Noise in Sub-Micron CMOS Image Sensors

Noise in Sub-Micron CMOS Image Sensors

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op maandag 3 november om 15:00 uur.

door

Xinyang WANG

Master of Science
University of Southampton, Southampton, UK
Bachelor in Electrical Engineering
Zhejiang University, Hangzhou, P.R.China

geboren te Harbin, P.R.China

Dit proefschrift is goedgekeurd door de promotor:

Prof.dr.ir. A.J.P. Theuwissen

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof.dr.ir. A.J.P. Theuwissen,	Technische Universiteit Delft, promotor
Prof.dr. P.J. French,	Technische Universiteit Delft
Prof.dr. E. Charbon,	Technische Universiteit Delft
Prof.dr. P. Magnan,	ISAE, France
Prof.dr. B.J. Hosticka,	Fraunhofer ISM, Germany
Dr.ir. I.M. Peters,	DALSA Professional Imaging, Eindhoven
Dr.ir. P. Centen,	Grass Valley, Breda

Reserve lid:

Prof.dr.ir. G.C.M. Meijer	Technische Universiteit Delft

Printed by PrintPartners Ipskamp, Enschede

ISBN: 9789081331647

Het onderzoek beschreven in dit proefschrift in financieel ondersteund door de Stichting voor Technische Wetenschappen (STW).

to yanxia, and my parents

Table of Contents

Chapter 1

Introduction

The first image created in the mankind's history maybe untraceable, but most likely it appeared even before the formation of actual languages. Through thousands of years, human beings' demand for creating visual images has never stopped and the techniques for capturing such images have continuously been refined, from the prehistoric cave hand-drawing to the latest 52 mega-pixel image captured by a Canon digital camera [1.1]. The invention of such digital cameras is the most recent revolutionary development in imaging-capture devices. The heart of these digital cameras is a so-called image sensor which converts the light intensity to electronic signals. The quality of the captured image is mostly determined by the pixel design and semiconductor technology of the image sensor. The main goal of this thesis project has been to improve the image quality by improving the noise generated in the pixels.

In this chapter, a brief introduction will first be given of the historical background of different types of image sensors in section 1.1. Next, in section 1.2, the scaling of CMOS image sensors (mega-pixel race) in the last decade is introduced. Next, the challenges in designing a large CMOS imager with a very small pixel pitch will be discussed, which is also the motivation for this

thesis. In the end, the structure of this thesis will be presented in section 1.4.

1.1 Background of Image Sensors: CCD vs. CMOS Image Sensor

Two types of semiconductor image sensor technologies are used in modern digital cameras, namely the charge-coupled device (CCD) and the complementary metal-oxide-semiconductor (CMOS) image sensor. Both devices were born during the booming of the semiconductor industry, which started with the invention of the first transistor in November 1947. Since the intention is to replace film-based cameras with electronic devices which can be made in available semiconductor processes, the first attempt to create image sensors was based on existing nMOS or pMOS processes, i.e. a MOS image sensor.

The first successful MOS image sensor was invented by Morrison in 1963 [1.2], followed by Horton from IBM in 1964 [1.3], and Schuster from Westinghouse in 1966 [1.4]. In the early 60s, most of the photosensitive elements used in the image sensors were either phototransistors or n-p-n junctions (scanistors). The use of photon flux integration mode, which is predominant in the CMOS imagers used today, was first proposed by Weckler from Fairchild in 1967 [1.5], when, for the first time, a reverse-biased p-n junction was used for both photosensing and charge integration. This approach built the foundation of the photo-sensing principle in modern CMOS imagers. Based on his method, Noble developed the first 100x100 pixel array in 1968 using an in-pixel source follower transistor for charge amplification [1.6]. In fact this approach is still being used today. Thus, throughout the 1960s, significant improvements were already achieved in terms of the photosensing principle development and the pixel design. However, these early MOS imagers suffered from immature fabrication processes, e.g. a large non-uniformity between pixels due to the process spread,

which introduced extremely high fixed-pattern noise. Therefore, the applications of these MOS imagers were limited.

In 1970, a different type of solid-state imaging device, CCD, was first reported by Boyle and Smith from Bell Labs [1.7]. Compared to MOS imagers, CCDs had the advantage of a simpler structure and a much lower fixed-pattern noise, which made them more suitable for imaging applications. However, although the CCD began to appear in the imaging market in the mid-1970s, its vast commercialization only came 15 years after its birth because of fabrication and reliability issues. The first major success of CCD imagers was in video cameras after which CCDs quickly dominated almost all digital imaging applications.

Although there were several attempts to improve the MOS imagers during the years between the late 1970s and early 1980s [1.8][1.9], the development of MOS imagers was almost completely abandoned because of the success of CCDs. However in the early of 1990s, MOS imagers started to make a comeback[1.10].

Although CCDs had excellent imaging performance, their fabrication processes are dedicated to make photosensing elements instead of transistors. Consequently, it is very difficult to implement well-performed transistors using CCD fabrication processes. Thus, to co-integrate circuitry blocks on a CCD chip is very challenging. However, if the similar imaging performance can be achieved using CMOS imagers, it is even possible to implement all the required functionality blocks together with the sensor, i.e. a camera-on-a-chip, which may significantly improve the sensor performance and lower the cost. In 1995, the first successful high-performance CMOS image sensor was demonstrated by JPL [1.12]. It included on-chip timing, control, correlated double sampling, and fixed pattern noise suppression circuitries.

Since then, the use of CMOS imagers has increased very rapidly and has replaced CCDs in many fields, particularly for applications which require complex functionalities, low power consumption and low cost. However, although CMOS imagers have continued to gain share in the imaging market over the last few decades, CCDs have not become completely obsolete because of their still-superior imaging performance. Figure 1-1 shows the trend of CMOS

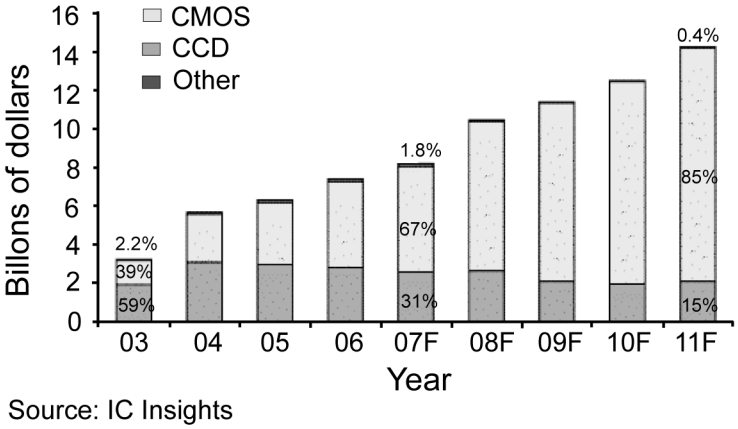


Figure 1-1:CMOS image sensors overtake CCDs, redrawn from [1.11].

imagers overtaking CCDs in the image sensor market [1.11]. As can be seen, even in 2003, the CCD imagers were still the majority in image sensor sales. Although the percentage of CMOS imager sales has increased drastically as indicated and predicted in Figure 1-1, this is mainly due to the growth of novel applications and not the taking over the existing CCD market.

Since 2000, CMOS imagers have stepped into their “golden age” because of the rapidly growing demand from cameras used in mobile telephones. CMOS image sensors are a perfect fit for these kinds of portable electronic device applications because of their small feature size and low-power consumption. Because the CMOS imagers naturally benefit from the fabrication process scaling, their resolution is capable of increasing significantly while maintaining the same sensor size. The continuous demand for higher sensor resolution and the feasibility of scaling down the pixel pitch together sparked the so-called “mega-pixel race” of the last few years.

1.2 CMOS Image Sensor Scaling: Mega-Pixel Race

From 1995, when the first successful 128x128 CMOS imager was made by JPL [1.12], until 2007, when a 52-mega pixel array was announced by Canon [1.1], the resolution of CMOS imagers was increased by more than 3000 times. The ever-shrinking pixel size and the drastically increasing imager resolution have literally brought the development of modern CMOS imagers into a new revolutionary era: a race of making mega-pixel sensors.

The engine behind this race has been the rapid development of semiconductor processes over the last decade, which make it possible to create much smaller pixels. By using a more advanced CMOS process, CMOS imagers naturally benefit from higher resolution, lower power consumption and less cost.

Figure 1-2 shows the roadmap of the state-of-the-art CMOS process, the mainstream CMOS imager process and the pixel pitch over the last two decades. As can be seen, compared to the state-of-the-art CMOS processes which are mainly used to make

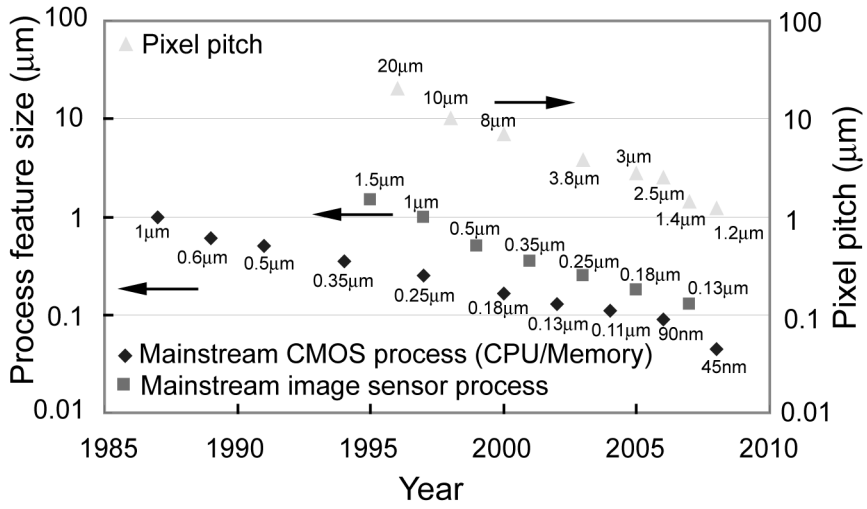


Figure 1-2:Roadmap of mainstream CMOS process, image sensor process and pixel pitch.

CPUs or memories, the imaging fabrication technology is approximately two generations behind. The pixel pitch also shrinks significantly together with the imaging fabrication process scaling, from 20μm in 1996 to 1.2μm in 2008. As can also be seen, between the late 1990s and 2003, the pixel pitch became approximately 20 times the minimal feature size used in the process. However, this ratio between the pixel pitch and process feature size has been reducing and is approaching to ten nowadays. This change shows that the pixel shrinkage speed is faster than that of the process scaling. In other words, people intend to use the current available process as much as possible and shrink the pixel pitch to its absolute minimum before moving to the next technology generation. This raises a very interesting question: why don't CMOS imager designers like to rush to the latest process?

Although the image sensor resolution benefits from the process scaling, new technologies sometimes create significant challenges to the imager performance. For example, the use of shallow trench isolation beyond a 0.18μm technology node introduces significantly increased dark current. More importantly, despite all the benefits of higher pixel resolution, the shrinking of pixel pitch is fundamentally not preferred in terms of the photo-response. Smaller pixel size leads to a reduced photo-sensing area, which ultimately limits the pixel full-well capacity. As will be explained in the next chapter, decreasing pixel full-well capacity damages the image quality by reducing the maximum pixel signal-to-noise ratio and the dynamic range. Consequently, a shared pixel structure is often used when the pixel pitch shrinks below 2μm [1.13].

However, in spite of the standing challenges associated with the shrinking of the pixel pitch, this mega-pixel race still continues. It is difficult to predict when the CMOS imager scaling will end. Although the pixel pitch nowadays can be as small as 1.2μm, it is still able to maintain relatively good imaging performance [1.14]. Moreover, even when the pixel pitch stops shrinking because of certain ultimate constrains, e.g. the optical limit [1.15], its fabrication process can still scale down in order to gain space inside the pixel and integrate more transistors for extra functionalities.

1.3 Challenges and Motivations

As mentioned above, making CMOS image sensors with extremely high resolution or small pixel pitch does involve many technical challenges, both from a micro-fabrication and design point of view. In this section, a few existing challenges will be identified. By addressing these issues, the main motivation of this thesis will be explained as well.

A typical challenge of fabricating such large image sensors stems from the limited exposure area of modern lithography tools. The drastically increased sensor size, which is a result of the multi-mega resolution, may require multiple lithography exposures on one device with stitching options, which therefore introduces variance and non-uniformity [1.16].

Besides the process constrains, the pixel pitch shrinking also introduces some physical limits, which sometimes severely compromise the sensor performance. One important example is the reducing of the pixel full-well capacity, as explained previously. Figure 1-3 shows an example of how pixel capacity and maximal signal-to-noise ratio change as the pixel shrinks [1.17]. As shown, when the pixel pitch shrinks from 5.6μm to 1.7μm, its full well capacity reduces from 30k electrons to 9k electrons, and the maximum signal-to-noise ratio reduces from 44.7dB to 39.5dB. Although there are specific techniques to improve the pixel full-well capacity [1.18], its decrease is in fact a natural consequence of smaller pixel design. Thus, increasing or even just maintaining the same pixel capacity while reducing the pixel pitch is extremely difficult.

The pixel dynamic range is another parameter that is compromised by the decreasing pixel full-well capacity. The dynamic range defines the ratio between the saturation level and the dark noise level. Since the saturation level (i.e. the pixel full-well capacity) reduces, the dynamic range decreases as well. Since maintaining the pixel capacity for smaller pixel is very difficult, the most straightforward approach to maintain the sensor dynamic range is to reduce the noise level. Also, the pixel signal-to-noise

ratio under low illumination conditions is determined by the dark noise level as well. Thus, it would be beneficial when the noise floor of CMOS imagers could be lowered.

The amount of noise from the imager's output signal depends on a number of noise sources. The origins of these noise sources are indeed complicated and often technologically dependent. In other words, adapting a new CMOS imager fabrication process may very well introduce new noise sources. Thus, to reduce the noise level of imagers made in modern processes, it is crucial to first understand what is the dominant noise source and its relationship to some specific process-dependent parameters. Knowing this makes it possible to find an approach to actually reduce the noise level. This is indeed the motivation of this thesis: to address the dominant noise sources in CMOS imagers made in deep sub-micron CMOS processes and to improve the sensor performance by means of reducing sensor dark noise level.

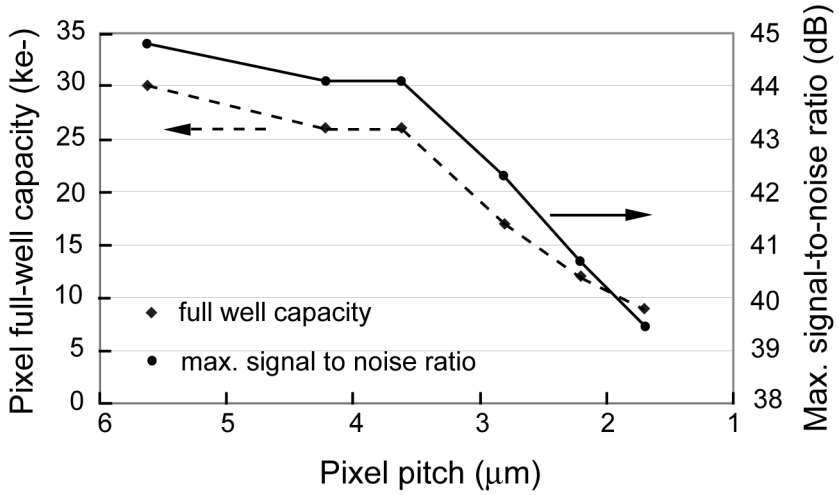


Figure 1-3:Pixel full-well capacity and signal-to-noise ratio as a function of the pixel pitch.

1.4 Thesis Organization

This thesis consists of five chapters. Chapter 2 gives an overview of the architecture and performance of CMOS image sensor pixels. The purpose is to briefly introduce the advantages and disadvantages of CMOS imagers with different pixel structures. The chapter starts with the explanation of some crucial characteristics used to evaluate the performance of a CMOS image sensor. Next, it provides an overview of the physical origin and characterization approach of the fixed-pattern noise (FPN) in CMOS image sensors. Thirdly, the temporal noise in CMOS imager pixels is discussed, and in the end, several commonly used pixel structures are described.

In chapter 3, the dark current of CMOS imagers is analyzed in detail. A description is provided of what the physical mechanisms are of the various dark current sources associated with a CMOS imager pixel. In this chapter, the mechanisms of different types of dark current is first explained. Their generational dependencies are shown using theoretical modeling of the dark current density. Next, different dark current sources of conventional CMOS imager pixels are analyzed in detail. The individual dark current contribution from the photodiode, the transfer gate, the floating diffusion and other elements are shown. Finally, conclusions are drawn on important considerations of designing low dark current pixels. Some basic design trade-offs are presented as well.

In chapter 4, the focus is shifted from the fixed-pattern noise to the pixel temporal noise. Conventionally, the $1/f$ noise is believed to dominate the pixel random noise floor in a pinned-photodiode 4T sensor. However, when the process scales down, a kind of "Lorentzian noise" is actually exhibited instead of the well-known $1/f$ noise, which can be characterized as random telegraph signal (RTS) noise. In chapter 4, the RTS noise of CMOS imagers is analyzed. First, a discussion is presented on the noise measurement results of a pinned-photodiode 4T CMOS imager, which reveal the existence of the RTS noise. This is followed by a theoretical modeling of this noise, which also explains the noise origin. Then,

the RTS noise is further analyzed with varying pixel front-end read-out timings and operation temperatures. It is shown how the properties of interface traps that induce the RTS noise are extracted during experiments. Finally, the relationship between the RTS and the 1/*f* noise in CMOS imagers is briefly discussed.

When the dominant noise source and its origin are known, the next task is to find an approach to reduce the noise level. In chapter 5, a buried-channel source follower is introduced to replace the standard surface-mode nMOS transistor as the in-pixel amplifier. It will be shown that the sensor dark random noise is significantly reduced, for both the 1/*f* and RTS noise components. Moreover, the pixel output swing is increased by almost 100% because of the negative threshold voltage of the buried-channel source follower transistor. The basic operation principles of the new source follower transistor and the fabrication considerations are first discussed in chapter 5. Next, the improved noise behavior measured from image sensors made in 0.18μm CMOS process is presented.

Finally, chapter 6 presents the main conclusions of this thesis and gives suggestion for future works.

1.5 References

[1.1] M. Iwane et al., “52 Mega-Pixel APS-H-Size CMOS Image Sensor for Super High Resolution Image Capturing”, *International Image Sensor Workshop*, pp. 295-298, Ogunquit, US, June 2007.

[1.2] S. Morrison, “A New Type of Photosensitive Junction Device”, *Solid-State Electron*, Vol. 5, pp. 485-494, 1963.

[1.3] J. Horton et al., “The Scanistor-A Solid-State Image Scanner”, *Proceeding of IEEE*, Vol. 52, pp. 1513-1528, 1964.

[1.4] M.A. Schuster et al., “A Monolithic Mosaic of Photon Sensors for Solid State Imaging Applications”, *IEEE Transactions on Electron Devices*, Vol. ED-13, pp. 907-912, 1966.

[1.5] G.P. Weckler, “Operation of P-N Junction Photodetectors in a Photon Flux Integration Mode”, *IEEE Journal of Solid-State Circuits*, Vol. 2, pp. 65-73, 1967.

[1.6] P. Noble, “Self-Scanned Silicon Image Detector Arrays”, *IEEE Transactions on Electron Devices,* Vol. 14, pp. 202-209, 1968.

[1.7] W.S. Boyle et al., “Charge-Coupled Semiconductor Devices”, *Bell System Technical Journal,* Vol. 49, pp. 587-593, 1970.

[1.8] S. Ohba et al., “MOS Area Sensor: Part II-Low Noise MOS Area Sensor with Antiblooming Photodiodes”, *IEEE Transactions on Electron Devices*, Vol. ED-27, pp. 1682-1687, 1980.

[1.9] K. Senda et al., “Analysis of Charge-Priming Transfer Efficiency in CPD Image Sensors”, *IEEE Transactions on Electron Devices*, Vol. ED-13, pp. 1324-1328, 1984.

[1.10] F. Andoh et al., “A 250,000-Pixel Image Sensor with FET Amplification at Each Pixel for High-Speed Television Cameras”, *Technical Digest ISSCC*, pp. 212-213, San Francisco, US, Feb. 1990.

[1.11] “Image Sensors will Reach Record Sales After Weak Start in 2007”, http://www.icinsights.com/news/bulletins/bulletins2007/bulletin20070529.html, *IC Insights Research Bulletin*, 2007.

[1.12] R.H. Nixon et al., “128x128 CMOS Photodiode-Type Active Pixel Sensor with On-Chip Timing, Control and Signal Chain Electronics”, *Proceeding of SPIE*, Vol. 2415, pp. 117-123, 1995.

[1.13]X. He et al., “CMOS Image Sensor Using Shared Transistors Between Pixels with Dual Pinned Photodiode”, US Patent 7087883.

[1.14]“Aptina Imaging Enhances Technology and Product Portfolio”, http://www.aptina.com/news/press/aptina_imaging_enhances_technology_and_product_portfolio/, Aptina Imaging, 2008.

[1.15]H. Wong, “Technology and Device Scaling Considerations for CMOS Imagers”, *IEEE Transactions on Electron Devices*, Vol. ED-43, pp. 2131-2142, 1996.

[1.16]S.U. Ay et al., “A 76 x 77 mm^2, 16.85 Million Pixel CMOS APS Image Sensor”, *Digest of Technical Papers of Symposium on VLSI Circuits*, pp. 19-20, Honolulu, US, 2006.

[1.17]G. Agranov et al., “Optical-Electrical Characteristics of Small, Sub-4μm and Sub-3μm Pixels for Modern CMOS Image Sensors”, *IEEE Workshop on CCDs and Advanced Image Sensors*, pp. 206-209, Karuizawa, Japan, June 2005.

[1.18]Y. Lim et al., “Stratified Photodiode a New Concept for Small Size High Performance CMOS Image Sensor Pixels”, *International Image Sensor Workshop*, pp. 311-314, Ogunquit, US, June 2007.

Chapter 2

Overview of CMOS Image Sensor Pixels

This chapter gives an overview of the architecture and performance of CMOS image sensor pixels. The purpose is to briefly introduce the advantages and disadvantages of CMOS imagers with different pixel structures. Although the intention of this thesis is to analyze the noise in CMOS imagers, often other performance parameters are involved as trade-offs for noise considerations. Thus, it is essential to first clarify what the mechanisms and limiting factors of these performance characters are.

Section 2.1 takes a look at some crucial parameters that are used evaluate the performance of a CMOS image sensor. Next, section 2.2 provides an overview of the physical origin and characterization approach of fixed-pattern noise (FPN) in CMOS image sensors. In section 2.3, the temporal noise in CMOS imager pixels is discussed. Finally, in section 2.4, several commonly used pixel structures are described. The advantages and disadvantages of each type of pixel are also explained. The relative importance of various noise sources among different pixel structures is explained as well.

2.1 Performance Evaluation of CMOS Image Sensor Pixels

There are quite a lot of parameters used to evaluate the performance of a CMOS image sensor. Although some of them are mainly limited by the readout circuitries, the vast majority of them are either determined by or already limited by the pixel design, i.e. the quantum efficiency, dynamic range, saturation level, signal-to-noise ratio, dark current, image lag, non-uniformity and non-linearity of the photon response. This section gives detailed explanations of these important performance characteristics.

Since these parameters serve as objective criteria to evaluate an imager's performance, this section will not focus on any details regarding the exact pixel structure.

2.1.1 Quantum Efficiency and Spectral Responsivity

Quantum efficiency (QE) is a quantitative parameter that reflects the photon-sensitivity of an image sensor as a function of the wavelength (i.e. the energy) of impinging photons. It is defined as the percentage of the photons hitting the photodetector surface that produce an electron-hole pair. It is given by:

$$QE(\lambda) = N_{sig}(\lambda) / N_{ph}(\lambda) \tag{2-1}$$

where N_{sig} is the collected video signal charge and N_{ph} is the number of injected photons; λ stands for the wavelength.

Often, spectral responsivity is also used to characterize the photon-sensitivity of an image sensor. It is defined as the ratio of the photocurrent to the optical input power and is given by:

$$R(\lambda) = \frac{I_{ph}}{P} = \frac{qN_{sig}(\lambda)}{E_{ph}N_{ph}(\lambda)} = QE(\lambda)\frac{q\lambda}{hc} \tag{2-2}$$

where I_{ph} is the photocurrent, P is the optical input power, q is an electron charge, E_{ph} is the photon energy, h is Planck's constant, and c is the speed of light.

As indicated by Eq. (2-1) and Eq. (2-2), the photo-sensitivity of an image sensor can be expressed in two ways. Figure 2-1 shows an example which illustrates the relation between the QE and the spectral responsivity [2.1]. As can be seen, assuming a constant QE at 0.5 in the range of 400 to 700nm wavelength, the spectral responsivity is not uniform because of the extra factor λ, as shown in Eq. (2-2).

Naturaly, the QE should be as high as possible in an imaging system. The ideal QE is one, which means that there is an electron-hole pair being generated and collected for each individual impinging photon. However, such an ideal case is obviously very difficult to achieve in reality. The total QE loss is mainly due to two limitations. The first is the impinging loss which represents the photon loss during the impinging procedures. It includes the loss from the optical system, and the absorption and reflection by the structures above the photodiode (e.g. the metal and dielectric layers). In other words, the impinging loss stands for the missing photons that do not make it to the surface of the photo-sensing region. In order to minimize this loss, an anti-reflection coating (ARC) layer can be added on top of the sensor. In addition, the ratio of the photodiode to the total pixel area, i.e. the fill factor, should be as high as possible.

Secondly, the collection of the photon-generated carriers is not one hundred percent efficient, which thus introduces a QE

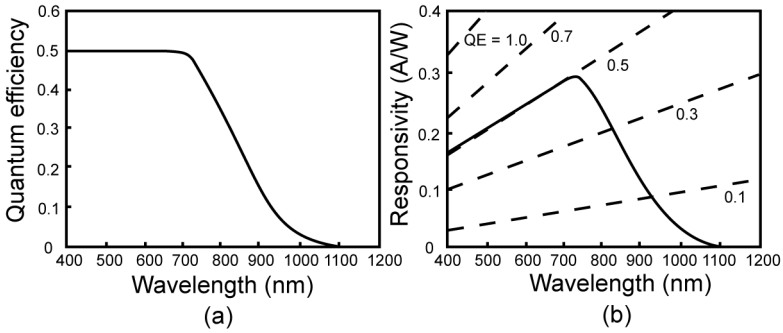


Figure 2-1:Photo-sensitivity: a) quantum efficiency, b) spectral responsivity, redrawn from [2.1].

reduction. To have a better understanding of this collection loss, it is necessary to first go through the photon carrier generation process.

In principle, as long as the energy of the impinging photon is higher than the bandgap of silicon (1.124eV), an electron-hole pair will be generated. Obviously, the absorption efficiency of the impinged photons is determined by the photon energy. Figure 2-2 (a) shows how electron-hole pairs are generated from photons with different energies. As can been seen, the lower the photon energy is (i.e. the longer the wavelength), the deeper the photon can penetrate into silicon before being absorbed.

A p-n junction is used to collect the photon-generated carriers, as shown in Figure 2-2 (b). Ideally, if all carriers can be collected regardless of their depth, there will be no collection loss. However, in most cases only the carriers generated within the depletion region of the p-n junction will be collected without any loss because of the existence of the build-in electrical field (V_{bi}). The carriers generated outside the depletion region may be recombined before diffusing to the depletion region. This collection loss, because of recombination, often introduces a significant QE reduction, particularly for photons with a longer wavelength.

In conclusion, QE and spectral responsivity represent how an imager responds to the impinged photons. To minimize the QE reduction due to impinging loss, an ARC layer can be used while

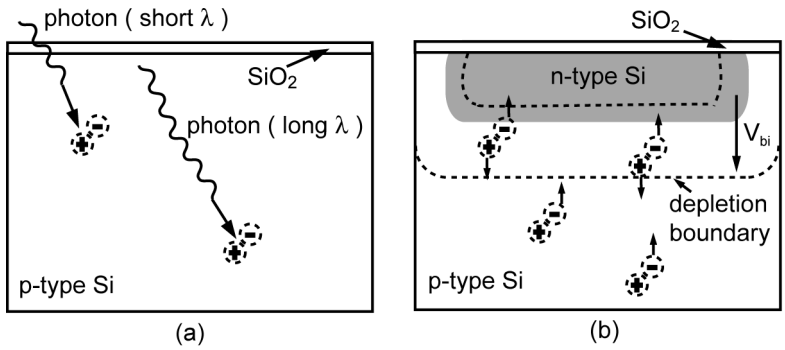


Figure 2-2:a) Electron-hole generations by photons with different wavelength, b) Photon-generated carriers collection by a p-n junction/photodiode.

the fill factor of the pixel design should be as high as possible. In order to avoid significant collection loss, it is essential to maintain a wide and deep depletion region of the photodiode.

2.1.2 Dynamic Range and Full-Well Capacity

A dynamic range (DR) is defined as the ratio between the pixel saturation level and its noise floor. It can be given as:

$$DR = 20\log\left(\frac{N_{sat}}{n_{dark}}\right)[dB] \tag{2-3}$$

where N_{sat} is the signal charge at saturation (which is also called pixel full-well capacity), and n_{dark} stands for the pixel noise level without illumination [in electrons]. As can be seen from Eq. (2-3), there are two ways to increase DR: by either improving pixel full-well capacity or reducing the dark noise level. A detailed analysis on noise in CMOS imagers is given later in this chapter. In this sub-section, only the approaches used to increase pixel full-well capacity are discussed.

As mentioned in the previous sub-section, a p-n junction is oftenly used as the photo-sensing component to collect the photo-generated carriers. Obviously, this photodiode has a maximum capacity of restoring the charge. This maximum charge saturation level is its full-well capacity. Figure 2-3 shows a simplified circuit of a photodiode operating in the charge integrating mode. V_{res} is the reset voltage of the photodiode, i_{ph} is the

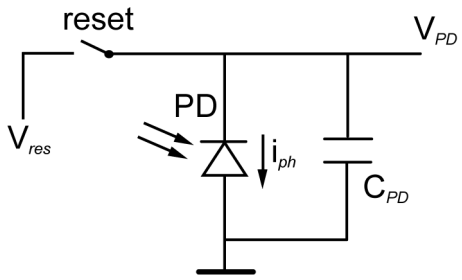


Figure 2-3:Simplied circuit of a photodiode operating in charge integrating mode.

photocurrent, C_{PD} is the photodiode capacitance, and V_{PD} is the photodiode voltage. For the photodiode shown in Figure 2-3, its full-well capacity is given as:

$$N_{sat} = \frac{1}{q}\int_{V_{res}}^{V_{PD(\min)}} C_{PD}(V_{PD}) \cdot dV \qquad (2\text{-}4)$$

where q is electron charge and $V_{PD(\min)}$ is the minimum value of V_{PD}. As can be seen from Eq. (2-4), for a given photodiode, the easiest way to increase N_{sat} is to increase the voltage swing between V_{res} and $V_{PD(\min)}$, i.e. V_{res} - $V_{PD(\min)}$.

