrano-Gotarredona, T.; Linares-Barranco, B., "A 3.6 μs Latency Asynchronous Frame-Free Event-Driven Dynamic-Vision-Sensor," in *IEEE Journal of Solid-State Circuits*, vol.46, no.6, pp.1443-1455, June 2011

[15] Serrano-Gotarredona, T.; Linares-Barranco, B., "A 128×128 1.5% Contrast Sensitivity 0.9% FPN 3 μs Latency 4 mW Asynchronous Frame-Free Dynamic Vision Sensor Using Transimpedance Preamplifiers," in *IEEE Journal of Solid-State Circuits*, vol.48, no.3, pp.827-838, March 2013

[16] MinHao Yang; Shih-Chii Liu; Delbruck, T., "A Dynamic Vision Sensor With 1% Temporal Contrast Sensitivity and In-Pixel Asynchronous Delta Modulator for Event Encoding," in *IEEE Journal of Solid-State Circuits*, vol.50, no.9, pp.2149-2160, Sept. 2015

[17] Kavadias, S.; Dierickx, B.; Scheffer, D.; Alaerts, A.; Uwaerts, D.; Bogaerts, J., "A logarithmic response CMOS image sensor with on-chip calibration," in *IEEE Journal of Solid-State Circuits*, vol.35, no.8, pp.1146-1152, Aug. 2000

[18] Loose, M.; Meier, K.; Schemmel, J., "A self-calibrating single-chip CMOS camera with logarithmic response," in *IEEE Journal of Solid-State Circuits*, vol.36, no.4, pp.586-596, Apr 2001

[19] Wei-Fan Chou; Shang-Fu Yeh; Chin-Fong Chiu; Chih-Cheng Hsieh, "A Linear-Logarithmic CMOS Image Sensor With Pixel-FPN Reduction and Tunable Response Curve," in *IEEE Sensors Journal*, vol.14, no.5, pp.1625-1632, May 2014

[20] Assaad, R.S.; Silva-Martinez, J., "The Recycling Folded Cascode: A General Enhancement of the Folded Cascode Amplifier," in *IEEE Journal of Solid-State Circuits*, vol.44, no.9, pp.2535-2542, Sept. 2009

[21] Ruoyu Xu; Bing Liu; Jie Yuan, "Digitally Calibrated 768-kS/s 10-b Minimum-Size SAR ADC Array With Dithering," in *IEEE Journal of Solid-State Circuits*, vol.47, no.9, pp.2129-2140, Sept. 2012

[22] Hao-Chiao Hong; Guo-Ming Lee, "A 65-fJ/Conversion-Step 0.9-V 200-kS/s Rail-to-Rail 8-bit Successive Approximation ADC," in *IEEE Journal of Solid-State Circuits*, vol.42, no.10, pp.2161,2168, Oct. 2007

[23] Sauerbrey, J.; Schmitt-Landsiedel, D.; Thewes, R., "A 0.5-V 1-μW successive approximation ADC," in *IEEE Journal of Solid-State Circuits*, vol.38, no.7, pp.1261,1265, July 2003

[24] Dai Zhang; Bhide, A.; Alvandpour, A., "A 53-nW 9.1-ENOB 1-kS/s SAR ADC in 0.13- μ m CMOS for Medical Implant Devices," in *IEEE Journal of Solid-State Circuits*, vol.47, no.7, pp.1585,1593, July 2012

[25] Saberi, M.; Lotfi, R.; Mafinezhad, Khalil; Serdijn, W.A., "Analysis of Power Consumption and Linearity in Capacitive Digital-to-Analog Converters Used in Successive Approximation ADCs," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol.58, no.8, pp.1736-1748, Aug. 2011

[26] Chun-Cheng Liu; Soon-Jyh Chang; Guan-Ying Huang; Ying-Zu Lin, "A 10-bit 50-MS/s SAR ADC With a Monotonic Capacitor Switching Procedure," in *IEEE Journal of Solid-State Circuits*, vol.45, no.4, pp.731,740, April 2010

[27] Shin, M.-S.; Kwon, O.-K., "14-bit two-step successive approximation ADC with calibration circuit for high-resolution CMOS imagers," in *Electronics Letters*, vol.47, no.14, pp.790-791, July 7 2011

[28] Matsuo, S.; Bales, T.J.; Shoda, M.; Osawa, S.; Kawamura, K.; Andersson, A.; Munirul Haque; Honda, H.; Almond, B.; Yaowu Mo; Gleason, J.; Chow, T.; Takayanagi, I., "8.9-Megapixel Video Image Sensor With 14-b Column-Parallel SA-ADC," in *IEEE Transactions on Electron Devices*, vol.56, no.11, pp.2380-2389, Nov. 2009

[29] B. Wicht, T. Nirschl and D. Schmitt-Landsiedel, "Yield and speed optimization of a latch-type voltage sense amplifier," in *IEEE Journal of Solid-State Circuits*, vol. 39, no. 7, pp. 1148-1158, July 2004

[30] Li, Chenghan, et al. ”Design of an RGBW color VGA rolling and global shutter dynamic and active-pixel vision sensor.” In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 718-721, 2015

This page is intentionally left blank.