Both V_{res} and $V_{PD(\min)}$ depend on the operation conditions, but they have their limits. Increasing V_{res} improves the voltage swing, but consequently it also results in an increase in dark current and the possibility of the photodiode breaking down. $V_{PD(\min)}$ is normally set by the pixel structure. It is important to notice that because C_{PD} is also a function of V_{PD}, the linearity of the photodiode response diminishes.

Besides the photodiode, other structures are also used as photon-sensing elements, e.g. photogates [2.2][2.3] or pinned photodiodes [2.4][2.5]. In the case of photogates, the photon-generated carriers are integrated in a MOS-capacitor, thus the full-well capacity is mainly determined by the doping profile of the silicon underneath the photogate. The charge saturation level of pinned photodiode can be acquired in the same way as that of the photodiode, which can also be calculated from Eq. (2-4). However, the reset voltage V_{res} in a pinned photodiode is normally set by the junction itself instead of the externally applied voltage.

In conclusion, increasing pixel full-well capacity is one way to improve the DR of imagers. However, for a given pixel with a fixed fill factor, increasing full-well capacity is rather difficult because of the restriction of the voltage swing. Because of this, high dynamic range CMOS imagers are normally realized through some specific pixel structures and operation principles, e.g. multi-exposure [2.6] or logarithm pixel response [2.7].

2.1.3 Signal-to-Noise Ratio

As analog circuitry, one of the most important parameters of a CMOS image sensor pixel is its signal-to-noise ratio (SNR). This is defined as the ratio between the signal and the noise at a given input level and can be given as:

$$SNR = 20\log\left(\frac{N_{sig}}{n_{sig}}\right)[dB] \tag{2-5}$$

where N_{sig} is the signal charge [in electrons], while n_{sig} is the total noise at the given signal level [in electrons].

Figure 2-4 shows the SNR as a function of the input photons in an ideal case, where n_{dark} is assumed to be equivalent to 20 photons. At the beginning under low illumination conditions, the dark noise level is dominant and the SNR is roughly given as:

$$SNR = 20\log\left(\frac{N_{sig}}{n_{dark}}\right)[dB] \tag{2-6}$$

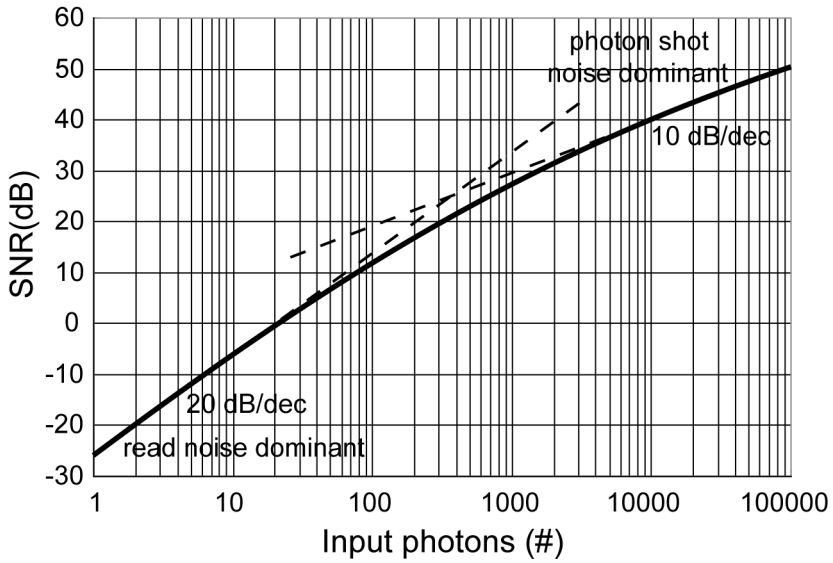


Figure 2-4:Ideal SNR as a function of input photons.

Because n_{dark} is a constant, the SNR increases linearly, i.e. 20dB/dec according to Eq. (2-6). At higher illumination levels, the dominant noise source is the photon shot noise, which is the square root of the input photons. Thus, the SNR is given as Eq. (2-7) and therefore increases in 10dB/dec:

$$SNR = 20\log\left(\frac{N_{sig}}{n_{sig}}\right) = 20\log\left(\frac{N_{sig}}{\sqrt{N_{sig}}}\right) = 10\log N_{sig}\,[dB] \quad (2\text{-}7)$$

As can be seen from Eq. (2-7), the maximum SNR appears when the photodiode is saturated and completely determined by the maximum signal charge N_{sat}, i.e. the full-well capacity. In theory, the maximum SNR can be improved as long as the full-well capacity is increased. But this conclusion is based on the assumption that only the temporal noise is included in the noise level. However, the acquired SNR in reality is normally extracted from an actual pixel array, the spatial noises/offsets of which also contribute to the total noise level. In particular, the photon response non-uniformity (PRNU) limits the maximum SNR because it grows linearly with the input photons while the photon shot noise is only the square root dependency [2.1]. For example, for cases in which PRNU is linear at 1%, the maximum SNR, including PRNU, can never exceed 40dB, no matter how large the full-well becomes. Details regarding PRNU and spatial noise of image sensors are discussed later in this chapter.

In conclusion, the SNR represents a fundamental criterium for the image quality in terms of noise. Although in theory the maximum SNR is determined by the pixel full-well capacity, in reality, particularly for still-imaging applications, it is important to improve the spatial noise distribution among the complete imager in order to achieve a higher SNR.

2.1.4 Conversion Gain

Up to now, the performance of CMOS image pixels has been analyzed and characterized in electrons or photons. However, the output of pixels is always an analog signal which in most cases is an

analog voltage. Thus, there is an important process that converts the light signal into an electronic signal inside the pixels. Conversion gain is the parameter which represents the efficiency of this process.

In general, the conversion gain expresses how much voltage change is produced by one electron, at either the photon-sensing node or the charge detection node, depending on the pixel structure. The conversion gain is given as:

$$CG = \frac{q}{C_{CG}} \left[\mu V / e^{-} \right] \tag{2-8}$$

where C_{CG} is the capacitance of the sensing node or the charge detection node.

The conversion gain may be one of the most important parameters of a CMOS imager pixel. The linearity and uniformity of the pixel response, light sensitivity, and the pixel random noise are all influenced by its value and distribution. The characteristics of the conversion gain among different pixel types are discussed in the last section of this chapter.

2.2 Overview of Fixed-Pattern Noise in CMOS Image Sensors

Usually, an image sensor continuously produces a two-dimensional stream of information. Therefore, there are two types of noise which represent the variation in both spatial and temporal domain. The variation of the output from different pixels under the same illumination condition is referred to as fixed-pattern noise (FPN), because that it is fixed in a spatial position. The noise which fluctuates over time from an individual pixel is called random or temporal noise.

In this section, FPN is discussed with a focus on its physical origin and its evaluation method.

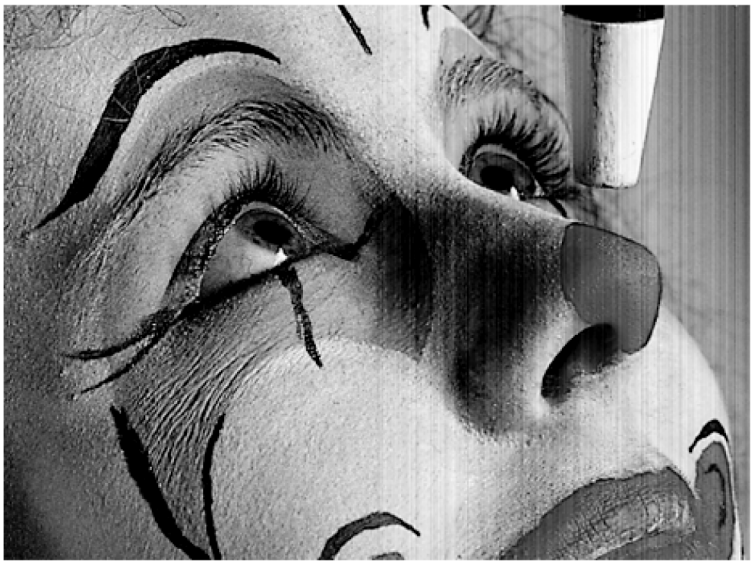


Figure 2-5:Simulated image containing both pixel and column FPN. The left half image contains 3% pixel FPN, the right half contains 3% of column FPN, taken from [2.9].

2.2.1 Fixed-Pattern Noise in Dark

FPN in dark is normally considered an offset variation of pixel outputs because it is a constant for a given pixel at a fixed integration time. There are two main sources causing this offset FPN, the mismatch of in-pixel or column-level transistors, and the dark current generated inside the pixel.

The imperfection of the fabrication process introduces significant mismatch to the transistor parameters, e.g. the threshold voltage spread of transistors made in a 0.18μm process is up to tens of milli-volt [2.8]. This non-uniformity causes spatial offset variations among the entire pixel array. In CMOS imagers, transistors are used inside the pixel to either reset the photodiode, or

to amplify the photon-generated charges. The mismatch of these transistors induces pixel-level FPN.

However, there is an efficient way of eliminating this type of FPN. It is called double sampling (DS): by sampling the pixel output twice both before and after the charge integration and subtracting these two samples, the offset caused by the in-pixel transistor mismatch can be removed completely.

Another typical mismatch-caused FPN appears in the column circuitry of the pixel array. Figure 2-5 shows a simulated image containing both pixel and column FPN. As can be seen, the column FPN introduces stripes onto the captured image. Unfortunately, compared to the pixel FPN, the column FPN is often more noticeable to the human eye and it is more difficult to be removed through circuitry solutions. Because of this, the column FPN is mostly suppressed or eliminated in the digital domain during the image processing procedures.

In terms of pixel FPN, the mismatch-induced FPN can be eliminated by the double sampling operation, where the actual primary FPN source is the dark current generated inside the pixel. Even without illumination, there are electron-hole pairs being generated from the photo-sensing region. This response from a pixel that is not illuminated is called dark current; the total amount of the collected dark charge is called dark count. Since the dark current of each individual pixel is not uniform over the complete pixel array, the induced FPN cannot be eliminated easily. Because of its importance, all of chapter 3 is dedicated to explaining and analyzing the exact origins and mechanisms of the dark current in CMOS imagers.

Dark FPN is normally evaluated by so-called dark signal non-uniformity (DSNU), which represents the distribution of the dark voltage output of each individual pixel of the whole array. Since the extracted DSNU is normalized with respect to the dark current, it is independent from the exposure time.

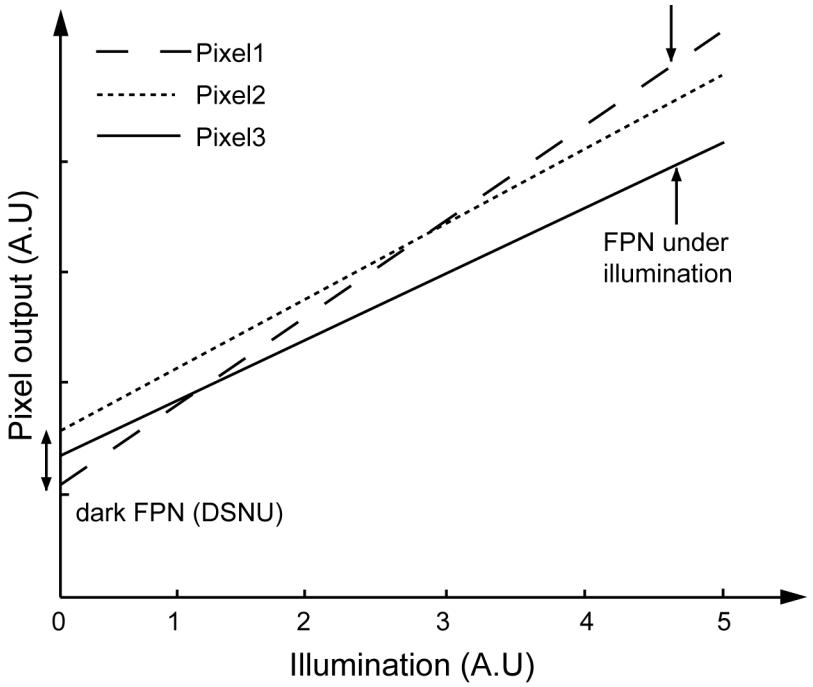


Figure 2-6:Pixel photon-responsivity in an ideal case, ignoring any non-linearity effects.

2.2.2 Fixed-Pattern Noise under Illumination

Contrary to dark FPN, the magnitude of FPN under illumination is often observed to be proportional to the illumination condition. Thus, instead of offset FPN, it is often treated as gain FPN. Figure 2-6 shows the photo-responsivity of several pixels in an ideal situation. It illustrates the relation between the dark FPN (offset) and the FPN under illumination. As can be seen, although the FPN under illumination is mainly due to the photo-response gain mismatch of different pixels, it does, however, also included the influence from the dark FPN as well. Thus, it is important to take DSNU into account when analyzing FPN under illumination.

Determining the sources of gain FPN is somewhat complex. They can be divided into three different categories. First, there are

light collection variations, e.g. the non-uniformity of the micro-lens efficiency. Secondly, the photon-electron conversion also introduces non-uniformities, e.g. the varying of the effective fill factor of each pixels. Third, gain FPN may also be induced by the variations during the electron-voltage conversion process, e.g. the non-uniformity of the conversion gain.

Therefore, to know exactly what the dominant source of the gain FPN is proves to be rather difficult. Because of this, the gain FPN is often corrected by using a gain map or a look-up table. This means that the gain of each individual pixel needs to be calibrated and stored in advance during the fabrication phase.

To evaluate FPN under illumination, the photo-response non-uniformity (PRNU) is used. The definition of PRNU is the same as for DSNU except that it is measured under an illumination condition instead of in the dark. However, as mentioned above, it is important to be aware that the FPN under illumination also includes the influence from the dark FPN. Thus, to obtain an accurate PRNU value, the DSNU needs to be subtracted from the original image data before calculating PRNU. Because PRNU represents the gain FPN under illumination, it should be proportional to the exposure time.

2.3 Overview of Temporal Noise in CMOS Image Sensors

As explained in the previous section, FPN is fixed for a given pixel, which makes it relatively easy to be eliminated by image processing steps in digital domain. This leaves temporal noise as the major limiting performance factor in terms of noise for CMOS imagers. In this section, the physical origins of different noise sources presented in the CMOS image sensor pixels are described. In addition, techniques to reduce or eliminate specific noise sources are briefly explained as well.

2.3.1 Photon Shot noise

Photon shot noise is the noise associated with the random arrival of photons. It is an expression of a natural process rather than pixel design or fabrication technology. Thus, photon shot noise is the most fundamental noise among all the noise sources found in all imagers.

The amount of photon-generated carriers in the photo-sensing area is also a random variable. If the photodetector is exposed to a perfectly uniform light source, the time between photon arrivals is governed by the Possion statistics [2.10]. Therefore, the magnitude of the photon shot noise equals the square root of the mean number of electrons stored in the photo-sensing area. It is given by:

$$n_{photon} = \sqrt{N_{sig}} \tag{2-9}$$

The rms noise voltage due to photon shot noise is therefore given by:

$$\overline{V_{photon}} = CG \cdot \sqrt{N_{sig}} = \frac{q}{C_{CG}} \sqrt{N_{sig}} \tag{2-10}$$

Interestingly, although Eq. (2-10) suggests that an increase in the capacitance C_{CG} lowers the photon shot noise, it can be seen from Eq. (2-7) that the imager SNR is in fact independent of C_{CG} and solely determined by the signal level when photon shot noise dominates the readout noise floor. In other words, the higher the signal level (i.e. the photo-generated charge), the higher the sensor's SNR.

Unlike other noise sources in CMOS imagers, photon shot noise is a unique noise which has a constant relationship to the illumination level. Moreover, because it is the result of a fundamental physical law instead of the actual sensor design, its existence is guaranteed in all image sensors. Therefore, its square root dependency to the signal level is used very widely to characterize sensor performance.

For example, the conversion gain of a pixel can be extracted based on Eq. (2-10). If photon shot noise dominates the noise floor,

the signal output voltage and the rms readout noise can be written as:

$$\begin{aligned} V_{sig} &= CG \cdot N_{sig} \cdot A \\ \overline{V_{noise}} &= CG \cdot \sqrt{N_{sig}} \cdot A \end{aligned} \tag{2-11}$$

where A is the voltage gain of the analog, or digital circuitry following photo-sensing element. Thus, the conversion gain CG can be calculated by:

$$CG = \frac{\overline{V_{noise}^2}}{V_{sig} \cdot A} \tag{2-12}$$

As shown, if the voltage gain A is known, the value of the conversion gain can be easily extracted through Eq. (2-12). An accurate calculation of the conversion gain is critical in imager characterization procedures since there are many performance parameters derived from it. Luckily, the unique property of photon shot noise offers the possibility of measuring the conversion gain.

2.3.2 Dark Current Shot Noise

As explained in the previous section, electron-hole pairs are generated in the photo-sensing element even without illumination. It is called dark current. This generation mechanism is a thermal process that depends exponentially on temperature. Similar to photon shot noise, dark current generation also obeys Poisson statistics. Thus, dark current shot noise can be given by:

$$n_{dc} = \sqrt{N_{dc}} \tag{2-13}$$

where N_{dc} is the mean value of the dark count.

The only approach to reduce the dark current shot noise is to lower the dark count. Details of the dark current generation mechanism and reduction techniques will be discussed in chapter 3.

2.3.3 Reset Noise

As shown in Figure 2-3, the photodiode needs to be reset by the switch "reset" every time before the charge integration starts. This reset operation effectively samples a bias voltage V_{res} onto the photodiode capacitance C_{PD}. Such a sampling operation obviously introduces sampling noise. It is normally referred to as "*kTC*" noise [2.14] in analogue circuitries or "reset" noise in CMOS imagers.

The reset noise, in fact, originates from the thermal noise of the the "reset" switch in Figure 2-3, which is often implemented by a nMOS transistor. During the "on" period, this nMOS transistor can be considered as a resistance which contains thermal noise. This noise is afterwards sampled and held by the capacitor C_{PD} after the transistor has been switched off. Thus, the noise power is given by integrating the thermal noise power over all frequencies. The reset noise in rms voltage can be given as:

$$\overline{V_{res}} = \sqrt{\int_0^\infty 4kT \cdot \frac{R}{1+(2\pi f RC_{PD})^2} \cdot df} = \sqrt{\frac{kT}{C_{PD}}} \qquad (2\text{-}14)$$

where R is the on-resistance of the nMOS switch, T is temperature, and f is frequency.

The noise charge in number of noise electrons can therefore be given as:

$$\overline{e_{res}} = \frac{C_{PD} \cdot \overline{V_{res}}}{q} = \frac{\sqrt{kTC_{PD}}}{q} \qquad (2\text{-}15)$$

At first glance, Eq. (2-14) and Eq. (2-15) seem controversial since they suggest a totally opposite dependency of the noise magnitude on the photodiode capacitance. This is because C_{PD} modulates not only the noise magnitude itself but also the efficiency of noise charge conversion to noise voltage. In Eq. (2-15), although the reset noise in electrons is proportional to the square root of C_{PD}, the noise charge to noise voltage conversion ratio is in inverse proportional to C_{PD}. Thus, the acquired noise voltage decreases if C_{PD} increases. Since the pixel output is eventually in voltage, the photodiode capacitance is expected to be as big as possible in terms of lowering the reset noise voltage.

However, although reset noise does benefit from a higher C_{PD}, there are imager performance parameters which may be damaged by increasing the photodiode capacitance, e.g. the light sensitivity. Moreover, in CMOS imagers, the required (small) pixel size usually constitutes an upper limit to C_{PD}. Thus, it is not really practical to significantly reduce the reset noise by increasing C_{PD}. These constrains make reset noise the dominant noise source in most CMOS imager pixels under low illumination conditions.

There is, however, a very efficient approach to eliminate this noise source, which is called correlated double sampling (CDS) [2.13]. The concept of CDS is based on the following analysis, for which is it assumed that $x_1(t)$ and $x_2(t)$ are two noise waveforms in the time domain and P_1 and P_2 are their noise power, respectively. If these two noise waveforms are subtracted from each other, the average of the resulting noise power is:

$$\begin{aligned} P_{av} &= \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}[x_1(t)-x_2(t)]^2\,dt \\ &= \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}x_1^2(t)dt + \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}x_2^2(t)dt \\ &\quad - \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}2x_1(t)x_2(t)dt \\ &= P_1 + P_2 - \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}2x_1(t)x_2(t)dt \end{aligned} \qquad (2\text{-}16)$$

where T stands for the period in time domain to extract the noise power. If both $x_1(t)$ and $x_2(t)$ originate from the same noise source, i.e. they are correlated, the noise power P_1 and P_2 are equal. Also, the integral term in Eq. (2-16) becomes $2P_1$. Thus, the average of the resulting noise power P_{av} becomes zero, or in other words, the noise is eliminated. If the two noise sources are independent from each other, i.e. non-correlated, the integral term in Eq. (2-16) vanishes [2.14], and the resulting noise power P_{av} is in fact the sum of both noise sources.

As a conclusion, if the two noise components are correlated, this noise can be eliminated completely by subtracting one from the other. In order to do so, two samples containing correlated noise sources are required. In CMOS imagers, the first sample is often the

pixel output taken right after the reset operation so that the reset noise can be measured. The next sample is taken after the photo-generated charge integration. Thus, the second sample contains the video signal voltage as well as the same reset noise. Since the reset noise from these two samples comes from the same reset operation, they are "correlated" and can be eliminated by CDS.

However, this technique is unfortunately not practical for all imager pixel structures. Its application and limitation on different pixel types will be discussed in the next section.

2.3.4 1/*f* Noise

Besides reset noise, 1/*f* noise is also a major noise source, which mainly appears from the in-pixel source follower transistor [2.11] in CMOS imagers. It was in 1955 [2.12] that the first 1/*f* type noise spectrum was shown by McWhorter. It is explained by McWhorter that the cause of this type of noise is due to the lattice defects at the interface of the Si-SiO_2 channel of the MOS transistor. These defects trap and de-trap the conducting carriers and therefore introduce a random current variation, which is the 1/*f* noise.

From a circuit designer's point of view, a simplified 1/*f* noise power can be given by [2.14]:

$$\overline{V_n^2} = \frac{K}{C_{ox} W L} \cdot \frac{1}{f} \tag{2-17}$$

where K is a process-dependent parameter, C_{ox} is the gate capacitance, and W and L are the width and length of the transistor. In fact, Eq. (2-17) seems quite simple since the only design consideration is the transistor dimension. However, it is important to be aware that it is only a simplified estimation of the 1/*f* noise power. In reality, particularly as the CMOS process scales down to deep sub-micron meter, the actual 1/*f* noise power becomes much more complex and involves more design factors [2.15].

The complexity of the 1/*f* noise spectrum is mainly due to an unclear noise mechanism. Although the origin of the 1/*f* noise is commonly accepted to be what McWhorter explained, it is still a

mystery how exactly such trapping and de-trapping processes manipulate the conducting current amplitude. In order to derive an accurate model that predicted $1/f$ noise power, the physical mechanism of this noise needs to be understood. McWhorter first proposed a so-called ΔN model, which illustrates that the conductivity variation due to $1/f$ noise is caused by the fluctuation of the number of the conducting carriers in the channel [2.16]. Unfortunately, this ΔN model cannot fully explain the $1/f$ noise spectrum, particularly in pMOS transistors [2.17]. In 1969 [2.18], Hooge proposed a so-called $\Delta\mu$ model, which considers $1/f$ noise to be caused by the fluctuations in the mobility of the charge carriers in silicon. The debate between the ΔN and $\Delta\mu$ models went on for years. There are also theories which intend to integrate the two models together [2.19][2.20]. Nowadays, although a unanimously accepted model is not yet available, it is commonly accepted that the ΔN model is better suited for n-type MOS transistors while the $\Delta\mu$ model is better suited for pMOS transistors [2.21].

Details about the influence of $1/f$ noise in CMOS image sensors will be discussed in Chapter 4.

2.3.5 Other Noise Sources

There are also other noise sources associated with CMOS imagers. Unlike the above-mentioned fundamental noise sources, these other sources depend significantly on the sensor design and fabrication technology. In other words, it is possible to avoid these noise sources through specific techniques.

Hot carrier (HC) effects may appear in the in-pixel source follower transistor. Because the source follower transistor is operated in saturation during the pixel readout, the conducting electrons may be accelerated by the high electrical field in the pinch-off region near the drain and become “hot” electrons. If the energy of these hot carriers goes beyond a certain threshold, excess electrons are generated through the impact-ionization process [2.22]. These excess electrons can be easily collected/absorbed by the photodiodes close by which thus introduce noise.

However, the HC noise only occurs when there is a conducting current present in the source follower transistor, i.e. only during the pixel readout period. To reduce the noise, or in other words, to reduce the possibility of HC effects, the pixel output sampling time can be reduced. Furthermore, the power supply of the source follower transistor can also be lowered to reduce the electrical field of the pinch-off region so that the impact-ionization process becomes less likely to occur.

Power supply coupling may also introduce pixel-level noise. For example, the supply coupling between the gate of the reset transistor and the photodiode introduces offset from the reset signal, i.e. pixel FPN. It may be removed through CDS, however that will introduce a problem for the global shutter operation [2.23].

In conclusion, because the pixel temporal noise varies in time instead of in a spatial domain, the reduction or elimination of this noise is often difficult. The resulting pixel readout noise floor sets the fundamental limit on imager performance, especially under low illumination conditions. In order to achieve superior image quality, it is essential to understand the origins of these temporal noises, find the dominant noise source, and reduce its noise power accordingly.

2.4 CMOS Image Sensor Pixel Circuits

Among CMOS imagers, two types of pixels are commonly used, i.e. the passive pixel sensor (PPS) and active pixel sensor (APS). The main difference is that an additional amplifier is used inside the APS pixels. APSs are able to offer lower noise levels and higher readout speeds. Since APSs have became the technology of choice for most of the CMOS imager applications, only APS pixel circuits are introduced here. This section is organized according to the different photo-sensing elements used in the pixel.

2.4.1 Photodiode Three Transistor (3T) Pixel

The three transistor (3T) pixel uses a p-n junction (photodiode) as the photon-sensing node. It was the most commonly used pixel

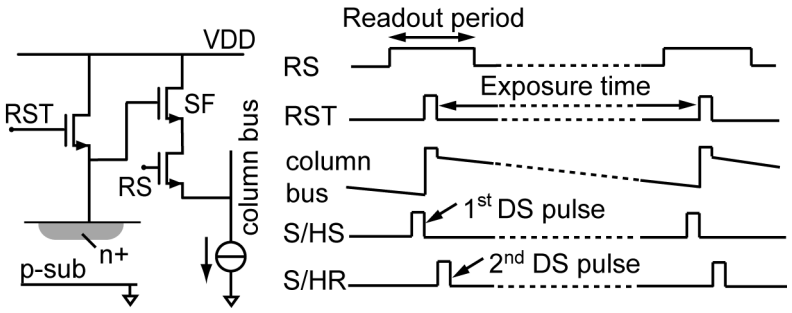


Figure 2-7:3T Pixel schematic with cross-section of the photodiode and timing diagram.

structures among all APS sensors. Although the photodiode-type pixel was first described already in 1968 [2.24], the first high-performance photodiode APS was implemented by JPL only in 1995. This revolutionary design adapts a 3T pixel structure and is still used today.

Figure 2-7 shows the pixel schematic with a cross-section of the photodiode and its timing diagram during exposure and readout periods. As can be seen, the pixel consists of three nMOS transistors. The potential of the photodiode is to reset to VDD through a reset transistor (RST). After that, the photon-generated charges are collected and converted into a voltage signal directly by the photodiode. The conversion gain is determined by the photodiode capacitance. The signal charge is amplified afterwards by the source follower transistor (SF) and readout through a row select transistor (RS).

As shown in Figure 2-7, the RST is switched off during exposure. The photodiode potential decreases because of the integration of the photon-generated electrons. The exposure operation ends when the RST is switched on. Before and after the photodiode is reset, the video signal and reset level on the column bus are readout sequentially by the sample-hold reset (S/HR) and sample-hold signal (S/HS) pulses from the double sampling circuitry in the column. By subtracting the reset level and video signal, the light intensity can be determined. Because of the double

sampling operation, the threshold mismatch of the SF transistors is removed so that the pixel FPN is lowered.

Since only three transistors are used inside the pixel, the fill factor of 3T pixels is improved compared to most of the other APS pixels. Moreover, because the photodiode can be reversely biased using a strong positive potential through RST, which results in a wide depletion region, both the quantum efficiency and full-well capacity for 3T pixels are excellent.

However, the temporal noise of 3T pixels is rather high. Because the pixel array is readout row-by-row and stored in the column structure, the double sampling operation, i.e. S/HR and S/HS pluses, needs to be completed within the rather short readout period, as shown in Figure 2-7. The two samples have to be implemented right before and after the photodiode reset operation. Thus, the two sampled signals, in fact, contain reset noise from different reset operations. As explained in the previous section, since the reset noise is non-correlated, this double sampling operation actually increases the resulting noise power. Therefore, in 3T pixel CMOS imagers, the *kTC* noise appears to be the dominant noise source.

As a result of this, a lot of effort has been spent on investigating and improving the reset noise in 3T pixels. Recent research proves that it is possible to reduce reset noise through a so-called “soft reset” techniques [2.26]. It is shown that if RST is switched on using the same voltage amplitude on its drain and gate, the resulting reset noise power in voltage square is actually less than *kT/C* but *kT/2C* because of a non-equilibrium transistor operation. A further noise reduction can be obtained by using an “active reset” technique [2.27][2.28]. A noise power reduction of five or six times lower than *kT/C* is reported. However, although these methods are able to reduce the reset noise significantly, they introduce limitations for other imager performance parameters. For example, the use of a soft reset may introduce image lag or non-linearity of the photo-response[2.29].

Moreover, although both soft reset and active reset are capable of lowering the reset noise, the remaining noise power is still the dominant noise source that limits the overall noise floor. Therefore,

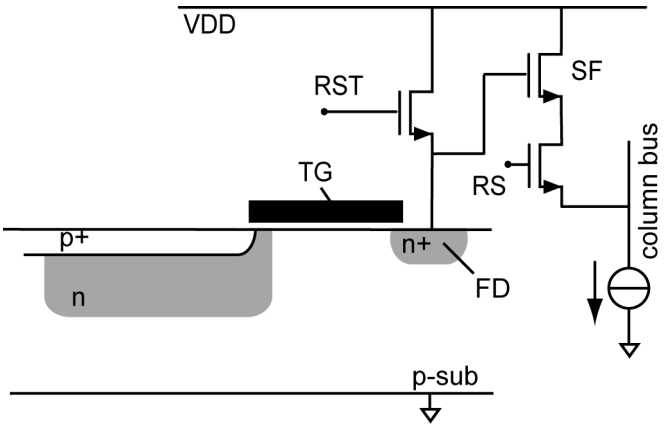


Figure 2-8:4T Pixel schematic with cross-section of the photo-sensing and charge transfer gate region.

the performance of 3T pixels is rather compromised in terms of temporal noise. This is exactly the reason why a pinned-photodiode 4T pixel is more commonly used for low noise applications.

2.4.2 Pinned-Photodiode Four Transistors (4T) Pixel

Pinned-photodiode (PPD) was first used as a photo-sensing element in CCD imagers to avoid incomplete charge transfer from the photodiode [2.30]. This structure was afterwards implemented in CMOS imagers in 1997 [2.31], when achieved a good spectral response and low dark current level.

Figure 2-8 shows the schematic of a PPD 4T pixel with the cross-section of the photo-sensing element, the charge transfer gate (TG), and the floating diffusion (FD). As can be seen, the photo-sensing element consists of two p-n junctions: the p+/n junction close to the surface and the n/p-sub junction in the silicon bulk. Compared to the photodiode in 3T pixels, the operation of this PPD photon-sensing component is rather complex and deserves extra attention.

Figure 2-9 shows the potential diagram of the PPD, the TG and FD during charge integration, and the FD reset and charge transfer/

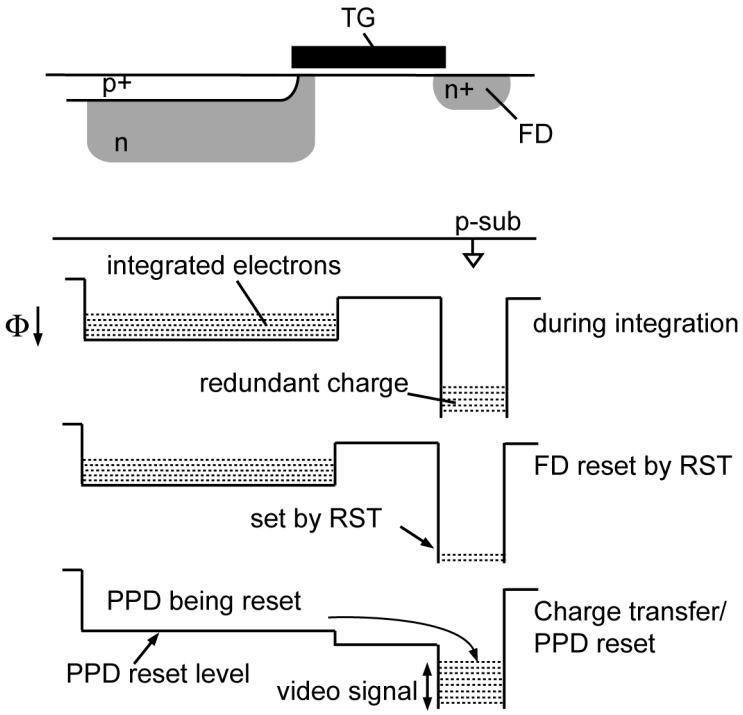


Figure 2-9:Potential diagram of the PPD, TG and FD during charge integration, FD reset and charge transfer/PPD reset.

reset operation. As shown, the photo-generated electrons are generated and collected in the PPD during the exposure time. After that, the FD needs to be reset first to remove any redundant charges. The reset level of the FD is determined by the reset mode of RST transistor, e.g. a soft reset, as mentioned previously. In the end, the TG is switched on so that the electrons stored in the PPD flow to the FD. Meanwhile the PPD is automatically reset and ready for the next integration operation.

The PPD reset level (also called pinning voltage), shown in the Figure 2-9, is completely determined by PPD itself instead of RST operation or FD potential, as long as the photo-generated charges are completely transferred. This operation principle indeed establishes a rather strict requirement on the PPD fabrication. To

acquire a well-controlled PPD reset level and avoid transfer inefficiency, the PPD must be fully depleted, i.e. the depletion region of the surface p+/n junction needs to merge with that of the n/p-sub junction in the bulk. In order for this to happen, the doping profiles of both the p+ pinned layer and the n region have to be accurately controlled and optimized.

Besides PPD itself, the charge transfer efficiency also depends on the FD potential. After the charge transfer operation, the potential of the resulting signal level on the FD needs to be higher than that of the PPD reset level. Otherwise, charges in FD may flow back to the PPD and cause so-called “charge sharing”. Because of this, the FD reset level should be as high as possible. In addition, the conversion gain needs to be adjusted as well, since it modulates how much potential is generated from the transferred charges. The conversion gain of PPD 4T pixels is determined by the FD capacitance. Thus, compared to 3T pixels, of which the conversion gain is set by the photodiode capacitance, the conversion gain of 4T pixels is normally much higher, which is attractive when obtaining high light sensitivity.

Although the fabrication of PPD 4T APS is sometimes a considerable challenge, this type of pixel is becoming the most popular design for high-quality image applications [2.32]. That is due to its significantly improvement on sensor performance, particularly in terms of temporal noise.

As explained above, 3T pixels suffer from reset noise because of the non-correlated double sampling. However for a PPD 4T APS, the reset noise can be eliminated completely. As shown in Figure 2-9, the FD is reset immediately before the charge transfer operation, simultaneously while this reset level of FD is sampled and held for CDS operation. After the charge is transferred from the PPD, the resulting video signal is sampled again. In this way, the reset noise in these two samples is from the same reset phase and therefore can be removed completely by subtracting the two samples from each other. By eliminating the reset noise, the dark temporal noise level of PPD 4T APS is dramatically reduced. The remaining noise is dominated by the $1/f$ noise from the in-pixel source follower transistor [2.33][2.34].

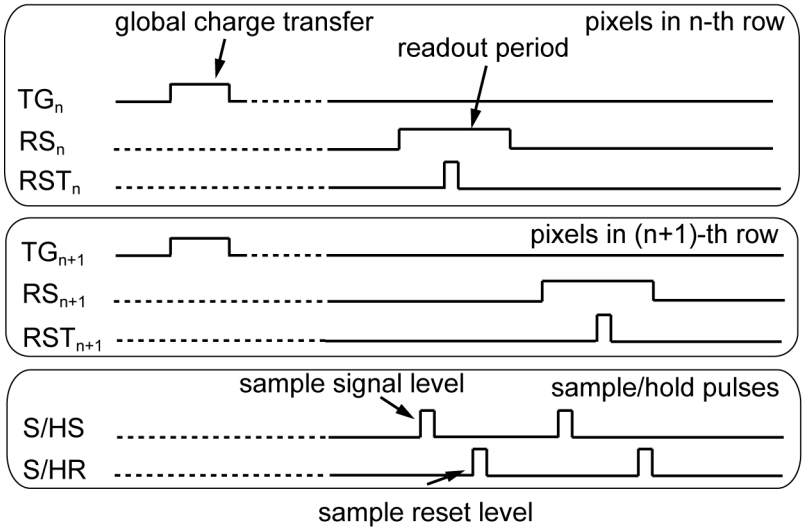


Figure 2-10:Timing diagram during pixel readout period for two adjacent rows in a PPD 4T APS opearating in global shutter mode.

Another important advantage of PPD 4T pixels is that they can operate not only in a rolling shutter but also in a global shutter mode (snapshot). This feature is very important for high-speed imaging applications, since it enables the ability to capture fast-moving objects without image distortion. Figure 2-10 shows the readout timing diagram of two adjacent rows of a PPD 4T APS in the global shutter operation mode. As can be seen, since the integration time of all rows has to start and end at exactly the same moment, the charge transfer operations (TG pulses) for all rows happen simultaneously. However, regardless of whether this occurs in rolling shutter mode or global shutter mode, the pixel readout scheme has to follow a row-by-row sequence. That means when the n-th row is selected, the video signal has already been stored on the FD and has to be sampled first. After that, the FD is reset and the reset level is sampled again. Clearly, such a readout scheme produces non-correlated samples in terms of reset noise. Thus, in order to perform global shutter operation, the pixel temporal noise is sacrificed.

Because the pinning voltage of the PPD is decided by its doping profile, its depletion region cannot be adjusted with a biasing voltage as the photodiode in a 3T pixel. Thus, the full-well capacity of a PPD is generally smaller than that of a reverse-biased photodiode, giving the same pixel size and fill factor. However, since the PPD depletion region is closer to the top surface, the QE is improved for light with a shorter wavelength. Moreover, the p+ pinned layer significantly reduces the dark current generated from the top Si-SiO_2 interface.

In conclusion, today PPD 4T pixels are one of the most commonly used pixel structures in CMOS imagers. They achieve good blue response (shorter wavelength), extremely low dark current and most importantly, very low dark temporal noise levels. Also, they can be implemented in a global shutter mode, but unfortunately at the expense of losing CDS capability which consequently increases the noise level.

2.4.3 Other Pixel Designs

Besides 3T and 4T pixels, other types of pixels are used in CMOS imagers as well. For example: a five transistor (5T) pixel

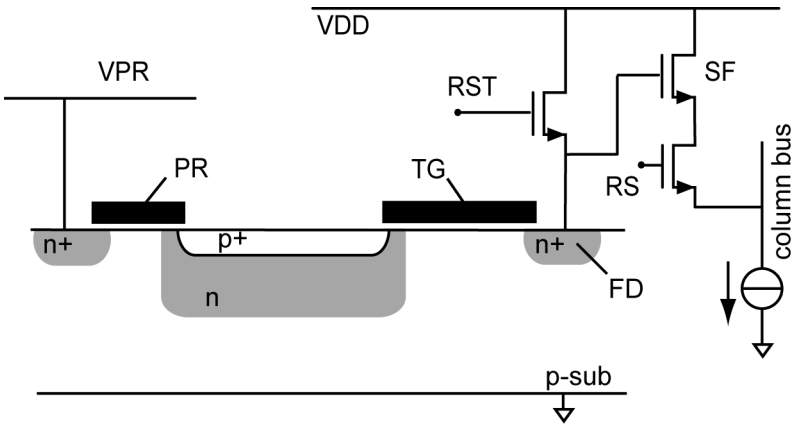


Figure 2-11:5T pixel schematic with cross-section of the photo-sensing, TG and PR transistors

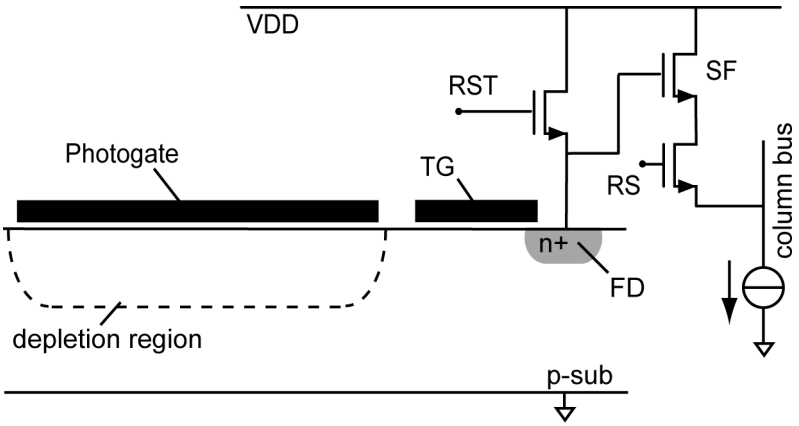


Figure 2-12:Photogate pixel schematic with cross-section of the photo-sensing element and TG transistors.

based also on a PPD photo-sensing element was developed in 2002 [2.35]. As shown in Figure 2-11, the pixel structure of a 5T pixel is very similar to that of a PPD 4T pixel, except that there is an extra switch PR. The imager is able to start an integrating operation by resetting the PPD through PR, while the previous frame is read out at the same time. This approach increases the imager frame rate significantly during a global shutter operation. However, it does not solve the non-correlated double sampling problem. Thus, the noise level of a 5T APS in global shutter mode is still high. There are also pixel designs that consist of even more transistors than 5T in order to add extra functionalities inside the pixel [2.36]. However, from the noise perspective, adding transistors equals adding more noise sources. Thus, the noise of the PPD 4T pixel is indeed the lowest.

Some pixel designs also implement different photo-sensing elements, e.g. the photogate pixel shown in Figure 2-12. As can be seen, the depletion region of the photo-sensing element of a photogate pixel is created by positively biasing the photogate, as in that of CCD sensors. In terms of pixel temporal noise, the reset noise of photogate pixels can also be eliminated with CDS. Thus, the overall noise floor is low. However, photogate pixels have one distinct disadvantage. The presence of a gate on top of the

photo-sensing region significantly decreases the QE of light with a shorter wavelength because photons are absorbed by the poly-silicon gate.

In conclusion, flexible pixel design is one of the most attractive advantages of using CMOS imagers. The pixel circuitry can be designed and optimized to satisfy certain performance parameters. Since the purpose of this thesis is to analyze and lower noise in CMOS imagers, in most cases sensors with PPD 4T pixel design are used for analysis and characterization because of its outstanding noise performance and excellent QE compared to photogate pixels.

2.5 References

[2.1] J. Nakamura et al., *Image Sensors and Signal Processing for Digital Still Cameras*, Taylor & Francis, pp. 79-91, 2006.

[2.2] S.K. Mendis et al., "CMOS Active Pixel Image Sensors for Highly Integrated Imaging Systems", *IEEE Journal of Solid-State Circuits*, Vol. 32, pp. 187-197, 1997.

[2.3] S. Mendis et al., "CMOS Active Pixel Image Sensor", *IEEE Transactions on Electron Devices*, Vol. 41, pp. 452-453, 1994.

[2.4] B.C. Burkey et al., "The Pinned Photodiode for an Interline Transfer CCD Imager", *Technical Digest IEDM*, pp. 28-31, San Francisco, US, Dec. 1984.

[2.5] R.M. Guidash et al., "A 0.6μm CMOS Pinned Photodiode Color Imager Technology", *Technical Digest IEDM*, pp. 927-929, Washington DC, US, Dec. 1997.

[2.6] T. Yamada et al., "A 140dB-Dynamic-Range MOS Image Sensor with In-Pixel Multiple-Exposure Synthesis", *Technical Digest ISSCC*, pp. 50-51, San Francisco, US, Feb. 2008.

[2.7] S. Chamberlain et al., “A Novel Wide Dynamic Range Silicon Photodetector and Linear Imaging Array”, *Journal of Solid-State Circuits*, Vol. 19, No. 1, pp. 41-48, 1984.

[2.8] *CMOS 18FLV Design Manual*, Version 3.04, Philips, Jan. 2004.

[2.9] M.F. Snoeij et al., “A CMOS Imager with Column-Level ADC Using Dynamic Column Fixed-Pattern Noise Reduction”, *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 12, pp. 3007-3015, 2006.

[2.10] J.R. Janesick, “Scientific CCDs”, *Optical Engineering,* Vol. 26, pp. 692-714, Aug. 1987.

[2.11] K. Findlater et al., “SXGA Pinned Photodiode CMOS Image Sensor in 0.35μm Technology”, *Technical Digest ISSCC*, pp. 218-219, San Francisco, US, Feb. 2003.

[2.12] A.L. McWhorter, *1/f Noise and Related Surface Effects in Germanium*, PhD dissertation, MIT, Cambridge, MA, 1955.

[2.13] A.J. Blanksby et al., “Performance Analysis of a Color CMOS Photogate Image Sensor”, *IEEE Transactions on Electron Devices*, Vol. 47, pp. 55-64, Jan. 2000.

[2.14] B. Razavi, *Design of Analog CMOS Integrated Circuits*, Boston, McGraw Hill, 2001.

[2.15] A.J. Scholten et al., “New 1/f Noise Model in MOS Model 9, Level 903”, *Compact Modeling report*, Nat.Lab. NL-UR 816/98, Philips Electronics, 1998.

[2.16] A.L. McWhorter, *Semiconductor Surface Physics,* Philadelphia, Univ. Pennsylvania Press, 1957.

[2.17] S. Tedja et al., “Noise Spectral Density Measurements of a Radiation Hardened CMOS Process in the Weak and Moderate Inversion”, *IEEE Transactions on Nuclear Science*, Vol. 39, pp. 804-808, 1992.

[2.18]F.N. Hooge, "1/*f* Noise is No Surface Effect", *Physics Letters A*, Vol. 29, No. 3, pp. 139-140, April 1969.

[2.19]H. Mikoshiba, "1/*f* Noise in n-Channel Silicon-Gate MOS Transistors", *IEEE Transactions on Electron Devices*, Vol. ED-29, pp. 965-970, June 1982.

[2.20]K.K. Hung et al., "A Unified Model for the Flicker Noise in Metal-Oxide-Semiconductor Field Effect Transistors", *IEEE Transactions on Electron Devices*, Vol. 37, pp. 654-665, March 1990.

[2.21]J. Chang et al., "Flicker Noise in CMOS Transistors from Subthreshold to Strong Inversion at Various Temperatures", *IEEE Transactions on Electron Devices*, Vol. 41, pp. 1965-1971, Nov. 1994.

[2.22]H.S. Wong, "Experimental Verification of the Mechanism of Hot-Carrier-Induced Photon Emission in n-MOSFET's Using an Overlapping CCD Gate Structure", *IEEE Electron Device Letters*, Vol. 13, No. 8, pp. 389-391, Aug. 1992.

[2.23]B. Pain. "Device Noise in CMOS Imagers", *ISSCC 2007 Forum, Noise in Imaging Systems*, San Francisco, US, Feb. 2007.

[2.24]P. Noble, "Self-Scanned Silicon Image Detector Arrays", *IEEE Transactions on Electron Devices,* Vol. 14, pp. 202-209, 1968.

[2.25]R.H. Nixon et al., "128x128 CMOS Photodiode-Type Active Pixel Sensor with On-Chip Timing, Control and Signal Chain Electronics", *Proceedings of SPIE*, Vol. 2415, pp. 117-123, Bellingham, US, Feb. 1995.

[2.26]H. Tian et al., "Analysis of Temporal Noise in CMOS Photodiode Active Pixel Sensor", *IEEE Journal of Solid-State Circuits,* Vol. 36, pp. 92-101, Jan.2001.

[2.27]B. Fowler et al., "Low Noise Readout Using Active Reset for CMOS APS", *Proceedings of SPIE*, Vol. 3965, pp. 126-135, San Jose, US, May 2000.

[2.28]B. Pain et al., “Reset Noise Suppression in Two-Dimensional CMOS Photodiode Pixels through Column-Based Feedback Reset”, *Technical Digest IEDM*, pp. 809-812, San Francisco, US, Dec. 2002.

[2.29]B. Pain et al., “An Enhanced-Performance CMOS Imager with a Flushed-Reset Photodiode Pixel”, *IEEE Transactions on Electron Devices*, Vol. 50, No. 1, pp. 48-56, Jan. 2003.

[2.30]N. Teranishi et al., “No Image Lag Photodiode Structure in the Interline CCD Image Sensor”, *Technical Digest IEDM*, pp. 324-327, San Francisco, US, Dec 1982.

[2.31]R.M. Guidash et al., “A 0.6µm CMOS Pinned Photodiode Color Imager Technology”, *Technical Digest IEDM*, pp. 927-929, Washington DC, US, Dec. 1997.

[2.32]A.J.P. Theuwissen, “The Hole Role in Solid-State Imagers”, *IEEE Transactions on Electron Devices,* Vol. 53, No. 12, pp. 2972-2980, 2006.

[2.33]B. Pain et al., “Excess Noise and Dark Current Mechanisms in CMOS Imagers”, *IEEE Workshop in CCD and Advanced Image Sensors*, pp. 145-148, Karuizawa, Japan, June 2005.

[2.34]J.Y. Kim et al., “Characterization and Improvement of Random Noise in 1/3.2” UXGA CMOS Image Sensor with 2.8µm Pixel Using 0.13µm-Technology”, *IEEE Workshop in CCD and Advanced Image Sensors*, pp. 149-152, Karuizawa, Japan, June 2005.

[2.35]E. Fox et al., “Pinned Photodiode Five Transistor Pixel”, *US Patent 6566697*.

[2.36]S. Lauxtermann et al., “Comparison of Global Shutter Pixels for CMOS Image Sensors”, *International Image Sensor Workshop*, pp. 82-85, Ogunquit, US, June 2007.

Chapter 3

Dark Current in CMOS Image Sensors

As previously explained, the primary source of pixel FPN in CMOS imagers is the dark current generation inside the pixels. Although it may not be noticeable during normal pixel operation, especially with an extremely small pixel size, it becomes a salient factor in applications which require long integration time and low illumination. In order to minimize the pixel-level FPN, it is essential to reduce the dark current generated in the pixels.

In this chapter, the physical mechanism of various dark current types will first be explained in section 3.1. By theoretical modeling the dark current density, their generation dependencies will be shown. Important approaches to distinguishing different dark current mechanisms will also be discussed also in this section. In section 3.2, different dark current sources of conventional CMOS APS pixels will be analyzed in detail. The individual dark current contribution from the photodiode, transfer gate (TG), and the floating diffusion (FD) will be shown. In section 3.3, conclusions will be drawn on important considerations of designing low dark current pixels. Basic design trade-offs will be indicated as well.

3.1 Dark Current Generation Mechanisms

Dark current stands for the signal response when a photo-detector is not exposed to light. It is a very complex process that is related to many design and technology factors, e.g. the silicon defect density, the electric field of the photo-sensing element, and operation temperature. The total dark current in CMOS imagers normally consists of several components which have different physical origins. Thus, it is essential to understand their individual generation mechanisms and characteristics in order to allow further optimization of dark performance.

In this section, all different dark current sources that could appear in CMOS imagers are divided into two categories according to their generation locations. Figure 3-1 shows the dark current composition in a standard n+/p-sub junction. Although terms like "diffusion" and "thermal generation" are widely used nowadays, they are somehow confusing when distinguishing different dark current components. "Diffusion" refers to the way in which the minority carriers are being collected. However, "thermal generation" directly describes how the free minority carriers are generated. Thus, in the following analysis, all dark current components will be divided according to the location from which

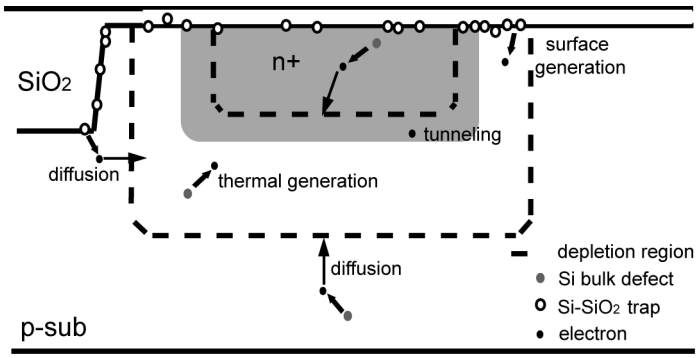


Figure 3-1:Different dark current generation mechanisms in a n+/p-sub junction.

they are being generated, as can be seen from Figure 3-1, i.e. either inside or outside the depletion region of the p-n junction.

3.1.1 Dark Current Generated in the Depletion Region

In CMOS imagers, the photodiode is normally reverse-biased in order to acquire a wide depletion region and to achieve good QE. The electric field across this depletion region is therefore very strong, which means that if there are free minority carriers being generated inside the depletion region, they are collected very efficiently through a drift process. Thus, it is reasonable to assume that all generated carriers from the depletion region (regardless of their different generation mechanisms) are collected before being recombined, which contributes to the total dark count.

Thermal generation: As shown in Figure 3-1, thermal generation stands for the dark current generation according to the conventional Shockley-Read-Hall mechanism. The minority carriers are thermally generated from the generation-recombination process in the silicon.

Because of the existence of the silicon bandgap, a direct generation-recombination process between the silicon valence band and the conduction band is very unlikely to occur without supplying extra energy, e.g. without illumination at room temperature. Instead, the dominant process is the indirect transition through an defect (energy state) located in the silicon bandgap [3.1]. In other words, those energy states act as stepping stones inside the silicon bandgap and assist the electrons in the valence band being excitated to the conduction band, therefore making it easier for dark current generation.

Because the junction is reverse-biased, the minority carrier concentrations in the depletion region are lower than the equilibrium concentrations, i.e. being depleted. Therefore, the capture (recombination) process is negligible and the emission (generation) process of minority carriers is dominant in order to restore the system to equilibrium. The rate of electron-hole pair generation inside the depletion region can be obtained with [3.2]:

$$G = \left[\frac{\sigma_p \sigma_n \upsilon_{th} N_t}{\sigma_n \exp\left(\frac{E_t - E_i}{kT} \right) + \sigma_p \exp\left(\frac{E_i - E_t}{kT} \right)} \right] n_i \qquad (3\text{-}1)$$

where σ_n is the electron capture cross section, σ_p is the hole capture cross section, υ_{th} is the thermal velocity of either electrons or holes (assuming they are equal), N_t is the density of the generation centers (silicon defects), E_t is the defect energy level, E_i is the intrinsic energy level, k is Boltzmann's constant, and T is the absolute temperature.

Considering a simple case where $\sigma_n = \sigma_p = \sigma_o$, Eq. (3-1) can be written as:

$$G = \frac{\sigma_o \upsilon_{th} N_t n_i}{2 \cosh\left(\frac{E_t - E_i}{kT} \right)} \qquad (3\text{-}2)$$

As shown in Eq. (3-2), the generation rate reaches a maximum value at $E_t = E_i$ and reduces exponentially as E_t moves away from the middle band. Consequently, only the silicon defects whose energy states are close to half bandgap contribute the most to the total generation rate.

The dark current caused by thermal generation in the depletion region is:

$$J_{gen} = \int_0^W q G \, dx \approx q G W = \frac{q n_i W}{\tau_g} \qquad (3\text{-}3)$$

where W is the depletion width, q is the electronic charge, n_i is the intrinsic concentration, and τ_g is the generation life time and can be expressed as [3.2]:

$$\tau_g = \frac{2 \cosh\left(\frac{E_t - E_i}{kT} \right)}{\nu_{th} \sigma_o N_t} \qquad (3\text{-}4)$$

As shown in Eq. (3-3), the dark thermal generation current is proportional to the intrinsic concentration n_i. The temperature dependence of n_i is given by [3.3]:

$$n_i = \sqrt{N_C N_V}\,\exp\left(-\frac{E_g}{2kT}\right) \tag{3-5}$$

where N_C and N_V are the carrier densities and E_g is the energy bandgap. By combining both Eq. (3-5) and Eq. (3-3), it can be concluded that the temperature dependency of thermal generation current is proportional to the exponential value of a half silicon bandgap.

This conclusion is already being used widely nowadays to address and verify the thermal generation dark current component in both CCD and CMOS imagers. Figure 3-2 shows the Arrhenius plot of the dark current density measured from a CCD. The average dark current density is plotted as a function of $1000/T$ and the activation energy ΔE_t can be found with:

$$D = D_0 e^{\Delta E_t / kT} \tag{3-6}$$

where D is the dark current density and D_0 is a pre-factor. By matching the Arrhenius plot function and Eq. (3-6), ΔE_t equals exactly half of the silicon bandgap, which can be extracted from the measurement. This means that the dominant dark current generation mechanism within the measured temperature range is the thermal generation from the depletion region.

Eq. (3-3) is often used to characterize and predict the dark current produced from thermal generation. Since the properties of the silicon defects are completely determined by the fabrication technology, the only design factors to be considered in Eq. (3-3) is the depletion width, which is related to the applied bias voltage and junction doping concentrations. In order to reduce the dark current, the depletion width of the photodiode junction should be as small as possible. However, as mentioned before, it is essential to maintain a wide depletion region of the photodiode in order to achieve a good QE. As a result, thermal generation is one of the primary dark current generation mechanisms in CMOS imagers [3.3], particularly in 3T pixel structures using photodiodes. The most efficient way to improve it is to optimize the technology to reduce the defect densities.

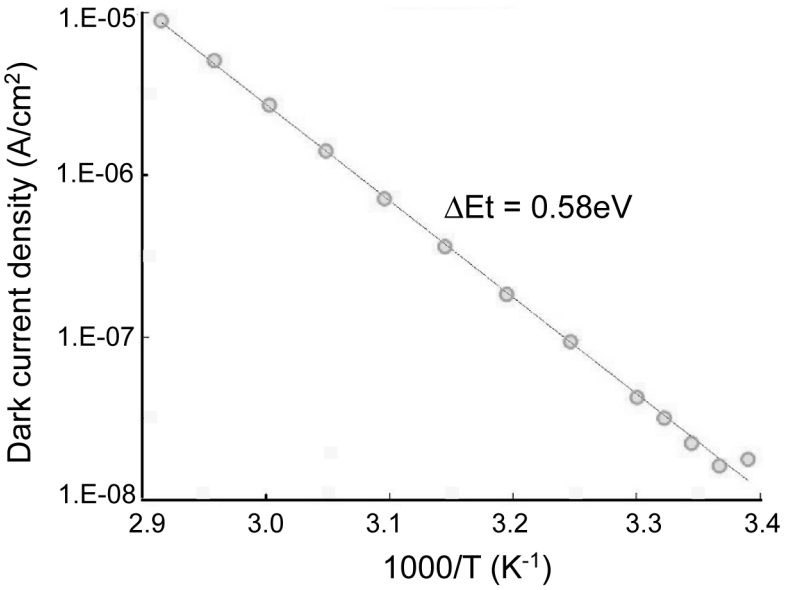


Figure 3-2:Arrhenius plot of the dark current measurement from CCD.

Surface generation: Physically the mechanism of surface generation dark current is the same as thermal generation. It is addressed separately because the density of the surface defects which are responsible for the surface generation is much higher than that of the energy states in silicon bulk. Thus, in modern CMOS imagers, surface generation dark current is also an important source of the total dark count and therefore deserves extra attention.

At the Si-SiO_2 interface, the lattice structure of silicon abruptly becomes discontinuous which introduces a large number of localized energy states. These energy states are normally called dangling bonds. The density of these surface states is mainly determined by the fabrication process. An efficient method for reducing its density is low-temperature hydrogen annealing [3.4]. However, it is rather an empirical procedure that requires extensive study. Thus, a perfect clean Si-SiO_2 interface is essentially impossible.

Moreover, the surface states density may change, sometimes even significantly increasing for certain application. For example, CMOS imagers are often used in a radiation environment in space applications, such radiation condition creates numerous interface states/trapped charges which therefore increases the dark current level [3.5][3.6]. Other process mechanisms can also affect interface state densities such as plasma etches and UV exposures [3.7]. Because of these limitations, studies on surface generation mechanism are crucial in terms of both performance and reliability issues.

The dark current generation rate because of surface generation can be expressed as:

$$J_{surf} = \frac{q S_0 n_i}{2} \tag{3-7}$$

where S_0 is the surface generation velocity which is determined by the trap energy and capture cross section.

Because the mechanisms for thermal and surface generation dark current are practically the same, it is hard to distinguish them from each other. For example, in the Arrhenius plot shown in Figure 3-2, both components exhibit a half-bandgap activation energy. However, because these two dark current sources are generated from different locations in the device, it is sometimes possible to separate them by biasing the device in certain conditions. For example, interfacial defects under the gate can be filled (passivited) by an inversion layer if the gate is biased in strong inversion. In that case, the measured dark current is mainly a contribution from thermal generation in the depletion region of the silicon bulk. Such an approach is applied in a so-called "gated diode", which has been widely used to characterize surface generation dark current [3.8]

Tunnelling: In the cases of thermal and surface generation, the carriers in the valence band gain enough energy, overcome the barrier (bandgap), and reach the conduction band through the intermediate energy states. However, under certain circumstances the valence band carriers can also "penetrate" the bandgap and reach the conduction band through a so-called "tunnelling" process.

Normally, a tunnelling process becomes pronounced when the depletion layer is narrow, thus appearing mainly in highly doped junctions. Figure 3-3 shows the tunnelling process in a heavily doped pn junction. As can be seen, if the valence band in the p-type is close to the energy level of the conduction band in n-type silicon, the two band edges are then positioned close together and the electrons can tunnel from left to right. This is called band-to-band tunnelling. If there is a mid-gap energy state in the depletion region, tunnelling becomes easier with the help of this intermediate state. This is called trap-assisted tunnelling process.

Both the band-to-band and trap-assisted tunnelling processes depend on the depletion width, i.e. the junction type and bias condition. Evidently even with a low voltage operation (around 3V), the tunnelling mechanism is the main dark current generation mechanism for a n+/pwell/psub junction [3.3].

The band-to-band and trap-assisted tunnelling current density dependency on the electric field can be expressed as [3.3]:

$$J_{btb} \sim V \left(\frac{F_m}{F_0} \right)^{3/2} \exp\left(-\frac{F_0}{F_m} \right) \qquad (3\text{-}8)$$

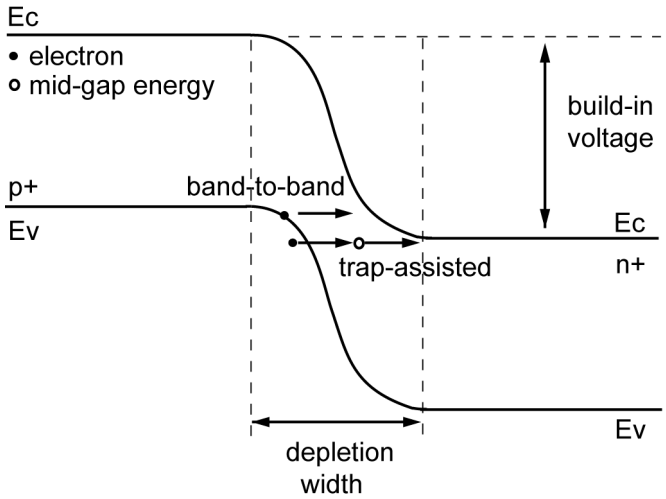


Figure 3-3:Band diagram for tunnelling process of a heavily doped pn junction.

$$J_{tat} \sim W \frac{F_\Gamma}{F_m}\left[\exp\left(\frac{F_m}{F_\Gamma}\right)^2 - \exp\left(\frac{F_m W_0}{F_\Gamma W}\right)^2\right] \qquad (3\text{-}9)$$

where V is the applied voltage, F_0 is a constant which depends on the temperature, W_0 is the depletion width at zero bias, F_Γ is a constant which depends on temperature and the effective mass of carriers, and F_m is the maximal electric field.

The most straightforward way to reduce the dark current caused by tunnelling is to reduce the electric field. In pinned 4T pixel design, tunnelling is very unlikely to occur because the electric field of PPD is rather low, as explained in the previous chapter. In the case of 3T pixels, since the reverse bias is crucial in terms of good photoresponse, dark current and QE is one of the basic standing trade-offs in determining the photodiode bias voltage.

Impact ionization: Compared to the tunnelling process, impact ionization occurs on a much higher electric field. At breakdown voltage, free carriers in the depletion region (either generated by thermal generation or tunnelling) are accelerated by the extremely high electric field. Some of these carriers are able to acquire sufficient energy, i.e. to become a "hot carrier", so that they are energetic enough to break covalent bonds and produce more hole-electron pairs. This process is called impact ionization.

Generally, the bias of the photodiodes in CMOS imagers is not strong enough to induce impact ionization. That means that the dark current contribution from such a process is negligible, at least from the photodiode side. However, later in this chapter, it will be shown that the dark current generation from the TG region of a pinned 4T pixel possibly involves an impact ionization process. Thus, it is worthwhile to also include this mechanism in this study.

In conclusion, both tunnelling and impact ionization processes can only exist inside the depletion region, and they depend strongly on the electric field across the junction. The Shockley-Read-Hall process, i.e. thermal and surface generation, may exist with or without an electric field. However, the presence of the depletion region significantly enhances the emission (generation) process

while suppressing the capture (recombination) process and therefore it increases the dark current generation rate drastically.

3.1.2 Dark Current Generated from Neutral Region

As shown in Figure 3-1, free electrons can also be generated in the neutral region, diffuse to the depletion region, and contribute to the total dark count. Compared to drift processes, this diffusion process is less efficient and depends strongly on the doping concentration. The diffusion length represents the average distance these electrons diffuse to before recombining and can be expressed as [3.2]:

$$L_n = \sqrt{D_n \tau_n} = \sqrt{\frac{kT}{q} \mu_n \tau_n} \tag{3-10}$$

where D_n is the electron diffusion coefficient, μ_n is the electron mobility, and τ_n is its lifetime; both μ_n and τ_n depend on the doping concentration. In fact, the higher the doping profile in the silicon neutral region, the smaller the diffusion length, thus the less significant the diffusion process is [3.2].

Diffusion Current: The minority carrier density at the edges of the depletion region under reverse bias is lower than that of equilibrium. Thus, the system trends to be restored by electrons diffused from the p-type neutral region and holes from the n-type neutral region. These processes are called diffusion currents. Because of the heavy doping concentration in the n-layer, the contribution to the total dark count from holes in the n-type neutral region is much less significant.

The continuity equation in p-type neutral region is given as [3.2]:

$$\frac{d^2 n_p}{dx^2} - \frac{n_p - n_{p0}}{D_n \tau_n} = 0 \tag{3-11}$$

where n_p stands for the electron concentration in the p-type region. In the boundary condition n_{p0} is n_p. Solving this question with a boundary condition of n_p (x=infinite) = n_{p0} and $n(0)$=0, yields the diffusion current:

$$J_{diff} = \frac{qD_n n_{p0}}{L_n} = q\sqrt{\frac{D_n}{\tau_n}} \cdot \frac{n_i^2}{N_A} \tag{3-12}$$

As shown in Eq. (3-12), the diffusion dark current is proportional to the square of the intrinsic concentration n_i. This means that the temperature dependency of the diffusion current is the exponential value of one silicon bandgap. Thus, if the Arrhenius plot is drawn as shown in Figure 3-2 for diffusion current, the activation energy extracted is expected to be close to the value of silicon bandgap. This conclusion is often used to distinguish between thermal generation and diffusion current mechanism.

Another approach to separate the thermal/surface generation and diffusion current can be drawn from Eq. (3-7) and Eq. (3-12). As can be seen, the temperature dependency of the thermal generation current and diffusion current is mainly determined by the intrinsic concentration of silicon (n_i), which is of the first order in Eq. (3-7) and of the second order in Eq. (3-12). By plotting the temperature dependency of the current dark density, it is possible to distinguish which mechanism is dominant at certain temperatures by observing the dependency slope with respect to the slope of n_i (T) [3.3].

Silicon defects certainly also exist in the neutral region. However, as explained previously, the emission through the generation-recombination process is much less significant without the presence of the electric field. Tunnelling and impact ionization process cannot occur either. Therefore, diffusion is the only pronounced dark current generation source in both the n-type and p-type neutral region of the silicon.

3.2 Dark Current Sources in CMOS Image Sensor Pixels

In the previous sub-section, the physical mechanism of various dark current sources were explained. In the section, some actual dark current sources inside different pixel structures will be pointed

out. Also existing or potential approaches on suppressing these dark current components will be introduced briefly.

3.2.1 Total Dark Current in Pixels

The total dark current generated in pixels mainly depends on pixel type, pixel size and fabrication process.

Figure 3-4 shows the change of the dark current from pinned 4T APS pixels among different pixel sizes with shrinking CMOS technologies. The solid line stands for the dark current normalized to the pixel area, i.e. the dark current density. As can be seen, the dark current may reduce with smaller pixel pitch under the same technology node through process optimization. Also, the dark current generation rates are improved with the shrinking of the technology nodes. Since it is the total dark current which is shown in Figure 3-4, the composition of each is un-known. Thus, in order to have a clearer understanding of this improvement, the dark

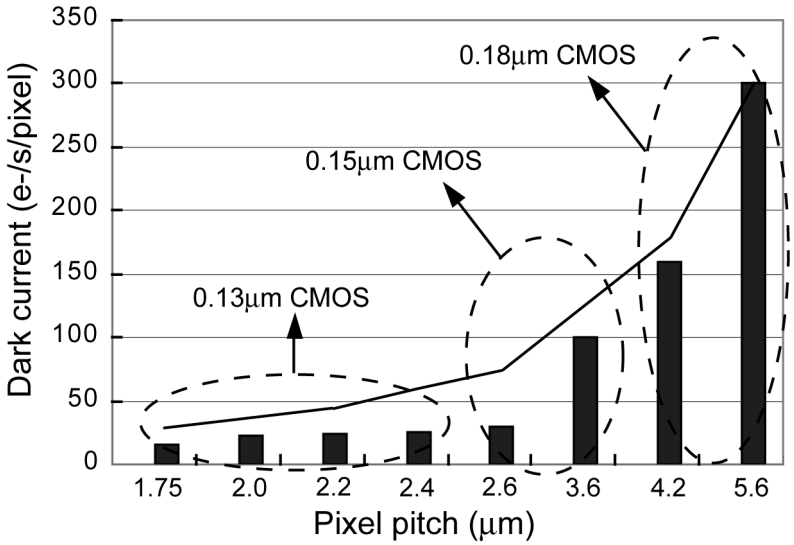


Figure 3-4:Dark current from various pixel sizes using different technologies, redrawn from [3.9].

current generated from different parts of the pixel will be analyzed in detail.

3.2.2 Dark Current from Photodiode

Dark current generated in photodiode of CMOS imagers is usually the most significant component of the total dark count. This sub-session will focus on how it is measured and how its behaviour changes among different pixel structures.

Dark current measurement: In principle, if the dark current is generated in the photodiode, the total dark electron count is proportional to the frame exposure time, which normally equals the integration time of the dark current.

Figure 3-5 shows the dark signal measurement with respect to the exposure time at different operation temperatures. As can be seen, the dark signal is proportional to the exposure time. Thus, it can be concluded that the total dark signal measured is dominated

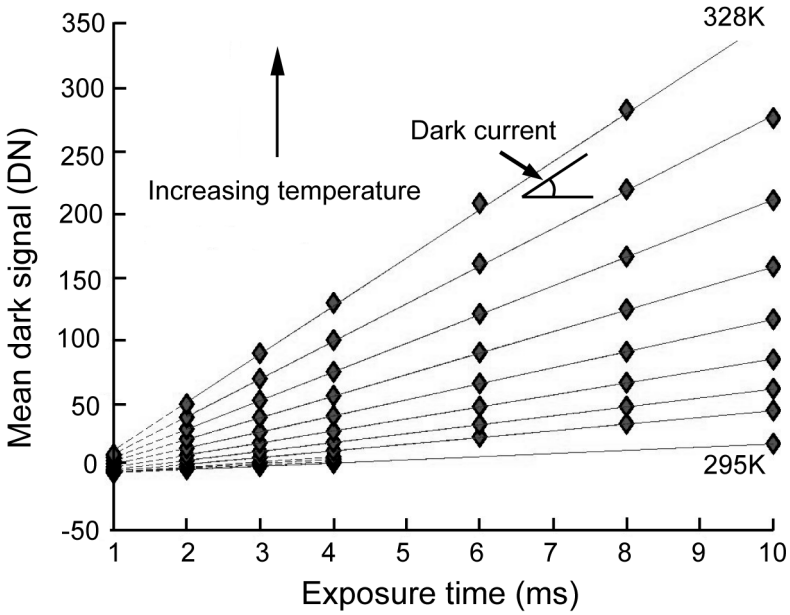


Figure 3-5:Mean dark current versus exposure time (for temperatures rising from 295K to 328K).

by the dark current generated in the photodiode. The generation rate depends on the operation temperature and can be extracted as the slope of the mean dark signal, as shown in Figure 3-5.

Photodiode dark current of 3T APS: In 3T APS pixels, the photodiodes are implemented as ordinary p-n junctions, similar to the one shown in Figure 3-1. The dark current density is mainly determined by the junction type and bias conditions.

Figure 3-6 shows the dark current density measurement with respect to different reverse bias conditions [3.3]. The exponential function of the dark current density from the n+/pwell/psub and p+/nwell/psub junctions are explained by the tunnelling process because of the strong electric field. The test structures are made in a 0.35µm CMOS process and in this particular technology node, the tunnelling process takes over the thermal generation when the reverse bias is above 2V. In other words, a trade-off exists between suppressing tunnelling processes and maintaining pixel photoresponse.

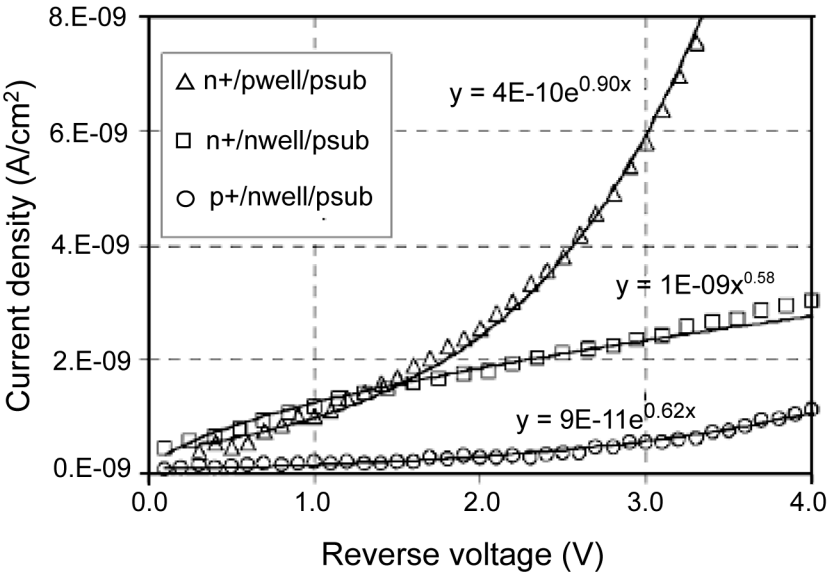


Figure 3-6:Measured reverse I-V characteristics for different junction types with fitted function, taken from [3.3].

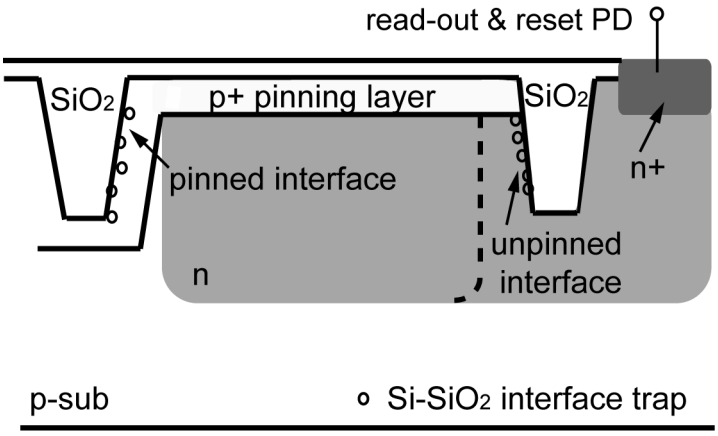


Figure 3-7:Cross section of a pinned photodiode in a 3T APS pixel.

If the tunnelling process can be avoided, the major dark current contribution in a photodiode is the thermal generation in the depletion region: as can be seen from Figure 3-6, the reverse characteristic of the n+/nwell/psub junction follows a square root dependency on the bias voltage. This proves that the thermal generation current is dominant because of the square root dependency of the depletion width 'W' in Eq. (3-3).

Furthermore, it is important to notice that the dark current of the p+/nwell/psub is significantly lower than the others. This is because of the partly pinned Si-SiO_2 interface in this junction type. Figure 3-7 shows the cross-section of such a pinned photodiode in a 3T pixel. However, as can be seen in Figure 3-7, because of the absence of a transfer gate in a 3T pixel structure, the photodiode needs to be reset and read out through an extra non-pinned diffusion region. If the n-region of the photodiode ends, as suggested by the dashed line in Figure 3-7, the photodiode is completely pinned so there is no contact between the depletion region and the Si-SiO_2 interface. However, because a n+ diffusion region is necessary, the n-layer of the photodiode has to be extended. Consequently, it is not possible to completely pin the Si-SiO_2 interface of the photodiode.

Thus, the surface generation dark current contributes to the majority of the total dark count in this type of photodiodes in 3T pixels.

Photodiode dark current of pinned 4T APS: One of the main motivations for designing pinned 4T APS is to reduce the surface generation dark current by a completely-pinned photodiode. Thus, generally speaking, the surface generation dark current component of PPD 4T APS pixels is significantly smaller than that of a 3T APS pixel.

Figure 3-8 shows the cross section of a typical pinned photodiode, which consists of a p+ pinning layer and a n/p-epi junction. As can be seen, the interface defects/energy states are filled by the holes in the p+ pinning layer. Because of the heavy doping of this layer, the depletion layer extension in the pinning layer is very narrow and is separated from the actual interface. Thus, the surface generation dark current is significantly reduced. This approach was first implemented in CMOS Imagers in 1997 by Guidash [3.10].

Although the photodiode seems completely pinned from the Si-SiO_2 interface, it has been discovered that the total dark count depends very much on the distance between the side wall interface of the shallow trench isolation (STI) and the photodiode [3.12]. To decrease the dark current, it is possible to increase the distance between the photodiode and the STI, but at the expense of lowering the pixel fill factor. One commonly accepted approach on this issue

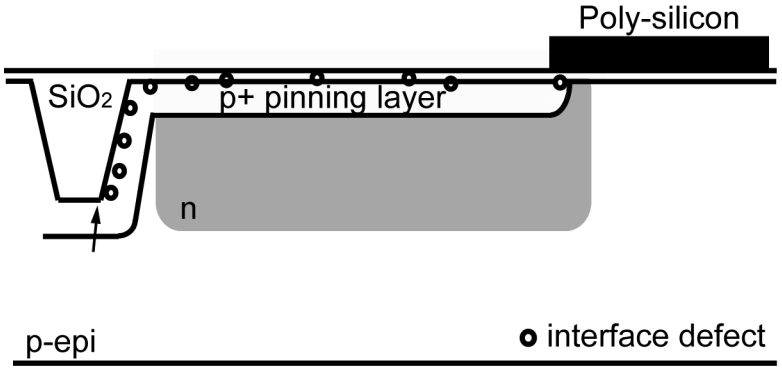


Figure 3-8:Cross section of a pinned 4T APS pixel photodiode.

is to extend the p+ pinning layer to cover the whole STI structure, as shown in Figure 3-8 [3.12]. Thus, the STI interface traps can be filled as well. Figure 3-9 shows the dark signal measurement of pixels with/without p+ pinning layer around the STI isolation. It shows that the dark current generation rate for pixels without a pinning layer around their isolation structures is significantly higher than that of pixels with protected STI.

In conclusion, if both the top and STI side-wall interfaces are well pinned by a heavily doped p-layer, the surface generation dark current can be drastically reduced.

In order to avoid image lag and transfer noise, it is important that the pinned photodiode region is fully depleted before starting integration (or after reset). Thus, both the depth and the doping profile of the n region of the photodiode need to be accurately controlled. This also results in a relatively small full-well capacity compared to the photodiode of a 3T APS pixel. Because of the smaller and shallower depletion region, the thermal generation dark current is also reduced.

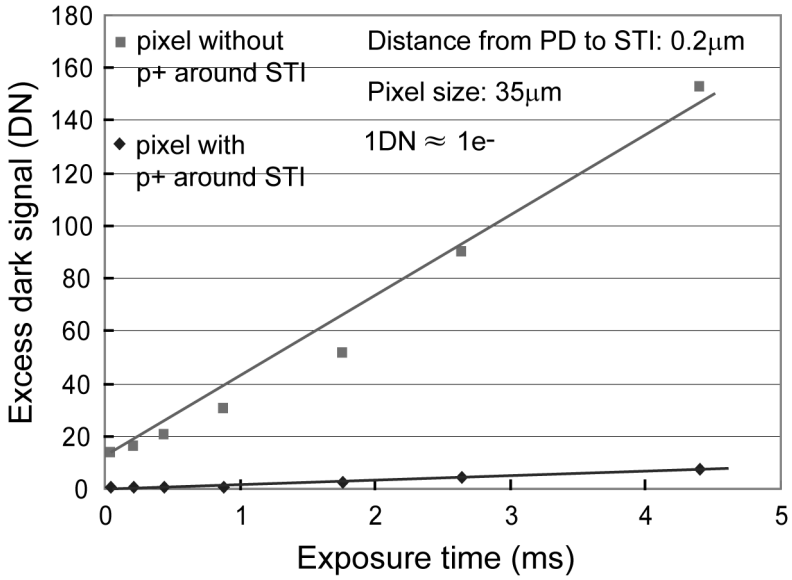


Figure 3-9:Dark signal measurements for pixels with/without p+ pinning layer around the STI.

That is to say, both the surface and thermal generation dark current in a pinned 4T pixel are relatively small. Consequently, the dominant component of the dark current from the photodiode is normally the diffusion current generated from the silicon neutral region. This analysis was confirmed in [3.9] through the activation energy extraction experiment. However, this conclusion does not apply to all pixels in a 4T image sensor. In reality, because of process spread, there are a few pixels that may have extraordinarily high dark current, the so-called hot pixels. The dark current of these hot pixels is normally introduced by higher defects densities inside their photodiode depletion regions, i.e. they are dominated by thermal generation current. Thus, for a 4T APS, the diffusion current is the dominate mechanism for the majority of pixels that have a rather low total dark count. However, for hot pixels, it is the thermal generation that dominates their dark current generation.

From Eq. (3-12) it can be seen that the diffusion current density depends mainly on the doping profile of the silicon neutral region. Increasing the doping concentration helps to reduce the diffusion current. However, it may result in a narrow depletion region, which damages the photoresponse and increases the chance of tunnelling.

Another approach to reduce the diffusion current is to use a hole-based pixel design instead of the normal electron-based photodiode. Because of the lower carrier mobility of holes, the diffusion dark current is expected to be smaller. Also, since pMOS transistors are used to make hole-based pixels, they are built inside an n-well, which in turn reduces the bulk diffusion dark current. Such a method has been implemented very recently in [3.13]. However, hole-based photodiodes normally suffer from a lower photoresponse, especially in a longer wavelength, because of the smaller diffusion length of holes.

3.2.3 Dark Current from Transfer Gate

As previously mentioned, the total dark count of pinned 4T pixel structures is generally smaller than that of 3T designs because of completely pinned photodiodes. In fact, if the 4T pixel pitch is extremely small, e.g. sub 2μm, the dark current contribution from

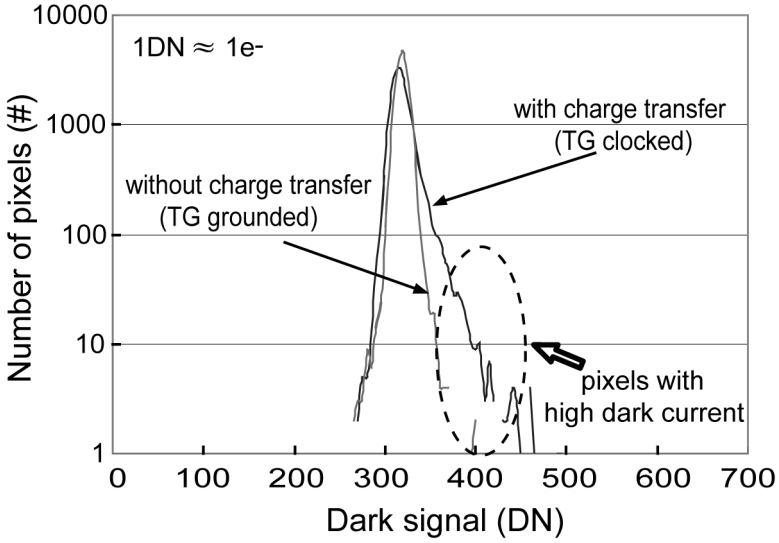


Figure 3-10:Histogram of dark signal measurement with and without charge transfer operation.

the photodiode becomes negligible. However, although the photodiode is optimized in terms of dark current generation, the presence of the TG structure introduces other dark current sources [3.14].

Dark current measurement: In 4T pixel structures, the total dark count can be measured after the integration and charge transfer operation. Thus, it includes the contribution from both the photodiode and the TG. Obviously, the first task in measuring the TG generated dark current is to separate it from the photodiode generation.

Figure 3-10 shows the histogram of the dark signal measurement with and without the charge transfer operation. The purpose of this experiment is to locate the "hot pixels", which have high dark current generation. The test sensor has pinned 4T pixel structures and is made using Philips' 0.18μm CMOS process. The pixel pitch is 3.5μm. As shown in Figure 3-10, the mismatch of the two histograms represents the pixels with relatively high dark current, which may be generated in either the photodiode or the TG.

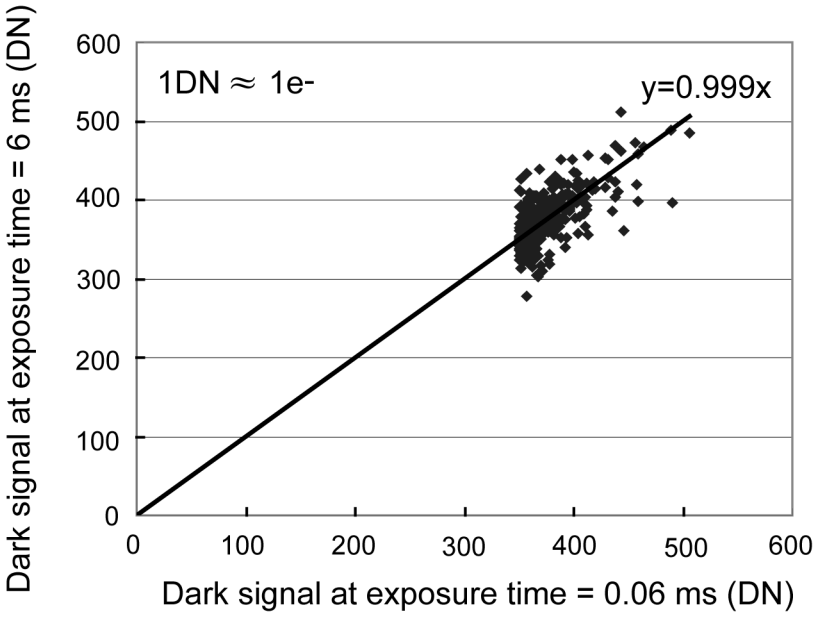


Figure 3-11:Pixel dark signal output at different exposure times.

If the dark current is generated in the photodiode, the dark signal is expected to be proportional to the exposure time. Figure 3-11 shows the dark signals of these "hot pixels" at the exposure times of 0.06ms and 6ms. It can be seen that the dark signals of these pixels are exactly the same under exposure times that differ by a factor of 100, which suggests that the dark current is generated from somewhere else than in the photodiode.

If the dark current is indeed generated in the TG region, it is possible that the total dark count (dark signal) depends on the charge transfer period. Figure 3-12 shows the histogram of the dark signals with different charge transfer times. As can be seen, the total dark current contribution is influenced by the period during which the TG gate is switched on. This observation suggests that the measured dark current is generated and collected during the charge transfer period.

To gain a different perspective Figure 3-13 shows the pixel excess dark signal with different TG-gate switched-on voltages as a function of the charge transfer period. During the charge transfer operation, the FD node potential is fixed at 2V and the TG switched-on voltage is changed from 3V to 4V with a 0.5V interval. As shown in Figure 3-13, it is clear that the pixel excess dark signal increases with a higher TG switched-on voltage.

Furthermore, it is important to notice that the dark current generation rates before t_1 are significantly higher than those after t_1. And after t_1, the dark current generation rates are actually independent of the TG gate bias condition. The mechanism behind this is rather complicated and may be related to various factors, e.g. image lag, interface defects passivation. Therefore, the exact dark current generation process from the TG gate region is still unclear and being study.

In the following discussion, simplified simulation studies will be shown in order to explain the mechanism of this new dark current source.

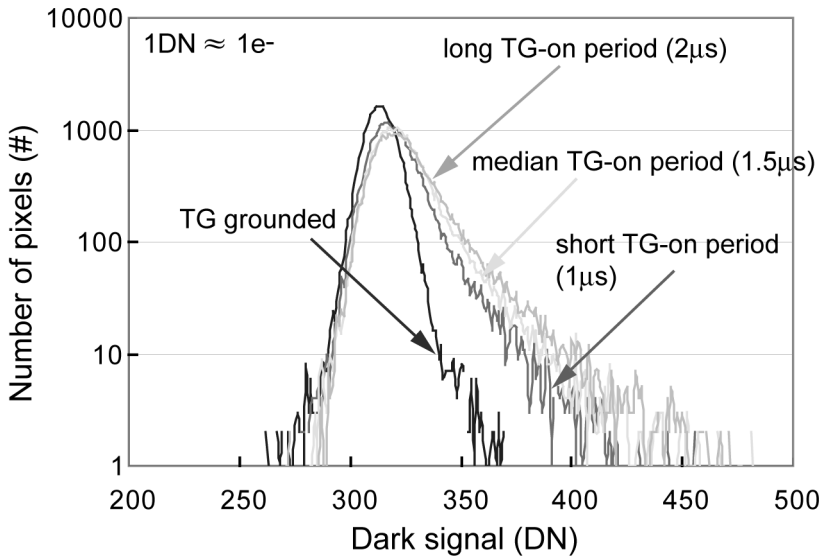


Figure 3-12:Histogram of dark signal measurement with different charge transfer period.

Simulation studies: As explained in sub-section 3.1.1, the dark current generation rate can be extremely high when the depletion region touches the bare Si-SiO_2 interface. In other words, interface defects are the "sources" of the dark electron generation and the electric field acts as an "amplifier". In pinned 4T pixel structures, the photodiode is well shielded from the interface. However, this is not true in the case of the TG region.

Figure 3-14 shows the TG potential diagram during a charge transfer operation simulated by MEDICI [3.11]. The dashed line stands for the depletion region boundary. As can be seen, although the PD depletion region is well pinned, its extension towards TG touches the Si-SiO_2 interface under the TG gate. Even worse, because of the heavily doped pinning layer, the electric field at the overlap between the TG gate and pinning layer is extremely stronge. Figure 3-15 shows the simulated electric field of this overlap region. It is shown that it is determined by the FD voltage and the TG gate voltage. This confirms the dark current generation rate dependency shown in Figure 3-13.

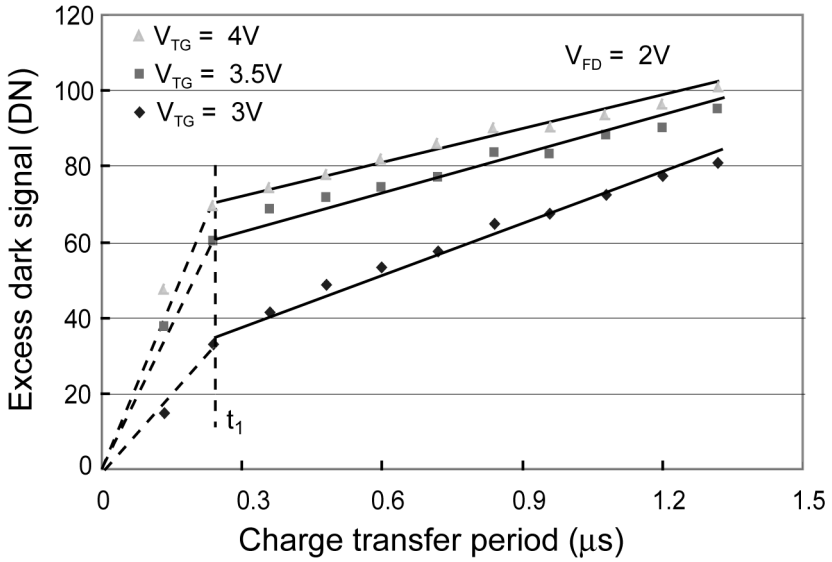


Figure 3-13:Excess dark signal with respect to charge transfer period with different TG switched-on voltages.

In general, the TG induced dark current fits into the category of surface generation mechanisms. However, it is complicated to compare it to the surface generation dark current from the photodiode because of its much stronger electric field. As shown in Figure 3-15, when the FD voltage is as low as 2V, the maximum lateral electric field along the channel reaches $3x10^5$ V/cm or

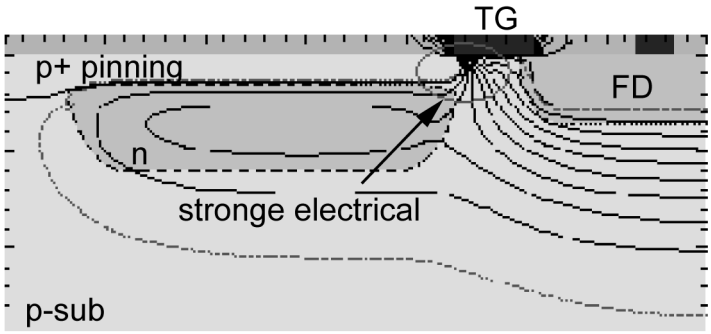


Figure 3-14:Potential diagram of the TG region during charge transfer operation.

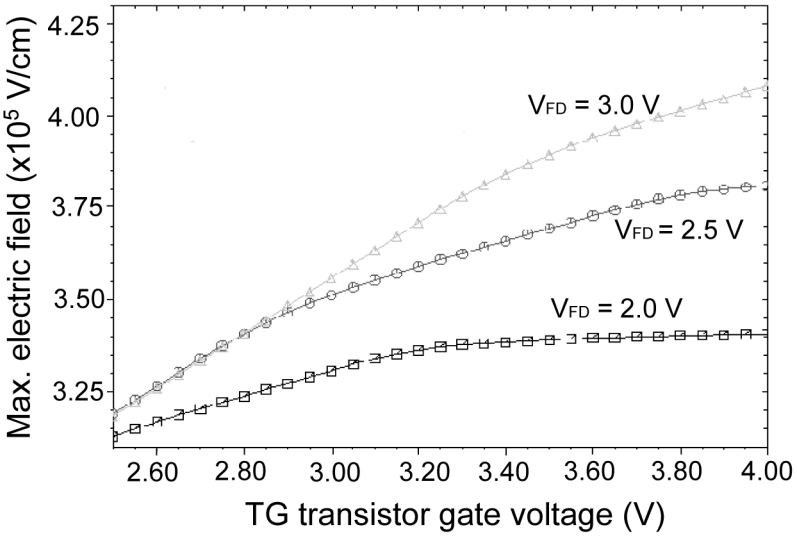


Figure 3-15:The simulated electric field of the overlap between the TG and the photodiode.

higher, which is strong enough to induce hot carrier (HC) effects [3.2]. Unfortunately, the exact influence of this HC effect to the total dark count is not yet fully understood.

Possible solutions: As mentioned, the TG-induced dark current is a type of surface generation dark current. Thus, a straightforward solution is to fill the interface defects with holes, e.g. adding an extra p-type doping layer under the gate of the TG transistor. However, a drawback is that creating such a layer may introduce leakage from the PD to the FD during integration.

Another approach is to use a negative bias on the TG gate during integration in order to attract holes to fill the interface defects. During the charge transfer period, the holes stay in those energy states for a certain period of time before being released. Thus, the trap-induced dark current can be reduced. However, in this case, the charge transfer operation needs to be accomplished before the holes are released. This requires a very short transfer time and may cause incomplete charge transfer, i.e. image lag.

3.2.4 Dark Current from Floating Diffusion

Besides the photodiode and TG region, the dark current can also have a contribution from the FD region. In fact, because of the severe process-induced damages that are caused by the contact etching process and high dose implantation, the dark current generation rate in the FD is rather high [3.12]. However, because the "integration time" of the FD node is generally very short in normal operation mode, its contribution to the total dark count is normally negligible. Nevertheless, for some particular operations, e.g a snapshot using global shutter in which the transferred charge is stored in the FD for a relative long period, this region can be the main source of total dark count.

For the normal operation of 4T pixel, the FD dark current starts integrating right after the first CDS sampling and ends after the second sampling. Since the time interval between these two CDS samplings is in the range of a micro-second, the total dark count is small. However, this may not be true if the read-out scheme is changed, e.g. in a multi-sampling configuration [3.15].

Besides the up-mentioned dark current sources inside the pixel, other sources may contribute to the total dark count, e.g. the excess carriers induced by the HC effect from the pinch-off region of the source follower transistor, or the sub-threshold leakage current from the drain of the reset transistor to the FD [3.16]. However, these dark current sources are either fairly insignificant or can be avoided completely by using particular design or operation principle. Therefore, this will not be explained in any further detail here.

3.3 Conclusions

The discussions presented in the previous sections has led to the following conclusions that are of importance to the design a low dark current pixel:

- From technology point of view, the process technology shrinkage is in favor of reducing the pixel dark current. Densities of both bulk and interface silicon defects are the most crucial parameters in terms of dark current generation rates. To ensure low dark current pixel designs, it is important to optimize the fabrication technology already from the very beginning of the process flow.
- Normally, pixel size shrinking helps to reduce the total dark count.
- The dark current of a hole-based pixel is normally smaller than that of an electron-based design, however this is at the expense of a lower QE, particularly in the long wavelength.
- In general, the dominant dark current mechanism in a 3T pixel is thermal generation (including surface generation) because of the wide depletion region of the photodiode. Thus, the total dark counts depends very much on the silicon surface and bulk defect density.

- In a pinned 4T APS, the dark current of the majority of the pixels which have a rather lower total dark count, is due to diffusion current. However, the bright/hot pixels are normally dominated by thermal generation because of a high defect density inside the photodiode depletion region.
- For both 3T and 4T pixel designs it is important to shield the Si-SiO_2 interface from the depletion region of the photodiode. However, because of the pick-up contact region, the photodiode of a 3T pixel structure cannot be completely shielded.
- In deep sub-micron CMOS processes, the use of STI contributes considerablyto the dark signal. The p+ pinning layer is required to isolate the side-wall interface of the STI from the photodiode. The distance between the STI and the photodiode can be increased in the layout to reduce dark current, but at the expense of reducing the fill factor.
- Significant dark current is generated from the overlap between the transfer gate and p+ pinning layer because of surface generation. Lower FD voltage helps to reduce the total dark count. However, doing so reduces the pixel output swing and may introduce image lag, as explained in chapter 2.
- FD or other types of in-pixel sensing capacitor structures normally result in a relatively high dark current generation rate. To avoid a high total dark count, it is important to ensure a short "integration time" of these dark current components. In other words, the charge stored on these sensing nodes needs to be removed within a very short period of time.

Certainly, the above-mentioned summary cannot cover all the aspects of dark current issues in CMOS imagers. However, it provides a basic guideline and presents trade-offs in terms of designing a low dark current pixel for most of the conventional CMOS imagers.

During the last decade, dark current issues have drawn great attention. Thanks to the well-optimized fabrication technology

nowadays and further optimized pixel structure, the dark current of CMOS imagers has been reduced significantly. Consequently, for most applications which do not require a very long exposure time, the dark current shot noise becomes so small that it can be ignored with respect to the total pixel random noise floor. Therefore, in the next chapters, the topic of random noise in CMOS image sensor pixels will mainly focus on other noise sources instead of the dark current shot noise, e.g. the $1/f$ noise of the source follower transistor inside the pixel.

3.4 Acknowledgement

Both Figure 3-2 and Figure 3-5 are taken from the dark current measurement experiment conducted by Gayathri G. Nampoothiri. The author would like to thank her for her valuable contribution to this chapter.

3.5 Reference

[3.1] W. Shockley et al., "Statistics of Recombination of Holes and Electrons", *Phys. Rev. 87*, pp. 835-842, Issue 5, 1952.

[3.2] S.M. Sze, *Semiconductor Devices. Physics and Technology*. New York, John Wiley & Sons, pp. 109, 2001.

[3.3] N.V. Loukianova et al., "Leakage Current Modeling of Test Structures for Characterization of Dark Current in CMOS Image Sensors", *IEEE Transactions on Electron Devices*, Vol. 50, pp. 77-83, No. 1, Jan. 2003.

[3.4] C. Leyris et al., “Evolution of RTS Source Activities in Saturation Range in N-MOSFETs for Different Oxidation Temperatures”, *18th International Conference on Noise in Physical Systems and 1/f Fluctuations*, pp. 213-216, Salamanca, Spain, Sep. 2005.

[3.5] P. Rao et al., “Gamma-Ray Irradiation Effects on CMOS Image Sensors in Deep Sub-Micron Technology”, *International Image Sensor Workshop*, pp. 70-74, Ogunquit, US, June 2007.

[3.6] A.J.P. Theuwissen, “Influence of Terrestrial Cosmic Rays on the Reliability of CCD Image Sensors - Part 1: Experiments at Room Temperature”, *IEEE Transactions on Electron Devices*, Vol 54, pp. 3260-3266, Dec 2007.

[3.7] J.R. Janesick, “Scientific CCDs”, *Optical Engineering,* Vol. 26, Aug. 1987.

[3.8] P. Rao et al., “Radiation Hardness of Deep-Submicron CMOS Technology Evaluated by Gated-Diode Measurements”, *Sense of Contact*, Zeist, the Netherlands, April 2007.

[3.9] G. Agranov et al., “Optical-Electrical Characteristics of Small, Sub-4μm and Sub-3μm Pixels for Modern CMOS Image Sensors”, *IEEE Workshop on CCDs and Advanced Image Sensors*, pp. 206-209, Karuizawa, Japan, June 2005.

[3.10] R.M. Guidash et al., “A 0.6 μm CMOS Pinned Photodiode Color Imager Technology”, *Technical Digest IEDM*, pp. 927-929, Washington DC, US, Dec 1997.

[3.11] http://www.synopsys.com/products/tcad/taurus_medici_ds.html, Synopsys.

[3.12]H.I. Kwon et al., “The Analysis of Dark Signals in the CMOS APS Imagers From the Characterization of Test Structures”, *IEEE Transactions on Electron Devices*, Vol. 51, pp. 178-184, No. 2, Feb. 2004.

[3.13]E.G. Stevens et al., “Low-Crosstalk and Low Dark Current CIS Technology Using a Hole-Based Detector”, *Technical Digest ISSCC*, pp. 60-61, San Francisco, US, Feb. 2008.

[3.14]X. Wang et al., “Fixed-Pattern Noise Induced by Transmission Gate in Pinned 4T CMOS Image Sensor Pixels”, *Proceedings of the 36th ESSDERC*, pp. 331-334, Montreux, Switzerland, Sep. 2006.

[3.15]S. Kawahito et al., “Column Parallel Signal Processing Techniques for Reducing Thermal and RTS Noises in CMOS Image Sensors”, *International Image Sensor Workshop*, pp. 226-229, Ogunquit, US, June 2007.

[3.16]P. Bain, “Device Noise in CMOS Imagers”, *ISSCC 2007 Forum: Noise in Imaging Systems*, San Francisco, US, Feb.2007.

Chapter 4

Random Telegraph Signal Noise in CMOS Image Sensors

Conventionally, 1/*f* noise is believed to dominate the pixel random noise floor in a conventional pinned photodiode 4T APS if the noise of the periphery circuitry is sufficiently small. However, when the process scales down, instead of the well-known 1/*f* noise, a kind of "Lorentzian noise" is exhibited, i.e. the noise spectra in the frequency domain shows a Lorentzian function. It can also be characterized as random telegraph signal (RTS) noise. The RTS noise is induced by a single interface trap located in the Si-SiO_2 interface of the in-pixel source follower transistor.

In this chapter, the RTS noise of CMOS imagers is characterized. Section 4.1 provides noise measurement results of a pinned photodiode 4T APS, which reveal the existence of RTS noise. This is followed by a theoretical modelling of this noise in section 4.2, which also explains the noise origin. In section 4.3, the RTS noise is further analyzed with varying pixel front-end read-out timings and temperatures. Section 4.4 illustrates how the properties of the interface trap that induces the RTS noise are extracted through experiments. In section 4.5, the RTS noise amplitude is discussed. In the end, the relationship between the RTS and the 1/*f* noise in CMOS imagers are discussed in section 4.6.

4.1 Pixel Random Noise Measurement

4.1.1 Test Sensor Structure and Measurement Setup

The test sensors that are characterized in this chapter are made in 0.18μm CMOS image sensor technology by Philips (Batch number: CD7482 4T7). All pixels are provided with pinned photodiode 4T designs in order to acquire the best noise performance. The pixel pitch is 3.5μm and the source follower transistor size equals W/L = 0.46μm/0.34μm. The whole sensor is divided into six different regions with different design parameters which mainly serve the purpose of characterizing the dark current. In our experiments of random noise measurement, the pixels of a minimum dark current generation rate were selected in order to minimize the dark current shot noise contribution to the total noise floor.

Figure 4-1 is the pixel schematic diagram and its front-end read-out timing during ordinary pixel operation. During the read-out period, the RST is switched off. The RS and the S/HR are enabled

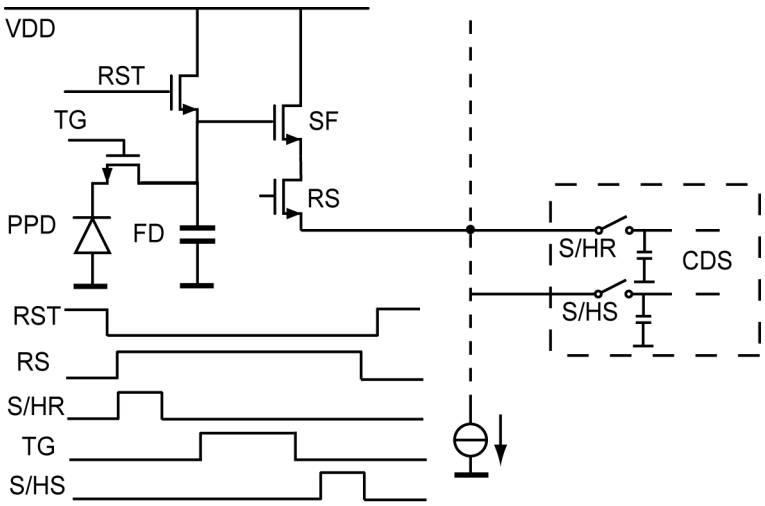


Figure 4-1:Pixel schematic diagram and front-end read-out timing of a pinned PD 4T APS.

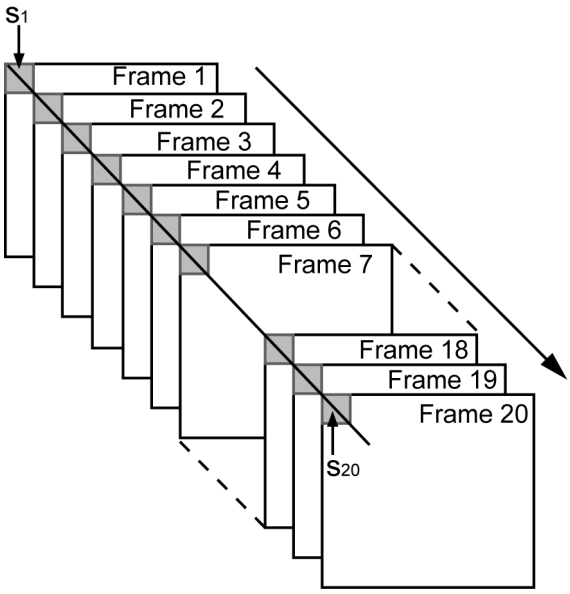


Figure 4-2:Measurement method of the random noise of an individual pixel.

first to sample the reset voltage of the floating diffusion (FD). After the charge transfer operation by TG, RS and S/HS are enabled to sample the video signal. By subtracting these two samples, the transistor offsets and the reset noise are canceled.

The noise measurement is done in complete darkness and during the experiment, the TG transistor is grounded, i.e. it is without the charge transfer operation, thus, the dark current shot noise from the PPD and the TG region are negligible.

The random noise of each pixel is obtained by calculating the standard deviation of its output from 20 frames. Figure 4-2 shows how the noise of one individual pixel is measured. As can be seen, assuming that the output of the pixel$_{1\text{-}1}$ at frame 'i' is S_i, the random noise of this pixel is:

$$\sigma = \sqrt{\frac{1}{19}\sum_{i=1}^{20}(S_i - S_a)^2} \tag{4-1}$$

where S_a is the average value of all the output samples. Because of this, the random noise of each individual pixel is available and the

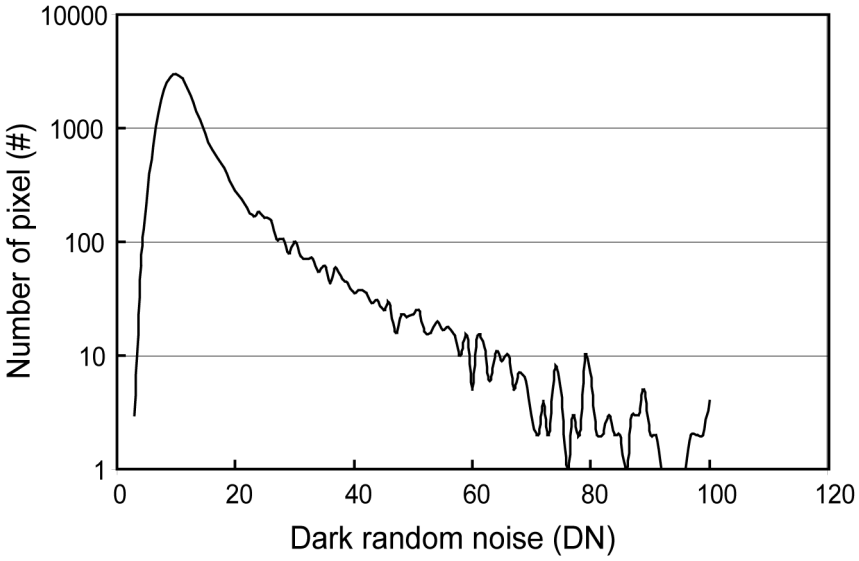


Figure 4-3:Histogram of the dark random noise measurement of a pinned photodiode 4T APS.

read noise floor of the whole imager can be measured precisely by plotting the noise histogram. This method is widely used nowadays and has proven to be very efficient and accurate in measuring the noise level of CMOS imagers [4.1][4.3].

4.1.2 Temporal Output Behavior of Noisy Pixels

As mentioned above, the contributions of the dark current and photon shot noise to the overall measured noise floor are negligible because the reset noise in 4T pixel designs can be canceled by the CDS operation. It can be concluded that the measured random noise is generated by the in-pixel source follower transistor. The detailed procedure of locating the noise sources through noise characterization will be explained in the next chapter with the introduction of a new prototype imager that was made using a different technology.

In modern CMOS imagers, the pixels become smaller and smaller, as does the in-pixel source follower transistor. Because the

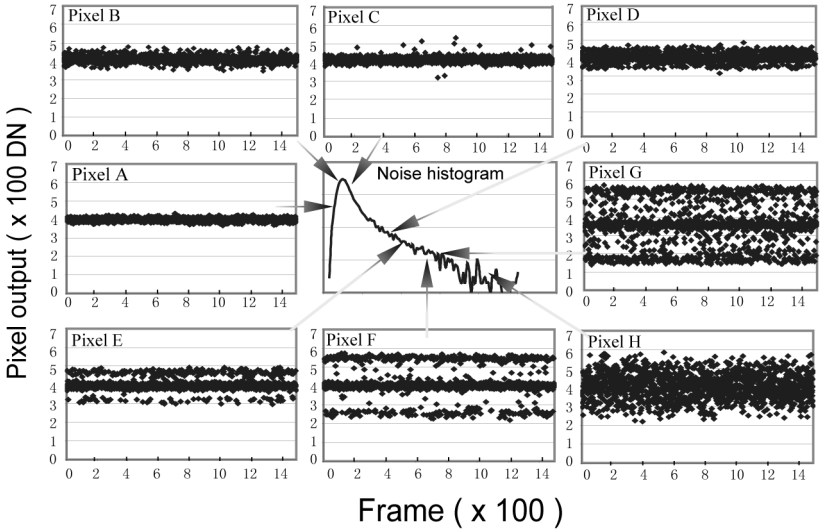


Figure 4-4:Temporal pixel output behavior of pixels belonging to different parts of the noise histogram.

1/*f* noise is reversely proportional to the transistor size [4.2], the pixel random noise caused by 1/*f* noise becomes more and more significant. Figure 4-3 shows the histogram of the dark random noise measurement. The histogram was acquired using the characterization method explained in section 4.1.1. The asymmetric distribution around the peak value is a typical sign of the 1/*f* noise from the source follower transistor [4.3].

With help from advanced computer power and the frame grabber nowadays, it is possible to address any single pixel of the whole array and study its output behavior through thousands of frames. To further investigate the characteristics of the dominating 1/*f* noise, several pixels are selected from the different regions of the histogram and their temporal output behavior is observed through 1,500 consecutive frames.

Figure 4-4 shows the temporal output behavior of these pixels. As can be seen, for most of the “quiet” pixels, e.g. pixel A and B, the output is fairly constant. The variance of their outputs is relatively small. Thus, the random noise of these pixels is not

significant, since it is generally below 20 digital number (DN). However, three discrete levels can already be clearly seen from the temporal output of pixel C. Because the frequencies of the appearance of the highest and lowest levels are small, the overall pixel random noise is small. Therefore, although the actual noise levels for pixel A,B and C are similar, the noise origins for each are clearly different.

Among all the "noisy" pixels, the majority also exhibit three discrete levels, as shown by pixels D, E, F and G. The frequencies and the amplitudes of the upper/lower levels, which determine the actual read noise level of these pixels are all different. Pixel H is one of the noisiest pixels, and the discrete levels are no longer visible.

The pixels which show three discrete levels among their outputs, such as pixel C, D, E, F and G in Figure 4-4, are called "blinking pixels". These blinking pixels are very visible to human eyes in most imaging applications. Moreover, it has recently found that the number of blinking pixels increases significantly when the technology shrinks to deep sub-micron CMOS processes [4.4][4.5][4.6]. Clearly, its physical origin deserves comprehensive studies.

4.1.3 Random Telegraph Signal Noise

This blinking output behavior is not unique in CMOS imagers. Very similar effects have also been observed and reported in analog/RF circuitry [4.7] and memory applications [4.8]. It has been proven that the discrete fluctuation of the conducting current or the threshold voltage are due to the so-called RTS effect of a single transistor.

As explained in chapter 2, $1/f$ noise is believed to be caused by the trapping and de-trapping of the carriers into the defects located in the Si-SiO_2 interface. The number of these defects is related to the total gate area and the technology. Since the gate area of the transistors shrinks when the process scales down, it is possible that there is only one interface defect located in or near the Si-SiO_2 interface if the transistor is extremely small. In such cases, the

trapping mechanism of a single conducting carrier into this trap introduces the RTS effect. In CMOS imagers, if the in-pixel source follower transistor contains only one active trap near the channel interface, it is possible to observe fluctuating pixel outputs.

Figure 4-5 explains why three discrete levels are produced in the pixel output. As can be seen, the pixel output before CDS exhibits two discrete levels caused by the trapping and de-trapping of a single minority carrier. During the CDS operation, the video signal is produced by subtracting S_2 and S_1. The lower value is generated if a falling edge of the RTS occurs between two samples. The higher value is generated in the case of a rising edge. If both samples are in the same RTS status, the CDS output falls into the middle level.

Figure 4-6 is the histogram of an RTS pixel output. The three discrete levels shown in the temporal output behavior are presented as three peaks. By plotting the histogram, the portion of samples

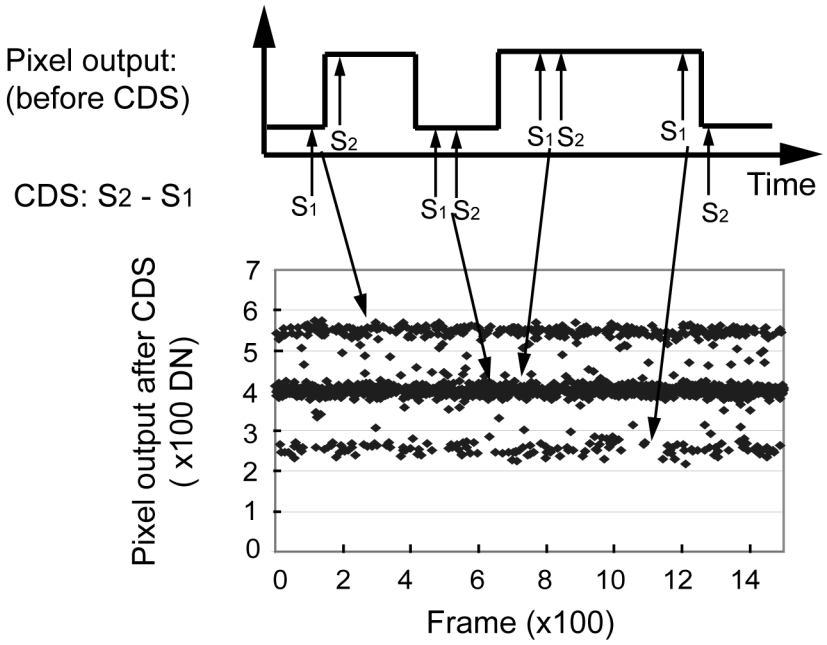


Figure 4-5:Correlated double sampling influence on the fluctuating pixel outputs.

belonging to each discrete levels to the total number of samples is clear. This histogram will be discussed in detail below since it is an important method to characterize the noise.

4.2 RTS Noise Modeling

This section will present a model of RTS noise in CMOS imager pixels that was built to predict its behavior and dependencies. In particular, a new technique to extract the RTS probability parameter through ordinary pixel operation will be introduced. In sub-section 4.2.1, historical studies on RTS noise will briefly be reviewed. In sub-section 4.2.2, the "classic" theory of the RTS noise will be introduced, and the probability parameter of RTS noise will be defined. Next, in sub-section 4.2.3, the existing noise model will be adapted to the pixel operations, i.e. the front-end read-out timing and the method used to extract the RTS probability parameters will be discussed.

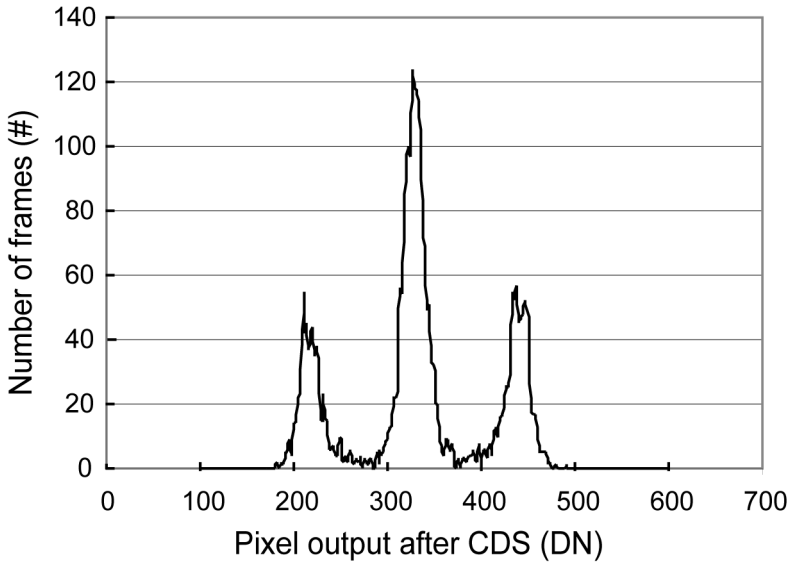


Figure 4-6:Histogram of RTS pixel temporal outputs

4.2.1 RTS Noise in Deep Sub-Micron MOS Transistors

Although the RTS effect in CMOS imagers only appeared very recently, the first detailed analysis of RTS behavior can be traced back to 1989 [4.9]. However, in spite of almost twenty years of research into the RTS noise phenomena in electronic devices, the exact physical origin of this noise is still unclear.

The existence of RTS was first predicted by McWhorter [4.11], who showed that the trapping and detrapping process of minority carriers can lead to a 1/*f* type spectrum. In his model, he predicted that a single Si-SiO_2 interface trap that introduces a Lorentzian spectrum exhibits RTS effect. As illustrated in Figure 4-7, if a transistor contains a large number of traps (assuming that these traps do not interact with one another), the power spectrum densities (PSD), i.e. the Lorentzian spectra of these individual traps, can be added to compose the shown 1/*f* noise spectrum of the transistor.

This prediction was confirmed by Kirton [4.9], who actually measured the Lorentzian spectrum of the RTS. In the work of

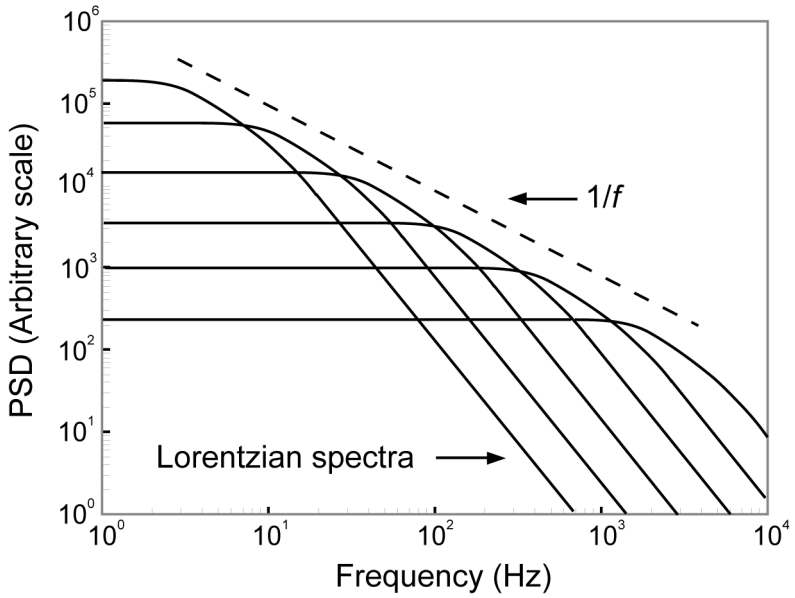


Figure 4-7:PSD of the RTS and 1/*f* spectrum.

Kirton, the RTS behavior in MOSFETs is analyzed as a function of gate voltage and temperature. Kirton's theoretical model is well established enough to extract the RTS trap timing constants. Nevertheless, by late 1980s, when Kirton made his RTS analysis, whether the exact origin of the $1/f$ noise was the so-called ΔN model or the $\Delta\mu$ model, was still being debated, as explained in Chapter 2. Kirton was able to predict the amplitudes of the RTS noise based on the ΔN model. His work showed that the average RTS noise amplitude can be reasonably represented by a simple theory. However, it is somewhat disconcerting to see that a large number of RTS amplitudes appear to be significantly greater or smaller than the predicted value. Unfortunately, he was not able to explain this significant RTS amplitude variance. Even though extensive efforts have been made in recent years to investigate the RTS behavior [4.12][4.14], there is still no convincing evidence that is able to reveal the exact mechanism of this noise and to provide an accurate method to model the RTS noise amplitude.

In the following discussion, the theory proposed by Kirton will be applied to the operation of CMOS imagers. It will be shown that the RTS noise dependency on read-out timing and temperature can be very well explained by this model. The RTS noise amplitude will also be briefly discussed at the end of this chapter.

4.2.2 RTS Noise Model

A trap is a localized energy state in the silicon bandgap whose energy level is normally between the conduction band (E_c) and the valence band (E_v) [4.13]. Depending on the trap energy level, its influence on CMOS imagers can be divided into two categories, as shown in Figure 4-8:

1) If the trap energy is near E_c, e.g. the trap E_{t1} in Figure 4-8, it can interact with the carriers in the conduction band by capturing or releasing an electron, which introduces the RTS behavior. As in the $1/f$ model, both the ΔN and the $\Delta\mu$ effects are involved with the noise amplitude. It has been reported that the $\Delta\mu$ effect often causes a much larger current fluctuation than the ΔN model [4.15].

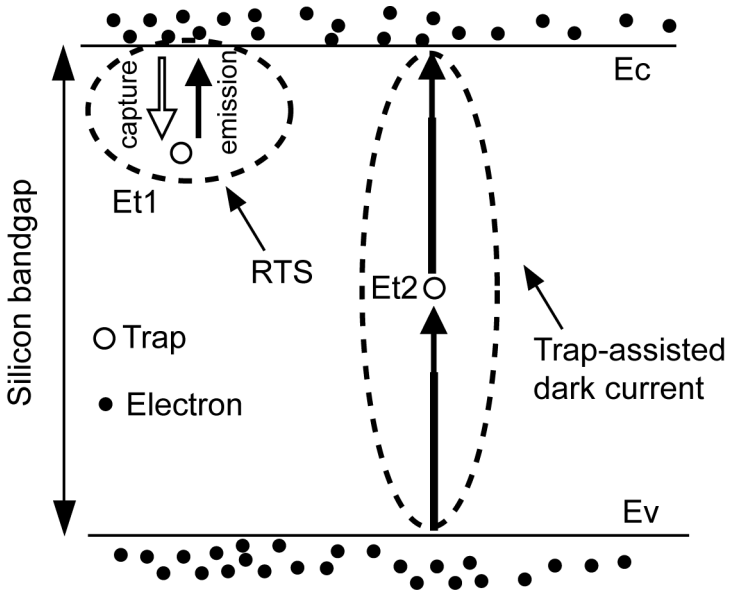


Figure 4-8:Traps in the silicon bandgap and their influences.

2) If a trap is close to the middle of the bandgap, the electrons in the valence band may gain enough energy and thermally stimulated to first the trap energy state and eventually to the conduction band. This mechanism is called trap-assisted dark current, as also explained in chapter 3.

Although the traps inside the silicon bandgap can be divided into two different categories, it is important to realize that these two categories are not completely separate. For example, the trap E_{t2} in Figure 4-8 may also introduce an RTS effect. However, because of the relatively large energy gap between E_{t2} and E_c, the capture (or the emission process) may become slow. The relation between the trap energy and the capture/emission efficiency will be explained below in detail. In CMOS imagers, the observed RTS noise is mainly due to the "fast" traps because of the high CDS frequency applied to the pixel output. Thus, the majority of RTS traps observed from the test imager is expected to be close to the conduction band.

The two most important parameters to describe the RTS trap behavior are the mean time before the capture of an electron (τ_c) and the mean time before the emission of an electron (τ_e), which are both expressed as [4.9]:

$$\tau_c = \frac{\exp(\Delta E_T / kT)}{I_d T \sigma_0 \chi}$$

$$\tau_e = \frac{\exp[(\Delta E_T + \Delta E_{ct}) / kT]}{T^2 \sigma_0 \eta} \qquad (4\text{-}2)$$

where ΔE_T is the trap energy level, ΔE_{ct} is energy difference between the conduction band and trap energy, σ_0 is the capture cross section of the trap, T is the temperature in Kelvin, I_d is the conducting current, k is Boltzmann's constant, η and χ are both process-related constants.

The rates of the capture and emission are therefore given as:

$$\gamma_c(t) = \frac{1 - P(t)}{\tau_c} \qquad \gamma_e(t) = \frac{P(t)}{\tau_e} \qquad (4\text{-}3)$$

In Eq. (4-3), $P(t)$ is the probability of the trap occupancy (PTO), which represents how likely the trap is to be occupied at a given time t. The net rate of capture is therefore given by:

$$\frac{dP(t)}{dt} = \gamma_c(t) - \gamma_e(t) \qquad (4\text{-}4)$$

Eq. (4-4) is considered to be one of the most important equation that characterizes the trap properties because it is valid in both the steady state as well as during a transient. This first order differential equation in P(t) can be solved [4.16] to yield:

$$P(t) = \frac{\tau_e}{\tau_e + \tau_c} + K \cdot e^{-\left(\frac{1}{\tau_e} + \frac{1}{\tau_c}\right)t} \qquad (4\text{-}5)$$

where K is a constant that depends on the initial state. In equilibrium, Eq. (4-5) becomes the well-known form:

$$P = \frac{\tau_e}{\tau_c + \tau_e} \qquad (4\text{-}6)$$

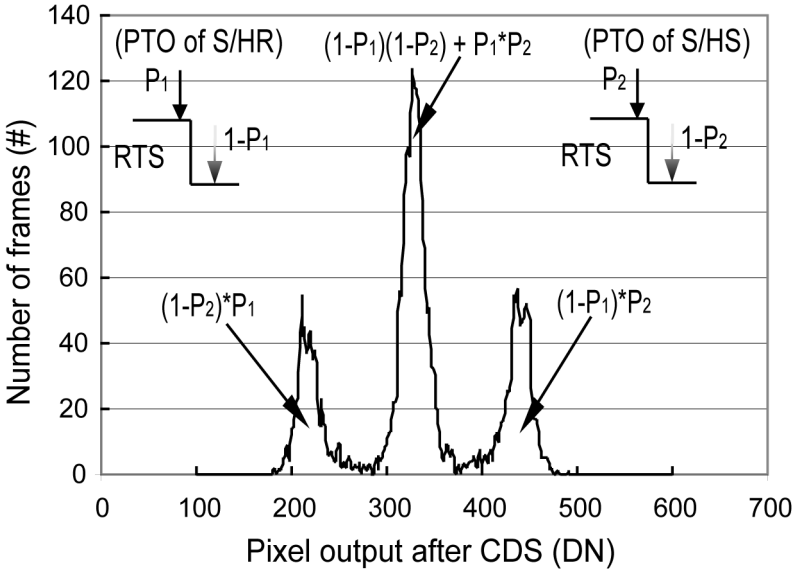


Figure 4-9:Temporal output histogram; P_x is the PTO during samplings.

In this chapter, the transient PTO analysis of RTS noise in CMOS imagers based on Eq. (4-5) and Eq. (4-6) will be used to extract the RTS trap energy using experimental results.

4.2.3 Probability of Trap Occupancy during Pixel Read-Out Operation

In section 4.1, the RTS effect in CMOS imagers was briefly described. Considering the RTS model explained in the previous sub-session, the next question is, of course, how to link this model to the actual RTS measurements in CMOS imagers.

Figure 4-8 is the same histogram of an RTS pixel temporal output behavior that is shown in Figure 4-6 together with the probability analysis of each peak. P_x is the PTO value during the 'x'-th sampling of the CDS operation. It points out how the pixel output level is determined by P_1 during the first sample and P_2 during the second sample. For example, the probability of the pixel output after CDS falling into the left peak is $(1-P_2)*P_1$. $(1-P_2)$ stands

for the probability of the second sample being in the lower level, and P_1 stands for the probability of the first sample being in the higher level. Similar equations can be formulated for the right and the middle peaks as well. Thus, the value of P_1 and P_2 can be extracted by calculating the ratio of the areas of each peak to the total area.

Eq. (4-5) explained how the PTO value can be theoretically calculated. However, several trap property parameters, e.g. the trap cross-section, and capture/emission time, are required in order to obtain an accurate PTO value. Unlike most of the other device property parameters, e.g. the gate oxide thickness, these trap parameters are highly process-related and randomly spread from device to device. In other words, the trap property parameters measured in some test structures cannot be taken as the default for other traps, even though they are all made by the same process. Moreover, these property parameters are difficult to measure directly from existing sensors because the pixel structure is fixed and cannot be modified easily for special trap-property measurement purposes. Therefore, it is essentially impossible to calculate an accurate theoretical PTO value of a specific RTS trap.

Although an accurate PTO value cannot be obtained from theory, its trend and dependency on pixel read-out timing can be analyzed and predicted using the model explained above. Figure 4-10 shows the change of the PTO value during the pixel readout operation. The values of the transient PTOs are shown to be determined by the source follower operation status.

During pixel read-out, the first CDS sampling (S/HR) samples the pixel output right after the reset pulse. During the sampling period, because the RS is switched on, the column current source is connected to the in-pixel source follower transistor. Thus, the source follower operates in the saturation region. The transient PTO of the source follower interface trap increases according to Eq. (4-5).

This increasing of the transient PTO can be qualitatively understood with Eq. (4-2). The trap capture time (τ_c) reduces because of a bigger conducting current I_d, which means that it takes

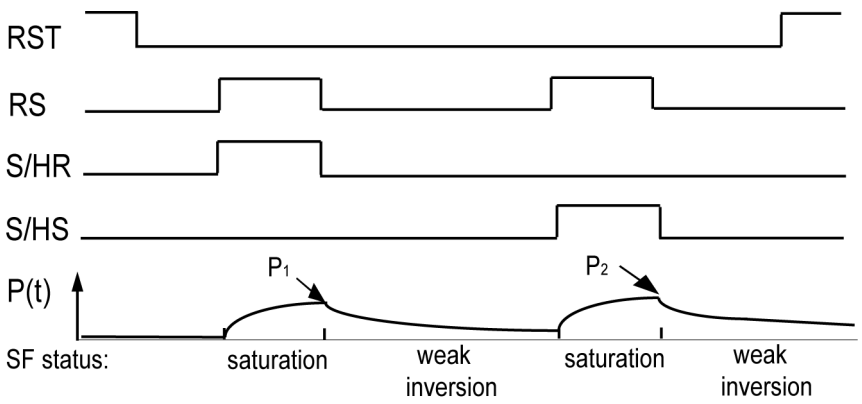


Figure 4-10:The readout timing diagram, the SF status, and the transient PTO (P(t)).

less time for the trap to capture one electron (i.e. it is easier to capture). However, the emission time (τ_e) is independent of the current and remains the same. Therefore, it is more likely that the trap is occupied by an electron at a given time. Thus, the transient PTO increases.

On the other hand, after the CDS sampling period, the RS is switched off and the current source is disconnected from the pixel. Thus, the source follower is switched off (or in the weak inversion region in the initial time). Consequently, the transient PTO value of the trap decreases because of the absence of electrons in the conduction band.

The theoretical prediction of both P_1 and P_2, which can be extracted from Figure 4-9, are also shown in Figure 4-10. They are the PTO value at the end of each sampling period. Linking the experiment and our model, the RTS noise dependency can be well explained even without knowing the actual value of the trap property parameters.

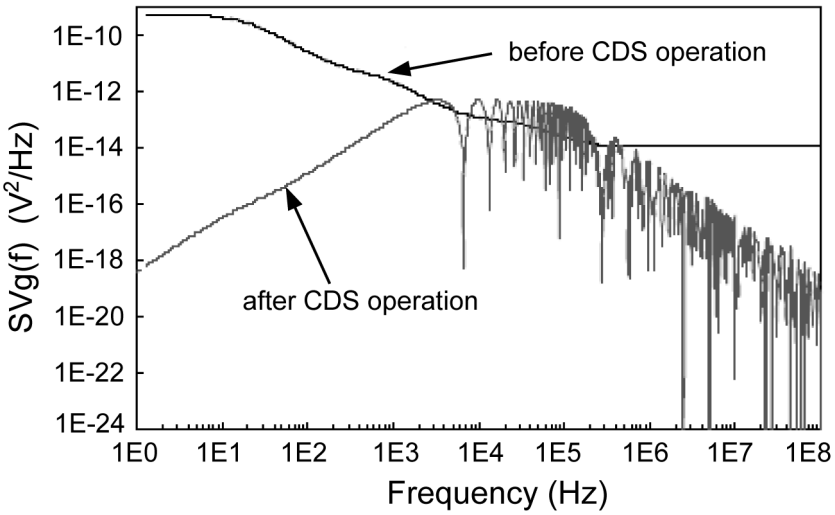


Figure 4-11:Input source follower RTS noise power spectral density before and after CDS operation, taken from [4.17].

4.3 RTS Noise Dependency

4.3.1 RTS Noise Dependency of CDS Operation

Nowadays, CDS architectures are applied on almost all CMOS imagers. As explained previously in chapter 2, the intention of the CDS is to cancel the offset/mismatch of the pixel-level circuitry and the kTC noise of the reset operation.

With respect to noise, the CDS operation acts as a high-pass filter that removes the noise component at low frequencies [4.18]. The response of this noise filter to the RTS noise is determined by the RTS characteristics and the CDS frequency. Figure 4-11 shows an example of a source follower RTS noise PSD before and after CDS operation in the frequency domain. The noise PSD before the CDS operation is a typical Lorentzian spectrum, as shown in Figure 4-7. After the CDS operation, the noise components in the

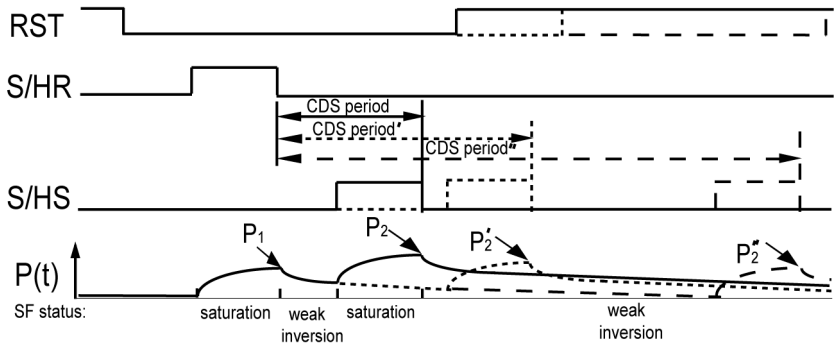


Figure 4-12:The source follower status and the RTS trap PTO values during pixel front-end readout.

low frequencies are reduced. In other words, the pixel RTS noise is suppressed by the CDS operation. A similar influence of the CDS architecture on the $1/f$ noise is also observed in [4.19].

However, the noise filtering, as explained in [4.18], is taken from a 3T APS with a very low CDS frequency. It is based on the assumption that the RTS trap properties do not vary during the CDS operation, i.e. the PSD of the RTS noise is fixed regardless of the CDS operation. Unfortunately, this assumption does not hold for all circumstances, particularly in the case of a 4T APS. As shown in Figure 4-10, if the row-select transistor switches simultaneously with the CDS sampling, the status of the source follower transistor changes, so do the RTS trap properties.

Furthermore, the PSD of the RTS noise shown in Figure 4-11 is based on the trap properties at equilibrium. However, in practice, the CDS frequency is so high that it is indeed difficult for the trap to reach equilibrium within the time interval of two CDS samples. Because of these uncertainties, the influence of the CDS operation on the RTS noise becomes complex and difficult to predict.

Figure 4-12 shows the source follower transistor status and the PTO values of the RTS trap during the pixel read-out periods with respect to different CDS frequencies. The transient PTO is predicted based on the model explained previously. The CDS period stands for the time interval between the first and the second CDS samples.

As shown in Figure 4-12, because of the switching of the row-select transistor, the source follower switches between saturation and weak inversion. Thus, the trend of the transient PTO values changes accordingly. As shown, if the 'on'-period of the CDS sampling pulse remains the same, the PTO value at the end of the first sampling (P_1) remains constant. However, the value of P_2 depends on when the second CDS sampling starts. It is clear that P_2 reduces if the CDS period increases simply because of a smaller PTO starting point at the beginning of the second CDS sampling period.

Consequently, the shorter the CDS period, the bigger the difference is expected to be between P_2 and P_1. If the CDS period is long enough for the RTS trap to reach its equilibrium during weak inversion (i.e. when it is switched off), P_2 will eventually be equal to P_1. Thus, in this case, symmetrical side peaks are expected from the RTS histogram.

Figure 4-13 is the measured RTS pixel output histogram with respect to different CDS periods. It can be seen that the right peak of the histogram does not change while the left peak grows with the increase in the CDS period. With a 160clk-long CDS period, the histogram displays two symmetrical side peaks, as predicted by our model.

Figure 4-14 shows the possibility that the pixel output falls in each peak as a function of the CDS period and the extracted PTO value obtained through experiments. As shown, the left peak of the pixel output histogram grows with the increase in the CDS period, while the right peak remains the same. The extracted value of P_1 is independent of the CDS periods, while P_2 reduces when the CDS period is increased, as predicted by our model.

With regard to noise, both the side peaks stand for the noise portion of the pixel output. Thus, the symmetrical histogram normally yields the highest noise portion, i.e. the pixel-read noise floor appears higher with longer CDS periods. Figure 4-15 shows the measurement of the pixel random noise with different CDS periods. It can be seen that an increase in the CDS periods results in a measured pixel random noise increase as well. If the CDS periods

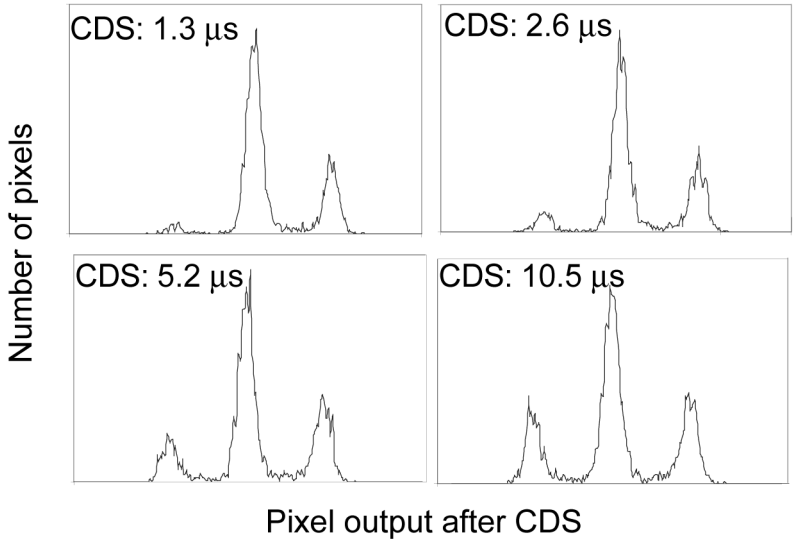


Figure 4-13:RTS pixel temporal output behavior as a function of the CDS period.

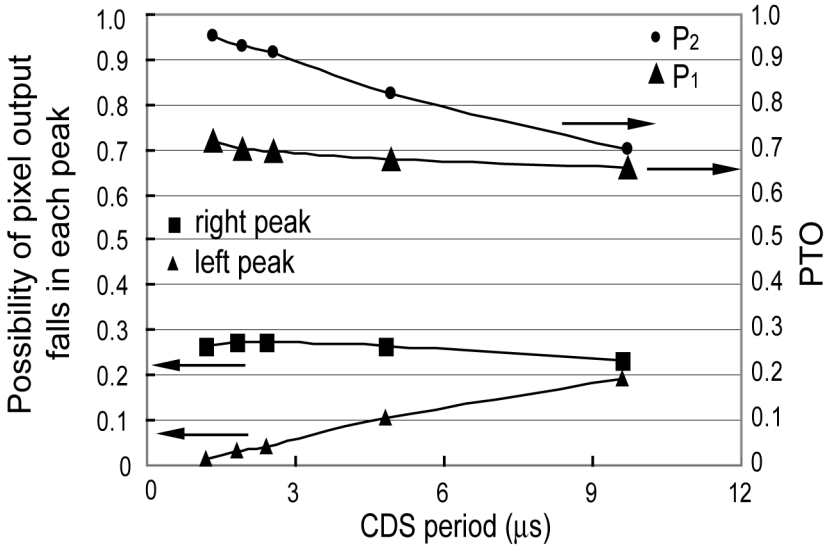


Figure 4-14:Extracted PTO values and the possibility of pixel output falls in each peak as a function of the CDS periods.

is long enough to yield a symmetrically temporal output histogram, the pixel random noise reaches its maximum and stays saturated.

In conclusion, the CDS period has a dramatic influence on the RTS trap time constant, and therefore also on the overall pixel read-out noise. The RTS noise reduces with the reduction of the CDS period, i.e. with faster CDS operation. However, the increase of the CDS frequency can only suppress the RTS noise and not eliminate the noise completely.

4.3.2 RTS Noise Temperature Dependency

CMOS imagers are very complex systems which may contain millions of transistors. Thus, the system temperature may increases during operation because of self-heating. Also, for many applications, e.g. space and medical imaging, there are very strict temperature requirements. Therefore, it is worthwhile to examine the RTS noise temperature dependency.

Figure 4-16 shows the RTS pixel output histograms as a function of the operation temperature. As shown, the noise

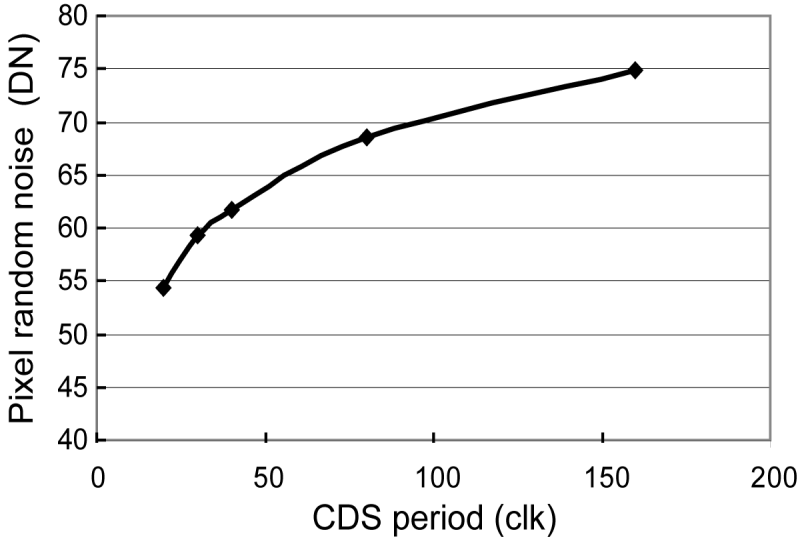


Figure 4-15:The pixel random noise as a function of the different CDS periods.

temperature dependency is rather complex. Both the left and the right peaks start growing from the beginning if the temperature is increased. However, the left peak grows much faster than the right one. At around 34^{o}C the two peaks are equal which yields the highest pixel read-out noise. After that, both peaks shrink if the operation temperature is further increased, and eventually the RTS noise disappears.

The RTS noise dependency shown in Figure 4-16 is quite un-common because, first of all, there is a clear turning point of both the side peaks' ratios, which suggests the existence of a similar turning point of the transient PTO around 34^{o}C. Secondly, unlike

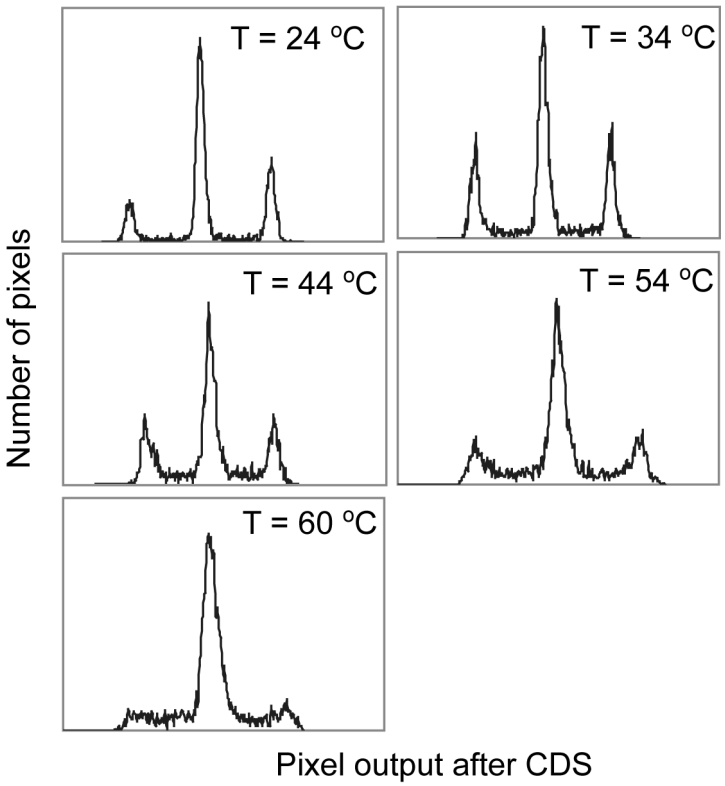


Figure 4-16:RTS pixel output histograms as a function of the operation temperature.

many other noise sources, the RTS noise reduces at higher temperatures (after passing the turning point).

To have a clear view, Figure 4-17 shows the extracted PTO values and the possibility of a pixel output fall in each peak as a function of the operation temperature. As can be seen, both P_1 and P_2 reduce when the temperature starts to increase. After the turning point, both P_1 and P_2 increase, and eventually approach one. This means that at a high operation temperature, the RTS interface trap is filled by an electron most of the time.

The decreases in the PTO values before the turning point can be explained using both classical theory and our model. As shown in Eq. (4-2), both τ_c and τ_e have their temperature dependencies. As the temperature increases, both τ_c and τ_e decrease, but τ_e reduces faster, according to Eq. (4-2). This was also confirmed by Kirton's experiment [4.9]. Thus, according to Eq. (4-6), the PTO in equilibrium reduces as well. Consequently, if the CDS pulse remains the same, the transient PTO values at the end of both samples decrease with increasing operation temperature.

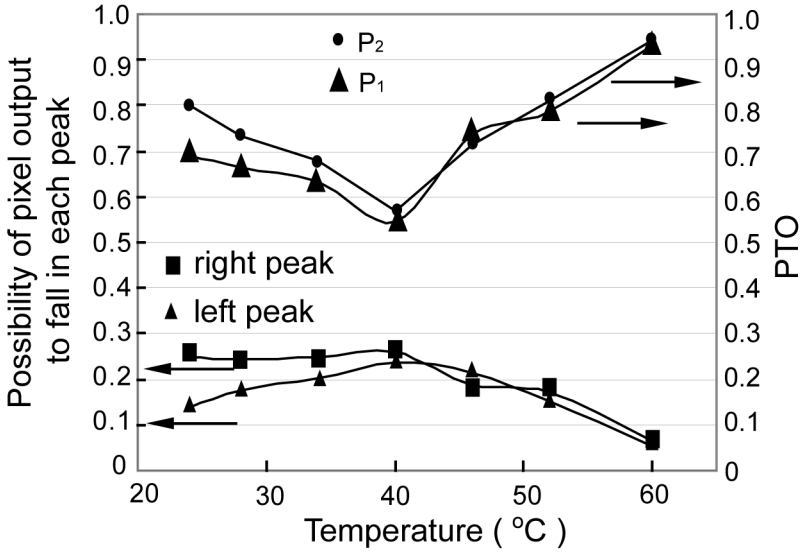


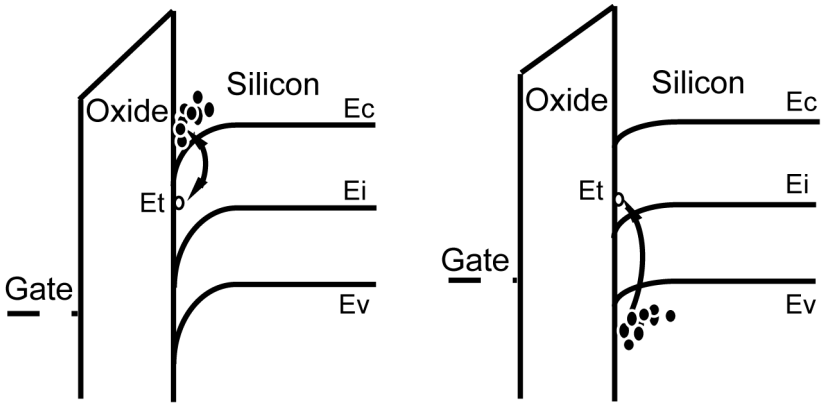
Figure 4-17:Extracted PTO values and the possibility of pixel output falls in each peak as a function of the operation temperature.

The increase in the transient PTO values after the turning temperature is somehow difficult to explain according to Eq. (4-6). However, it is important to note that Eq. (4-6) is based on the RTS behavior measurement when the transistors are biased in strong inversion. Nevertheless, as explained before, the in-pixel source follower transistor in fact switches between two states: weak inversion and saturation, as illustrated in Figure 4-10. In our previous analysis, it was assumed that the trend of transient PTO during weak inversion remains the same when increasing the temperature. However, this assumption may not be true when the temperature continues to increase past the turning point.

Figure 4-18 shows the interface band diagram when the source follower transistor is biased in the saturation and weak inversion. Because of the FD voltage, i.e. the gate voltage of the source follower remains the same during the two CDS pulses in dark, it is reasonable to assume that the trap energy is constant between the two states.

As can be seen, one of main differences in the band diagram between saturation and weak inversion operation is the number of carriers in the conduction band. During saturation, the carriers swap between the trap and the conduction band, introducing the RTS noise. During weak inversion, the trap is more likely to be empty, simply because there are no free carriers in the conduction band. Instead, if the trap energy remains the same, the energy gap between the trap and the valence band reduces, which makes it easier for the carriers in the valence band to jump and occupy the trap energy level. This carrier excitation mechanism is a thermal process. Clearly, the carriers in the valence band are able to gain more energy at a higher operation temperature, which therefore increases the chance that the trap is occupied by an electron and increases the PTO value.

Based on this analysis, our transient PTO model can be re-drawn for temperature dependency, as shown in Figure 4-19. The solid line of the transient PTO stands for the value at a relative low temperature and the dashed line stands for the case of a higher temperature. During saturation, it can seen that the PTO intends to



Saturation Weak inversion

Figure 4-18:Band diagram of the source follower transistor switching between the saturation and the weak inversion region.

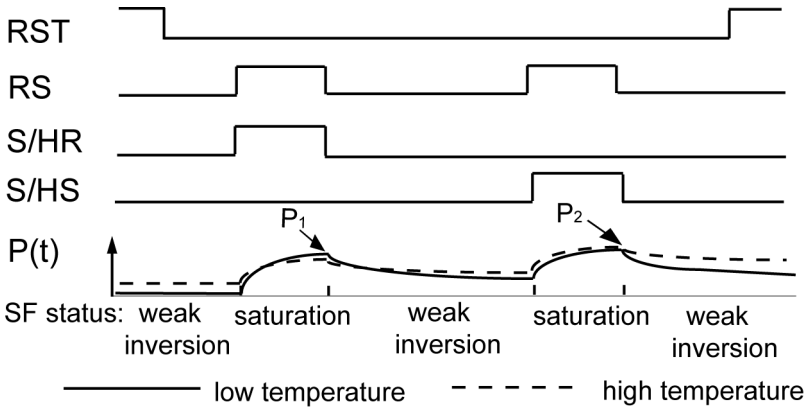


Figure 4-19:Transient PTO value at low and high temperature.

decrease at a higher temperature. On the contrary, the PTO intends to increase during weak inversion at higher temperature.

In conclusion, if the temperature is increased before the turning point, the PTO decreases during the saturation of the transistors. Therefore, the overall transient PTO value decreases. However, after the turning temperature, the thermal excitation mechanism

becomes too strong and the trap is likely to be filled with the electrons excited from the valence band, and thus the overall PTO increases.

This model is able to explain the temperature dependency of the RTS noise. As explained above, the complexity of this noise behaviour is mainly caused by the switching mechanism of the source follower transistor. The main purpose of this switching operation is to reduce the $1/f$ noise [4.20] and to save power consumption. However, it is worthwhile to note that the imagers can also be operated without switching the state of the source follower transistor. The RTS noise under this condition will be discussed into detail later together with an explanation of how the interface trap energy is extracted.

4.3.3 Infrared Light Effect on the RTS Noise

As mentioned previously, the temperature dependency of the RTS noise is rather complex mostly because the carriers exchange mechanism and the thermal excitation mechanism behave differently when the temperature changes. Even though the proposed model is capable of explaining the measurement data, the evidence supporting such a theory is not sufficient. Additional confirmative experiments are needed to verify the existence of these two different mechanisms and their influence on the RTS noise.

Infrared light can be used to verify the existence of defects located in the silicon bandgap. Because the energy of infrared photons is less than that of the silicon bandgap, in principle, pixels made in silicon will not respond to infrared illumination. However, a recent study has found that hot pixels which contain silicon defects do respond to infrared light [4.21].

In principle, if infrared light is used on RTS pixels during the weak inversion operation of the source follower transistor, the infrared photon will help to excite the carriers from the valence band into the trap state, and thus increase the chance of the trap being occupied. On the other hand, the exchange mechanism during

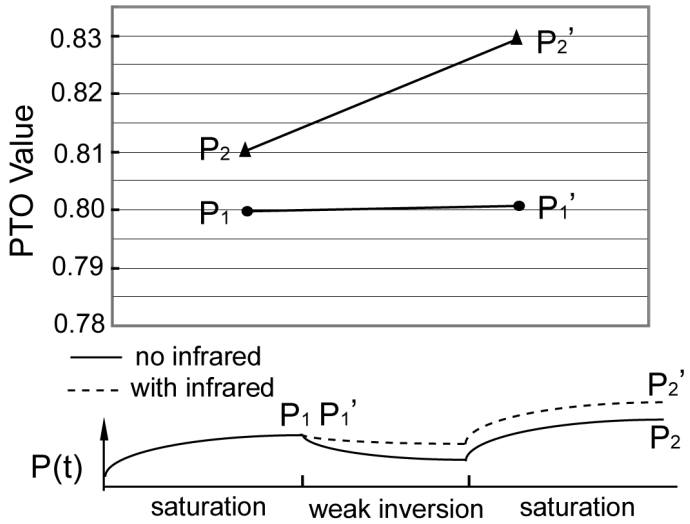


Figure 4-20:Extracted PTO values with and without infrared light illumination and their transient PTO dependency model.

saturation is not affected because the temperature change caused by infrared light is negligible.

Figure 4-20 shows both the extracted PTO values with and without infrared illumination (1330nm) from an RTS pixel and its explanation according to our model. As shown, when the infrared light is used, P_1 remains the same while P_2 increases, which is fully in agreement with our prediction according to the transient PTO model.

4.4 RTS Trap Properties Extraction

As suggested in sub-session 4.2.2, the pixel RTS noise is very much determined by the trap activation energy. For example, as shown in Figure 4-18, if the trap energy is closer to the conduction band level (i.e. it is at high activation energy), the temperature at which the read-out noise reaches its maximum may increase, because more thermal energy is needed for the excitation

mechanism to take over. Clearly, establishing a method to extract the trap energy will be very helpful in understanding and predicting the RTS noise behavior.

Such techniques that extract the activation energy of the gate interface traps from nMOS transistors do exist [4.22]. However, most of them are based on the experiments of a single test structure instead of the transistors inside the pixel. As explained previously, the spread of the RTS trap properties is extreme widely, and so simply knowing the trap energies of individual test structures may not help to predict the RTS noise behaviour of the actual pixels. Furthermore, the biasing conduction of the in-pixel source follower transistors is generally different compared to those test structures being used.

In this sub-session, a new approach to extract the activation energies of the RTS traps based on our model is introduced. All experiments were done during regular operation of an existing CMOS imager. Thus, this method offers a straightforward and convenient approach to evaluate the RTS noise behavior of the pixels directly from an image sensor. Each step of this method will be explained both theoretically and experimentally in the following section.

PTO Extraction in Equilibrium: At the first step, the PTO value of the RTS trap in equilibrium is extracted. The pixel front-end read-out timing explained in Figure 4-10 was therefore modified, as shown in Figure 4-21, of which the row-select continued during two CDS sampling periods. Thus, as shown, the transient PTO does not switch between the CDS time interval. If the CDS period is long enough, eventually, the PTO value reaches its equilibrium condition, which was expressed in Eq. (4-6).

Figure 4-22 shows the extracted PTO value in the experiment according to the method described in Figure 4-21. As can be seen, the value of P_2 stays the same with a continued increase in the CDS period. Therefore, it is safe to assume that the value of P2 is the PTO value of this RTS trap in equilibrium.

Temperature dependency: As shown in Eq. (4-6), the value of the PTO in equilibrium is completely determined by the value of the

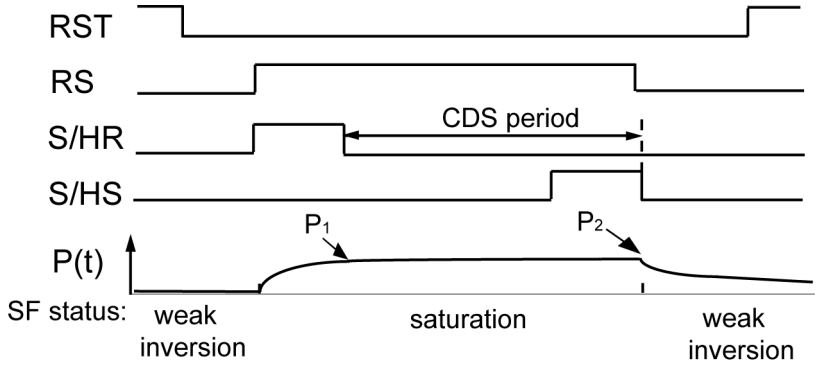


Figure 4-21:Transient PTO model during front-end reading period.

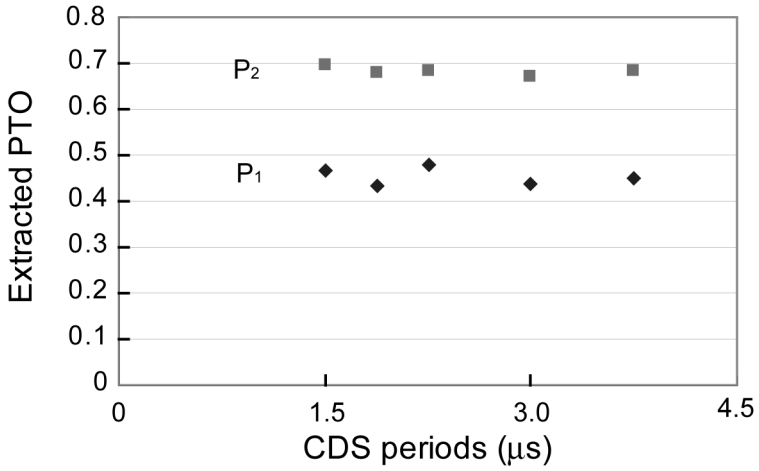


Figure 4-22:Extracted PTO in equilibrium.

trap capture and emission time. In other words, considering P2, we have:

$$\frac{\tau_c}{\tau_e} = \frac{1 - P_2}{P_2} \tag{4-7}$$

The ratio of the trap capture and emission time can also be expressed using detailed balance and is equal to [4.22][4.23]:

$$\frac{\tau_c}{\tau_e} = g \exp(\Delta E_{tf} / kT) \tag{4-8}$$

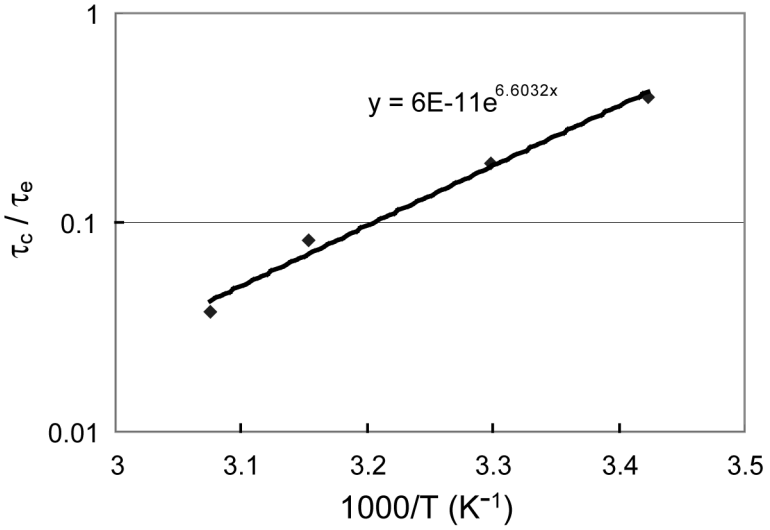


Figure 4-23:Temperature dependency of the ratio of the trap capture and emission time.

where g is the trap degeneracy factor and ΔE_{tf} is the trap energy level relative to the Fermi level. Therefore, through temperature dependence measurements, the ratio of the trap capture and emission time can be extracted and fitted to Eq. (4-8). Figure 4-23 shows the measurement result of the temperature dependency and its curve fitting to Eq. (4-8). As expected, the measurement points fit perfectly with the theoretical line.

Comparing the fitting curve and Eq. (4-8), it is easy to extract the trap energy level relative to the Fermi level, which in this case is about 0.569eV. This means that the trap is located 0.569eV above the Fermi level. Considering a roughly 1.2eV silicon bandgap, the trap is pretty close to the electron conduction band. This also confirms our explanation of Figure 4-8 where the RTS noise is mainly introduced by the interface defects which are very close to the E_c.

The Fermi level energy can easily be calculated or simulated knowing the doping concentration profile and the surface potential. Thus, in principle, the trap activation energy can be calculated as well. However, it is important to note that the surface potential of the source follower transistor is not uniformily distributed along the channel length. Moreover, for devices in a sub-micron process, even the doping concentration varies along the channel [4.12]. Thus, the absolute trap activation energy cannot be extracted without knowing the exact location of this trap (the horizontal distance to the drain or the source not the vertical distance to the interface). Unfortunately, until now there is no convincing method to acquire this information. And these uncertainties of the surface doping concentration, surface potential (therefore, the surface carrier mobility), and the trap location make it very difficult to modelling the RTS noise amplitude in CMOS imagers.

As mentioned at the beginning of this chapter, the timing and the frequency of the RTS noise are very well characterized and understood based on our PTO model. However, the noise amplitude cannot be explained using classical theory. Moreover, the McWhorter $1/f$ noise model will be challenged according to our experiment. In the next sub-sessions, some unsolved issues and conflicting facts related to the RTS and the $1/f$ noise will be discussed.

4.5 RTS Noise Amplitude

Clearly, one of the most complicated puzzles of the RTS noise is its noise amplitude. Before discussing possible reasons for the noise spread, it is important to be aware that the RTS noise is generated from the in-pixel source follower transistor, which is followed by a complex analog processing chain. The whole read-out chain has a significant influence on the noise amplitude as well. Thus, it is wise to first study the whole “noise chain” in order to have a clear view on what the factors are that determine or influence the read-out noise floor.

Starting from the source follower transistor, the conducting current of the transistor fluctuates because of the RTS effect. The fractional change of the current can be expressed as [4.24]:

$$\frac{\Delta I_d}{I_d} = \left(\frac{\eta g_m}{I_d}\right) \cdot \left(\frac{q}{WLC_{ox}}\right) \cdot \left(1 - \frac{x_t}{t_{ox}}\right) \tag{4-9}$$

where η is the constant due to the effect of multiple traps, g_m is the transistor transconductance, I_d is the conducting current, W and L are the transistor width and length, C_{ox} is the gate capacitance, x_t is the distance between the trap and the Si-SiO2 interface and t_{ox} is the gate oxide thickness. Eq. (4-9) is a simplified equation that ignores all the uncertain factors discussed at the end of sub-session 4.4. Based on Eq. (4-9), the corresponding PSD for the RTS noise can be described by the Lorentzian formula:

$$S_{RTS}(f) = \frac{4\Delta I_d^2}{(\tau_c + \tau_e)[(1/\tau_c + 1/\tau_e)^2 + (2\pi f)^2} \tag{4-10}$$

The CDS operation of the pixel output acts as a high-pass filter, thus, the read noise after CDS is given as [4.18]:

$$R(e-) = \frac{1}{S_V A_{SF}(1 - e^{-t_s/\tau_D})}[\int_0^\infty S_{RTS}(f) H_{CDS}(f) df]^2 \tag{4-11}$$

where $R(e-)$ is the rms read noise (e-), S_v is the FD (sensitive node) conversion gain (V/e-), A_{SF} is the source follower voltage gain, t_s is the CDS sample-to-sample period, and τ_D is the dominate time constant of the CDS processor (normally set at 1/2 t_s). $H_{CDS}(f)$ is the CDS transfer function given by:

$$H_{CDS}(f) = (\frac{1}{1 + (2\pi f \tau_D)^2})(2 - 2\cos(2\pi f t_s)) \tag{4-12}$$

As shown by Eq. (4-11), the pixel read-out noise is in fact determined by both the RTS PSD as well as the CDS transfer function. Eq. (4-11) is a well-known equation for modeling the RTS noise. However, this model is found to be inadequate in terms of predicting the noise amplitude [4.24].

Assuming that the CDS operation is fixed, according to this model the RTS noise spread is mainly caused by the spread of ΔI_d. It

is reported that the fractional change of the conducting current may spread from 1% to 70% [4.22][4.24]. And in our experiment, the RTS noise varied from tens of micro volts up to several milli volts, i.e. more than a factor of 100. According to Eq. (4-9), such a wide spread cannot be fully explained.

Secondly, in the model, it is assumed that the PSD of the RTS noise is constant at a fixed temperature. However, such an assumption is only true when the transistor is in equilibrium. As explained before, the RTS noise in a 4T APS is a non-equilibrium process because of the switching of the SF operation. Therefore, the $S_{RTS}(f)$ in Eq. (4-11) cannot be treated separately from the CDS operation.

Because of these limitations, recently a lot of effort have been spent on investigating the nature of the RTS noise in CMOS imagers. It is confirmed that not only the distance of the trap to the interface but also the position of the trap along the channel is very important for estimating the noise amplitude [4.5]. Moreover, the total channel length has a significant influence on the noise amplitude [4.12]. More research is necessary to fully understand this dominant noise.

4.6 RTS Noise and 1/*f* Noise

Although the exact mechanism of the RTS and 1/*f* noise is still being studied, the McWhorter model shown in Figure 4-7 has commonly been accepted. It is believed that the 1/*f* noise spectrum consists of many non-interfering RTS Lorentzian spectra. In other words, the RTS noise is caused by a single interface trap while the 1/*f* noise is caused by multi-traps. Thus, theoretically these two types of noise cannot co-exist. However, this conclusion is indeed being challenged recently [4.10]. Figure 4-24 shows the noise measurement and its reconstruction in the time and frequency domain of an n-type MOSFET [4.10]. As can be seen, Figure 4-24 (a) represents the raw noise measurement in the time domain. It is clear that this noise consists of two different sources. By re-constructing in the time domain, the continuous noise (b) and the

discrete noise (c) are shown. Next, all these noises are converted into the frequency domain. The continuous noise (b) yields a classical 1/*f* slope, while the discrete noise source in the time domain (c) yields a typical RTS Lorentzian spectrum. Since the RTS noise is dominant, the overall raw noise PSD shows as an RTS-like spectrum.

As opposited to the McWhorter model, Figure 4-24 suggests that the 1/*f* noise and the RTS noise are two separate noise sources. As shown by the PSDs of the frequency domain in Figure 4-24, the RTS noise is so significant that it fully covers the 1/*f* noise. This can be explained if there is a superposition of two distributions of traps, one with a distribution of traps leading to a 1/*f* behavior and another outside of the former, e.g. in a position or intensity which results in RTS behaviour.

This result in fact validates the puzzle we had regarding the big RTS noise amplitude in the noise measurements in the previous sub-section. However, this conflicting explanation brings another question to our measurement. As shown in the Figure 4-25, if the

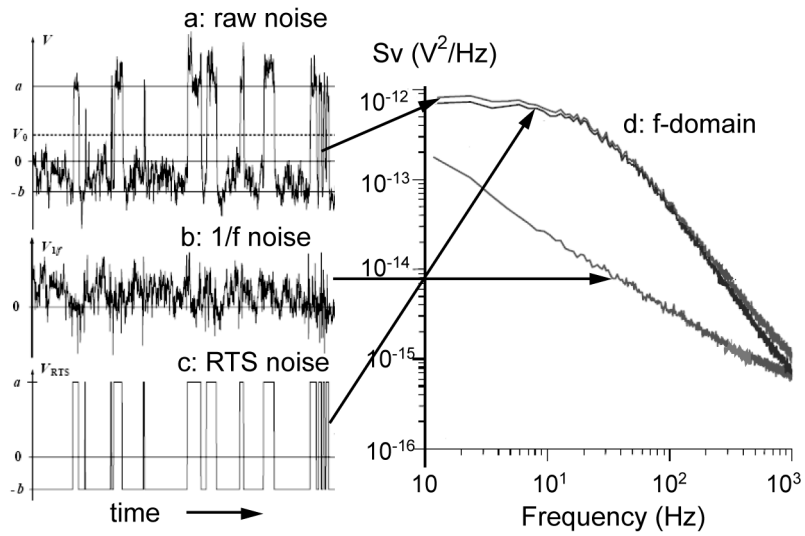


Figure 4-24:Noise measurement and construction of a nMOSFET, taken from [4.10].

side peaks in the noise histogram stand for the RTS noise, what is the cause of the noise distribution within each individual peak?

According to the McWhorter model, the RTS noise is caused by only a single interface defect. In this case, the unknown noise shown in Figure 4-25 can only be there because of the thermal noise. On the other hand, as suggested in Figure 4-24, if the RTS noise and the 1/*f* noise are separate noise sources, the noise distribution within each peak maybe because of the 1/*f* noise.

An easy approach to distinguish these noise sources is to observe their dependencies. Thermal noise increases with higher temperature while the 1/f noise dependency on temperature is rather weak. On the other hand, the read-out 1/*f* noise reduces with higher CDS frequency but the thermal noise is independent of the CDS operation, assuming the system bandwidth stays the same.

For the temperature measurement, as shown in Figure 4-16, the random noise is extracted from the central peak (therefore, neglecting the RTS affect) of an RTS pixel and is calculated at different operation temperatures. The result is shown in

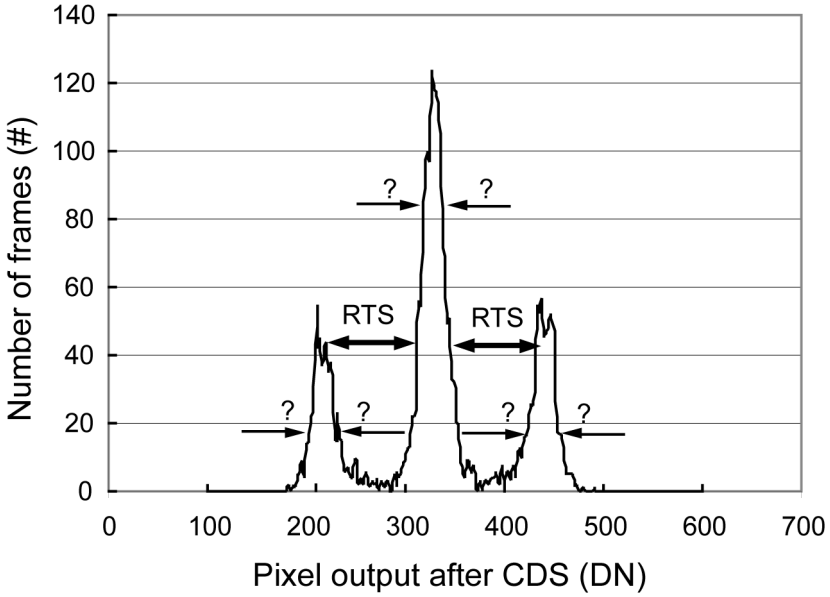


Figure 4-25:Noise histogram of the RTS pixel.

Figure 4-26. As shown, while increasing the temperature from 18 ^{o}C to almost 70 ^{o}C, no obvious increase of the random noise can be observed, which suggests that the thermal noise may not be the cause of the noise distribution.

Similarly, the noise dependency on the CDS frequency as shown in Figure 4-15 is re-calculated, ignoring the RTS side peaks. Figure 4-27 shows the measurement results. Interestingly, as shown, the random noise reduces when the CDS frequency increased, i.e. the CDS period reduces.

Although the McWhorter model is widely accepted and used for modeling the 1/*f* and the RTS noise, our experiment suggests something else. The trap-induced mechanism of the RTS noise stays true however, this RTS trap may not be the only interface defect inside the transistor channel, but it may very likely be the only "special" trap. Its special character may come from its location, distance to the interface, or even from the non-uniform doping of the channel. Certainly, future work is needed to have a clear understanding of the relation between the RTS and the 1/*f* noise.

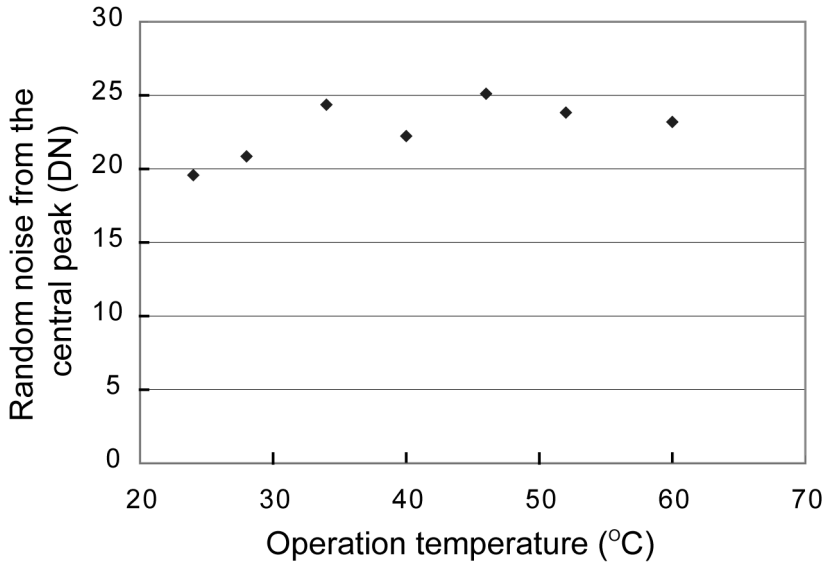


Figure 4-26:Noise temperature dependency from the central peak in the RTS noise histogram.

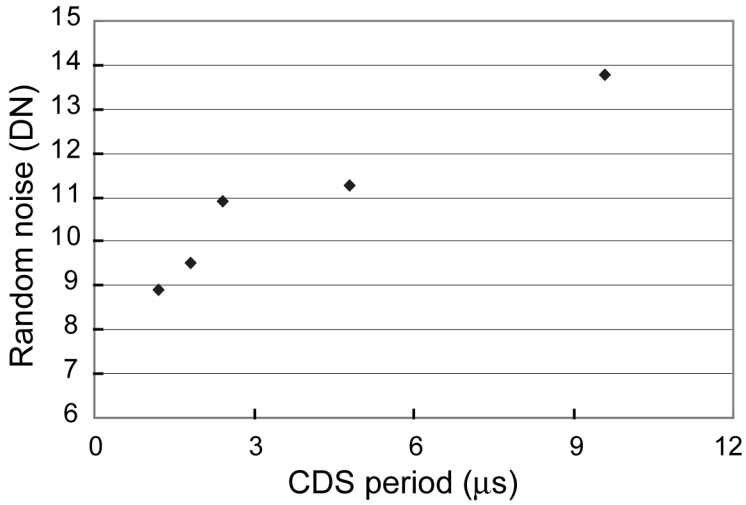


Figure 4-27:Noise dependency on CDS period.

4.7 References

[4.1] B. Pain et al., "Excess Noise and Dark Current Mechanisms in CMOS Imagers", *IEEE Workshop on CCDs and Advanced Image Sensors*, pp. 145-148, June 2005.

[4.2] B. Razavi, *Design of Analog CMOS Integrated Circuit*, McGraw-Hill International Edition, New York, 2001, pp. 215

[4.3] J. Kim et al., "Characterization and Improvement of Random Noise in 1/3.2" UXGA CMOS Image Sensor with 2.8µm Pixel Using 0.13µm-technology", *IEEE Workshop on CCDs and Advanced Image Sensors*, pp. 149-152, June 2005.

[4.4] X. Wang et al., “Random Telegraph Signal in CMOS Image Sensor Pixels”, *International Electron Device Meeting*, pp. 115-119, Dec. 2006.

[4.5] C. Leyris et al., “Modelling of the Temporal Pixel to Pixel Noise of CMOS Image Sensors”, *International Image Sensor Workshop*, pp. 219-223, June 2007.

[4.6] A. Lahav et al., “Optimization of Random Telegraph Noise Non Uniformity in a CMOS Pixel with a Pinned-Photodiode”, *International Image Sensor Workshop*, pp. 230-234, June 2007.

[4.7] R. Brederlow et al., “Fluctuations of the Low Frequency Noise of MOS Transistors and their Modelling in Analog and RTS-Circuits”, *Technical Digest IEDM*, pp. 159-162, Dec. 1999.

[4.8] N. Tega et al., “Anomalously Large Threshold Voltage Fluctuation by Complex Random Telegraph Signal in Floating Gate Flash Memory”, *Technical Digest IEDM*, pp. 491-495, Dec. 2006.

[4.9] M.J. Kirton et al., “Noise in Solid-State Microstructures: A New Perspective on Individual Defects, Interface States and Low-Frequency (1/f) Noise”, *Advances in Physics*, Vol. 35, No. 4, pp. 367-468, 1989.

[4.10] L.K.J. Vandamme, “1/*f* and RTS noise in MOSFETs, the Faster, the Noisier”, *AS-Bruit*, France, June 2004.

[4.11] A.L. McWhorter, *1/f noise and related surface effects in germanium*, PhD Thesis, MIT, Cambridge, MA, 1955.

[4.12] A. Asenov et al., “Random Telegraph Signal Amplitudes in Sub 100nm (decanano) MOSFETs: a 3D ‘Atomistic’ Simulation Study”, *Technical Digest IEDM*, pp. 279-282, Dec. 2000.

[4.13] R.M. Warner et al., *Semiconductor-Device Electronics*, pp. 70. Oxford University Press, UK, 1991,

[4.14]G.I. Wirth et al., “Modelling of Statistical Low-Frequency Noise of Deep-Submicronmeter MOSFETs”, *IEEE Transactions on Electron Devices*, pp. 1576-1588, July 2005.

[4.15]M. Schulz et al., “Single-Electron Trapping at Semiconductor Interfaces”, *Advances in Solid State Physics*, Vol. 35, pp. 229-241, 1996.

[4.16]A.v.d. Wel, *MOSFET LF Noise Under Large Signal Excitation Measurement, Modelling and Application*, PhD thesis, pp. 66, University of Twente, the Netherlands, 2005.

[4.17]C. Leyris et al., “Modelling of the Temporal Pixel to Pixel Noise of CMOS Image Sensors” *International Image Sensor Workshop*, pp. 219-222, June 2007.

[4.18]J. Janesick et al., “Fundamental Performance Difference Between CMOS and CCD Imagers”, *Proceeding of the SPIE*, Vol. 6276, 2006.

[4.19]K. Findlater et al., “SXGA Pinned Photodiode CMOS Image Sensor in 0.35μm Technology”, *Technical Digest ISSCC*, pp. 218-219, Feb. 2003.

[4.20]M.F. Snoeij et al., “The Effect of Switched Biasing on 1/*f* Noise in CMOS Imager Front-Ends”, *IEEE Workshop on CCDs and Advanced Image Sensors*, pp. 68-71, June 2005.

[4.21]M. Loch et al., “Infrared Response of Charge-Coupled Devices”, *Sensors and Camera Systems for Scientific and Industrial Applications VI*, *Proceedings of the SPIE*, Vol. 5677, pp. 201-208, 2005.

[4.22]Y. Shi et al., “Switching Kinetics of Interface States in Deep Submicronmeter SOI n-MOSFETS”, *Semiconductor Science and Technology*, Vol. 16, pp. 21-25, 2001.

[4.23]K.K. Hung et al., “Random Telegraph Noise of Deep-Submicrometer MOSFET’s”, *IEEE Electron Device Letters*, Vol. 11, pp. 90-92, 1990.

[4.24]P. Bain, “Device Noise in CMOS Imagers”, *ISSCC 2007 Forum: Noise in Imaging Systems*, San Francisco, US, 2007.

Chapter 5

Noise Reduction Using In-Pixel Buried-Channel Source Follower

As previously explained, the CMOS image sensor read-out noise floor is dominated by pixel-level noise, specifically the 1/*f* noise and the RTS noise of the in-pixel source follower. This chapter introduces a buried-channel source follower (BSF) that replaces the standard surface-mode nMOS transistor as the in-pixel amplifier. It will be shown that the sensor dark random noise is significantly reduced, for both the 1/*f* and RTS noise components. Moreover, the pixel output swing is increased by almost 100% because of the negative threshold voltage of the BSF.

Both the basic principle of the buried-channel transistor and its applications in CCDs are explained in the introduction, followed by process simulation studies on how to make such devices in a modern sub-micron CMOS process. After that, device simulations that confirm the buried-channel working principle, and the device operation and noise dependency are discussed. Next, the architecture of the prototype sensor is briefly described. In section 5.3, the measurement results of single transistors and test pixels are presented. Finally, noise measurement results of test sensors are shown and the trade-offs of using BSF in CMOS imagers are discussed.

5.1 Introduction

As previously mentioned, the dominating random noise sources in CMOS image sensors are caused by the Si-SiO_2 interface defects of the source follower transistor. However, the exact mechanism of the RTS and the 1/*f* noise are still not completely understood, which makes it very difficult to reduce these noise sources.

In the last several years, a great deal of efforts have been made to reduce the 1/*f* and the RTS noise of CMOS imagers. Research shows that improving the fabrication processes is preferred to using existing circuitry techniques, which are less efficient in reducing these noises [5.1]. One common method from technology point of view is to adjust the annealing process in order to optimize the gate oxide properties [5.2][5.3]. However, such an approach is very process-dependent and needs extensive studies on annealing temperatures and time. Moreover, when processes scale down, the transistors become extremely small, and thus simply reducing the amount of Si-SiO_2 interface traps using annealing optimization may not help with regard to noise. For example, as seen in chapter 4, a single interface defect can already introduce the RTS noise as high as several milli volts. Thus, as long as a perfectly clean gate interface cannot be guaranteed, the reduction of interface-defect-induced noise by improving the annealing process may become less significant.

As previously addressed, although the exact mechanism for these noise sources is not clear, it is commonly agreed on that both 1/*f* and RTS noise are caused by the trapping and de-trapping process of the conducting minority carriers. If the imperfection of the gate oxide has to be accepted, the alternative will have to take the conducting carriers away from the Si-SiO_2 interface. This is the exact motivation of using a buried-channel transistor as the in-pixel source follower. In this section, the basic principle of buried-channel nMOS transistor will be explained and several examples of applying buried-channel devices in CCDs will be introduced.

5.1.1 Working Principle of Buried-Channel nMOST

Literally, buried-channel transistors stand for the transistors that have the majority of their conducting carriers flowing far beneath the gate Si-SiO_2 interface during operation. In modern CMOS processes, the p-type MOS transistors are naturally buried-channel devices because of the threshold adjust doping process during fabrication.

In Figure 5-1, a cross-section of a buried-channel pMOS transistor and its band diagram across the p buried layer is depicted. As shown, the channel surface is fully depleted and the maximum electron potential in the inversion layer is in the bulk silicon.

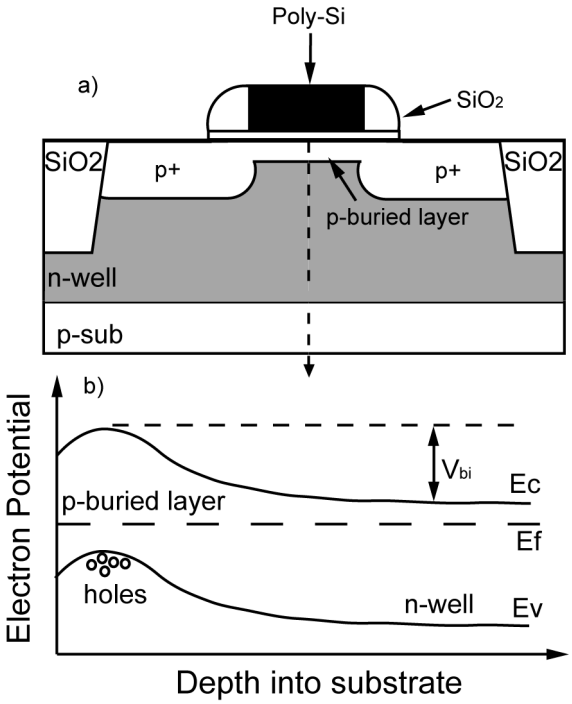


Figure 5-1:(a) Cross-section of a buried-channel pMOS transistor and (b) the electron potential diagram across the p buried layer.

Therefore, the conducting holes flow far beneath the oxide interface. The maximum electron potential along the channel is determined by the channel junction which consists of the p buried layer and the n-well.

In standard CMOS processes, the n-well of a pMOS transistor is heavily doped (in the range of $10^{18}cm^{-3}$) to prevent punchthrough, which makes its threshold voltage (V_t) too negative. The buried doping is to increase the V_t and normally determined as such that it is shifted to the same value as the positive V_t of the nMOS transistor. Thus, creating a buried-channel condition is not the purpose of such a processing step and therefore the channel depth is not optimized or characterized.

Furthermore, using pMOS transistors inside a pixel is essentially very difficult. The co-existence of both pMOS and nMOS transistors inside pixels consumes a great deal of extra space, which lowers the fill-factor of the pixel design. Therefore, in a modern CMOS process, it is worthwhile to create buried-channel nMOS transistors for the in-pixel source follower application.

The structure expected for a buried-channel nMOS transistor is very straightforward, i.e. a total region reverse of Figure 5-1 (a). Figure 5-2 shows the desired operation modes for such a device. It is simulated from an "ideal" CMOS process, which means all parameters and process flow can be adjust freely. The dashed lines stand for the boundaries of the depletion regions. As shown, when the transistor is switched off, the gate interface region is fully depleted and no current flows from the drain to the source. In the linear region, the two depletion regions are separated from each other, which allows current to flow. In the saturation region, the channel pinches off near the drain side. Through out the entire operation, the highest potential is in the silicon bulk. In this simulation, around 250-300nm is from the Si-SiO_2 interface.

Because of the buried-channel doping, the V_t of this nMOS transistor is shifted towards a negative value. This helps to increase the pixel output swing, which will be discussed in detail below.

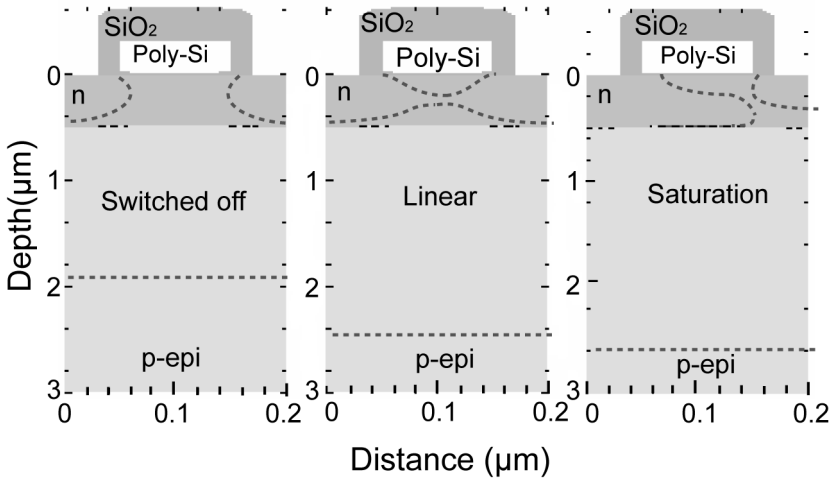


Figure 5-2:Expected operation modes of a buried-channel nMOS transistor.

5.1.2 Buried-Channel Devices in CCDs

Even though the use of buried-channel devices has not been demonstrated in CMOS imagers, this concept has been implemented in the imaging world. The idea of implanting a reverse dose to create a p-n junction which in turn builds a buried potential well is already widely used in CCDs.

The first publication of a buried-channel CCD (BCCD) was in 1973 [5.4]. By creating such a potential well beneath the Si-SiO_2 interface, the photon-generated charges are collected and transferred through the bulk of the silicon substrate. Therefore, the minority carriers cannot interact with the surface defects, so there is less transfer loss and less interface-defect induced dark current.

Figure 5-3 is taken from [5.5], which illustrates the device structure and the potential diagram of a surface-channel CCD (SCCD) and a BCCD. Similar to the potential diagram of

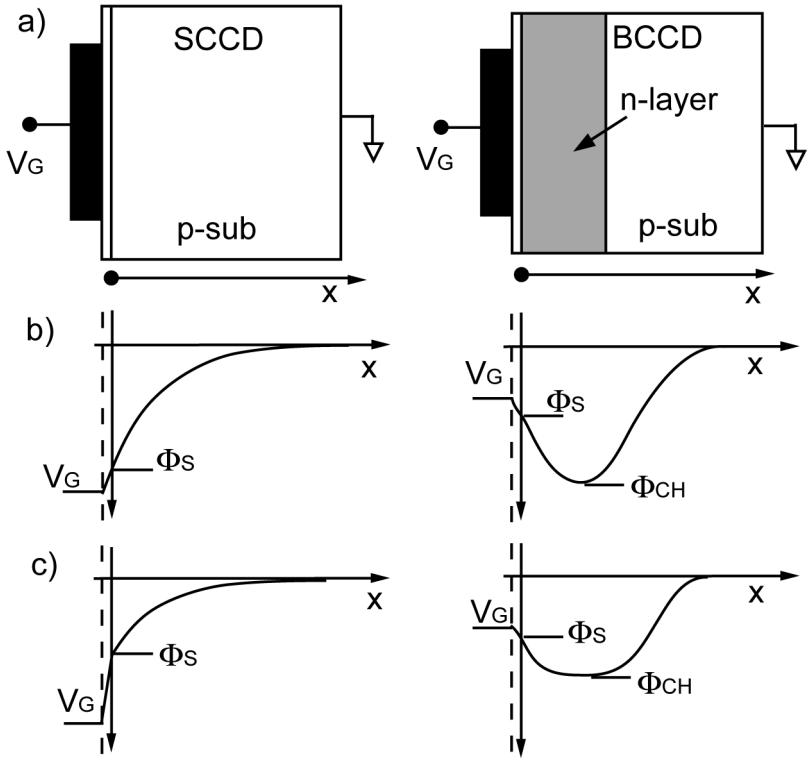


Figure 5-3:(a) Device structure of a SCCD and a BCCD, (b) the channel potential diagrams with empty wells (c) and filled wells.

Figure 5-1, the highest potential (Φ_{CH}) of the BCCD is in the bulk silicon instead of at the interface. The corresponding situation for a SCCD, with empty potential wells, is also shown. Observe that the surface potential Φ_S in the SCCD is lower than the gate voltage V_G, while, in the case of the BCCD, the channel potential Φ_{CH} is higher than the gate voltage V_G.

The situation in which a charge packet is stored in the BCCD and SCCD are schematically illustrated in Figure 5-3(c). For BCCD, the photon-generated electrons accumulate in the bulk potential well and therefore lower Φ_{CH}. As seen in (c), eventually, Φ_{CH} approaches the surface potential Φsand the carriers start interacting with the interface defects, which recreates the

disadvantages of the SCCD. This interaction is believed to be small or negligible if Φ_{CH}- Φ_S is greater than kT/q (-25.8mV at room temperature) [5.5].

Besides the BCCD, buried-channel transistors are also widely used to reduce noise in the output analog chain of CCDs. The output stage of CCDs commonly consists of two or three source followers. Normally, the first source follower is implemented as a buried-channel device to reduce the 1/*f* noise.

Understanding the roles of buried-channel devices in CCDs helps to "migrate" their advantages to CMOS imagers. However, the processing technology is significantly different from the device structure of CCDs and CMOS imagers (e.g, the gate oxide thickness, the channel doping profile, and the device dimensions). It is important to be aware of not only trap-related issues when using buried-channel devices in CMOS imagers, but also the influence on the overall imager performances. Therefore, in the following section, the complete process flow and the device performance will be examined using simulations.

5.2 Simulation Studies

As mentioned previously, the fabrication process of CMOS imagers is different from that of CCDs. Thus, it is important to predict and verify if the buried-channel device can be made in a CMOS process. And if so, if it can achieve the desired performance. In this section, both process and device simulations of buried-channel nMOS transistors will be discussed.

5.2.1 Process Simulations

By itself the fabrication of buried-channel nMOS transistors seems quite simple, with only an additional buried layer implantation compared to the standard process flow. However, the additional implantation needs to be done before the gate deposition. Therefore, it can only be implemented together with the channel/ well implantations, which occur in the very early stage of the whole

process. The annealing processing steps afterwards therefore introduce huge effects and uncertainties to the doping profile of the buried layer as well as to the junction build-in voltage. Thus, a simulation of the complete process flow is necessary.

Figure 5-4 is a simplified process flow chart that describes how the buried-channel nMOS transistors are made. As shown, the buried-channel implantation occurs right after the p-well and channel doping, and afterwards it goes through regular annealing process and the following fabrication steps. The key consideration when making a buried-channel nMOS transistor is to stay with the current technology as much as possible, which includes the fabrication steps, processing temperatures, and time, etc.

The simulation files are supplied by the foundry. The details of the additional buried implantation, e.g. total dose, dopant, and doping energy, need to be specified. The other processing parameters, however, cannot be changed.

Dopant: In current CMOS technology, both phosphorus and arsenic are commonly used as the dopants for n-type implantations. The main difference between these two dopants is their projected ranges, i.e. how far the ion is able to travel into the bulk silicon under the same ion energy. As explained in [5.6], with the same implantation energy, phosphorus has a much larger projected range, and therefore it is more suitable in order to create a deeper junction.

Total Dose: The total dose is important to determine the final doping profile of the buried layer. In general, increasing the total dose helps to bury the channel deeper. However, the transistor threshold voltage also shifts further toward the negative value. Although a negative V_t of the source follower transistor is preferred with the respect to pixel output swing, it is important to note that the enhanced pixel maximum output is limited by the row select transistor. Thus, if the V_t of the source follower transistor is too negative, the pixel cannot operate properly. This effect will be explained in detail in the last section of this chapter.

Implantation energy: As explained, the dopant ion travels deeper into the bulk silicon with the increase of the implantation

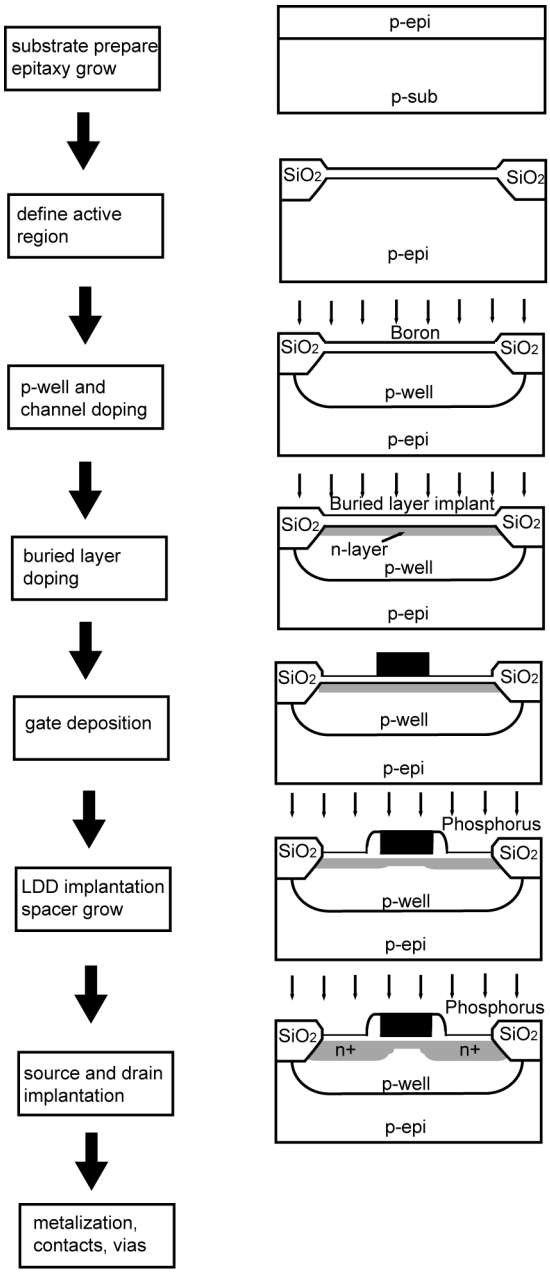


Figure 5-4:Simplified process flow of buried-channel nMOS transistor.

energy. However, as will be proven later with device simulations, the gate regulation of the transistor channel becomes weak with higher implantation energy, which may cause higher leakage current.

These are some general considerations with the implantation parameters selection. In practice, all these parameters are related to each other and cannot be treated in isolation.

Figure 5-5 is a simulation example illustrating how the channel doping profile changes after several key fabrication steps. As shown, the net doping profile is extracted from the middle of the transistor along the vertical distance into the bulk silicon. The simulation is performed with TSUPREM and the original simulation file is supplied by TSMC, based on a 0.18μm CMOS imaging process. In this case, phosphorus is used as the dopant, with a total dose of $8x10^{12}$ atom/cm^2, and the implantation energy is 80keV. Figure 5-5 (a) is the channel doping profile after the p-well implantation, and (b) illustrates how the profile changes right after the buried-channel implantation. It is interesting to note that the doping profile at this moment even exhibits an unexpected n-p-n channel junction. However, after the drive-in procedure, a few annealing steps and oxidations, the final doping profile (e) is completely different and indeed suitable for buried-channel transistor operation.

Even though the simulated channel doping profile can be considered a good indicator of the junction performance, it cannot be treated as accurate and reliable proof of many key parameters, e.g. the junction depth, and the maximum channel potential. Therefore, all the process solutions that are able to create acceptable doping profiles have to be followed by device simulations to verify the operation of the transistor in detail.

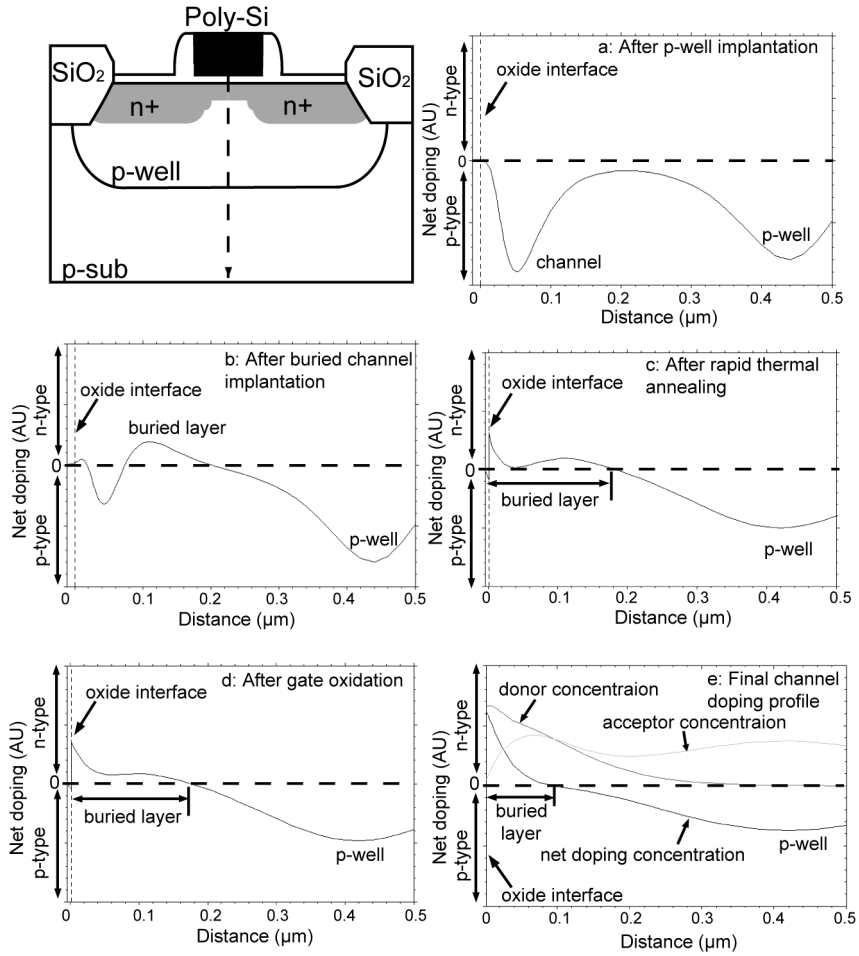


Figure 5-5:Process simulations of the channel doping profiles, after several key fabrication steps.

5.2.2 Device Simulations

The device simulations were conducted using MEDICI. The device structure, the material and doping information were generated by the process simulator TSUPREM.

Expected parameter ranges: Before evaluating each design parameter, it is necessary to define the desired range in order to determine which design is preferred. In the following analysis, three design parameters are discussed, namely the threshold voltage (V_t) of the transistor, the "potential distance" between the channel and interface, and the channel depth.

The working principle of a 4T pixel was explained in chapter 2. In order to increase the pixel output swing, a negative V_t is expected. However, if the V_t is too negative, the pixel output exceeds its maximum value, which is determined by the read-out row select transistor. In the selected process with a 3.3V power supply, the maximum pixel output is around 2.4V. Considering an FD signal swing from 2.6V to 1.4V, the V_t is expected to fit between 0.2V and - 1V. It is important to note that this is only a rough estimate of the suitable V_t values.

As explained in sub-section 5.1.2, to avoid the interaction between the carriers and the interface traps, the "potential distance" between the channel and interface needs to be greater than kT/q (-25.8mV at room temperature). In BCCD, the depth of the buried-channel profile is about 0.8μm [5.5]. Similar channel depth is therefore expected in our buried-channel nMOS transistor.

Threshold voltage: Clearly, the threshold voltage of the transistor is one the most straighforward requirements for our purpose and thus it is used as the "first criteria" for the process selection. Table 5-1 lists the simulated V_t with all the process choices we had.

The extracted V_t is the linear threshold voltage. Table 5.1 shows that if the dopant and implantation energy are kept constant, increasing the total dose lowers the threshold voltage. On the other hand, if the total dose remains the same, even though the junction is supposed to be buried deeper with higher implantation energy, the transistor threshold voltages actually become less negative. This shows that the gate regulation of the transistor threshold, i.e. the channel forming, becomes less efficient when the device converts from the surface-mode into buried-mode. In principle, such devices with high implantation energy are preferred in our application

Number	Dopant	Total Dose (atom/cm^2)	Energy (keV)	Simulated V_T (V)
1	Ph	6.0e12	70	-0.15
2	Ph	6.5e12	70	-0.257
3	Ph	7.0e12	70	-0.38
4	Ph	7.5e12	70	-0.52
5	Ph	8.0e12	70	-0.6656
6	Ph	8.5e12	70	-0.834
7	Ph	7.5e12	80	-0.3811
8	Ph	8.0e12	80	-0.5176
9	Ph	8.5e12	80	-0.6724
10	Ph	8.0e12	90	-0.3696
11	Ph	8.5e12	90	-0.5113
12	Ph	9.0e12	90	-0.6747
13	Ph	1.0e13	90	-1.043
14	Ph	9.0e12	100	-0.5067

Table 5-1.Simulated threshold voltage (V_t) with different implantation parameters.

because of their deeper junctions and mild negative V_t. However, in practice, we observe that these devices suffer from severe leakage.

Clearly, table 5-1 itself does not supply enough information to determine which choice is the most suitable, e.g. choice 4.8.11.14 all give roughly the same V_t. Therefore, further study/simulation is necessary to obtain transistor operation details.

Potential distance: As mentioned previously in sub-section 5.1.2, the interaction between the conducting carriers and the interface defects depends on the "potential distance" between the highest channel potential (Φ_{CH}) and the interface potential (Φ_S). Figure 5-6 plots the simulated device cross-section with source follower bias conditions. The boundary of the depletion region is shown along with the location of Φ_{CH} along the gate length. In this case, the No.2 implantation choice of Table 5-1 is taken, the bias current density used being 12μA/μm. It is important to note that the Φ_{CH} dashed line along the gate is a slanting line instead of a horizontal one. Figure 5-7 shows the middle gap potential curves extracted along X_1 and X_2, as indicated in Figure 5-6. It can be seen that the difference of Φ_{CH} at the various locations along the channel

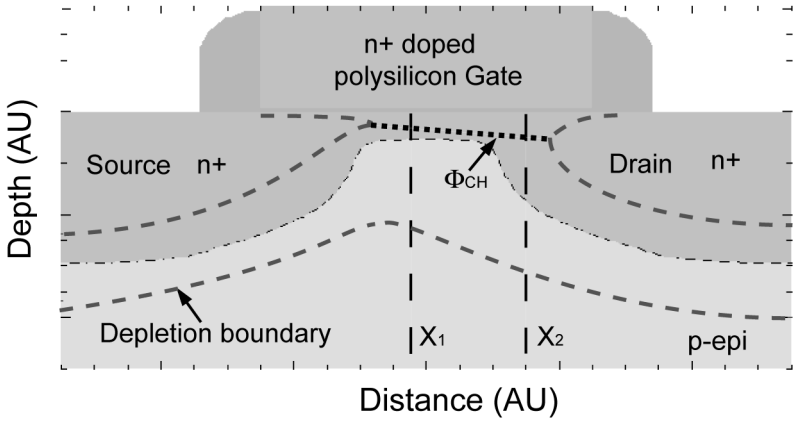


Figure 5-6:Cross-section of simulated BSF under source follower operation bias condition.

is rather significant, which means that the trap-related noise reduction efficiency highly depends on the trap location along the gate length.Although both the "potential distance" at X_1 and X_2 is greater than kT/q (25.8mV), the channel may turn back to surface mode near the source.

Certainly, the "potential distance" also depends on the implantation parameters. This simulation uses the No.2 device from Table 5-1, which is relatively lightly doped with respect to the threshold voltage. It is expected that the "potential distance" can be increased by using a larger total dose. However, from our simulation, such an increase is small. For example, raising the total dose from 6.5×10^{12} atom/cm^2 to 8.5×10^{12} atom/cm^2 with the same 70keV implantation energy, the increase of the "potential distance" along X_1 is less than 10mV (under the same gate bias and conducting current).

Therefore, the "potential distance" study may not help to distinguish which implantation solution is the "best doping". However, it does provide very important evidence that buried-channel devices can be made by means of the current technology and that the "buried" condition can be achieved. Furthermore, in contrast to the BCCDs, whose "potential distances" reduce with the integration of the photo-generated electrons, the

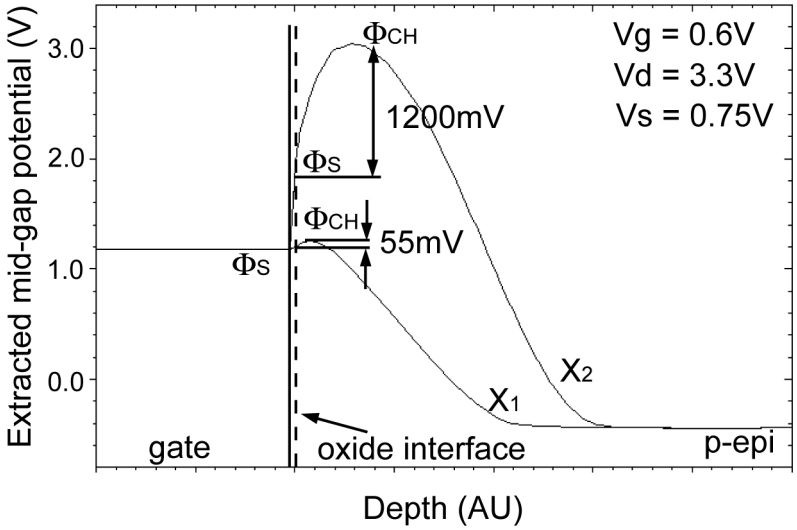


Figure 5-7:Middle-gap potential curve extracted along the different locations of the gate length.

"potential distance" of the buried-channel nMOS transistors is constant.

Channel depth: Besides the "potential distance", the actual physical distance between the channel (the maximum potential) and the interface, (i.e. the channel depth) is studied. Table 5-2 shows the channel depths of different doping solutions as a function of the gate bias. As shown, the channels are buried relatively shallower compared to the BCCDs. Such a channel depth is in fact too shallow to meet our goal. This is mainly due to the extremely thin gate oxide in modern CMOS technology, which is actually a fundamental technology limitation.

Although the channel depth cannot meet our goal, the simulation results of the threshold voltage and potential distance are prove the "buried" concept and the feasibility of creating such devices in the modern CMOS process.

Because the noise model for 1/*f* and RTS is not well established on the device level, the noise simulation can be extremely inaccurate and was therefore not included in the simulation studies.

After the simulation studies, five implantation solutions (No.2 4 6 8 14 from Table 5-1) were chosen mainly based on threshold voltage criteria. The test device and the prototype imager were made in a 0.18μm 1P3M CMOS process. In the following sections, the measurement results will be discussed.

5.3 Test Transistor & Pixel Characterization

5.3.1 Single Transistor Characterization

Test transistors were fabricated and measured. Figure 5-8 is the simulated and measured gate characteristics of the surface-channel and buried-channel transistors. These buried-channel devices were made with different total doses but with the same implantation energy. The transistor size is the same.

It can be seen that the simulated and measured I-V curves of the standard surface-mode device match very well. For the buried-channel devices, increasing the implantation dose lowers the V_t of the transistors. The V_t extracted from the measurement is slightly higher than the simulated one. This difference becomes greater for a higher implantation dose.

As shown in Figure 5-8, under the same gate bias condition, the slope of the I-V curve, i.e. the transconductance (g_m), increases with the increasing of the dose. For the source follower application in CMOS imagers, the transistors operate under a very small biasing

Implantation	Channel depth(nm) (Vgate = 2.6 V)	Channel depth(nm) (Vgate = 1.6 V)	Channel depth(nm) (Vgate = 0.6 V)
70/7.5 x 10^{12} (ph)	20	30	40
80/7.5 x 10^{12} (ph)	20	30	40
90/8.5 x 10^{12} (ph)	25	40	55
100/9.0 x 10^{12} (ph)	25	40	60

Table 5-2.Simulated channel depth of different dopings as a function of the gate biasing.

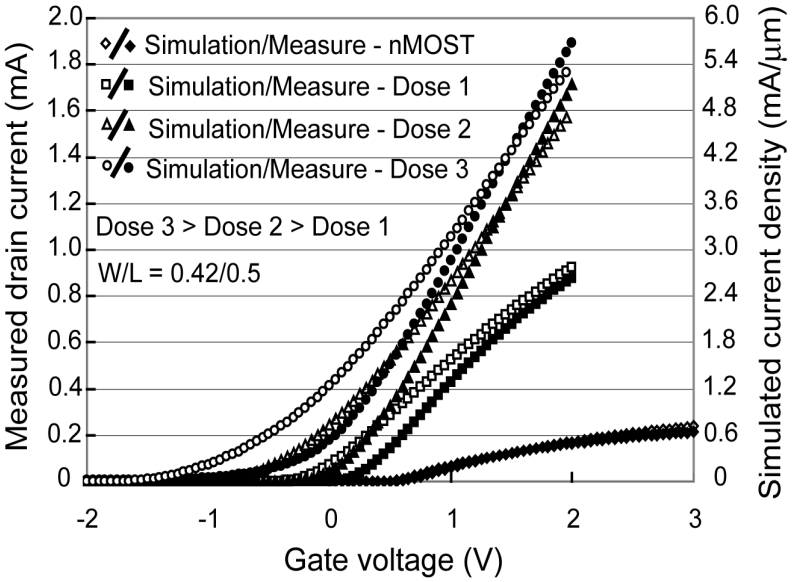


Figure 5-8:Simulated/measured gate characteristics of transistors with different implantations.

current in order to maintain a weak-inversion operation, it is therefore important to compare the g_m at the expected operation points. In Figure 5-9, the measured g_m of the surface-mode and the buried-channel transistors is plotted as a function of the bias current. The experiment was conducted using a fixed source voltage in order to include the body effect. The gate voltage is swept to acquire all the DC points. As shown, with the same bias current, the g_m of the buried-channel device is almost half that of the surface-mode device, which may cause a longer settling time.

Besides the standard I-V measurement, the 1/*f* noise of individual transistors has been characterized and analyzed. Figure 5-10 shows an example of the DUT current-referred noise measurement results for the surface-channel and buried-channel transistor. It is clear that the noise component at low frequencies is significantly lower for a buried-channel device than for a surface-channel device. However, the slope of the buried-channel device PSD is generally smaller than 1. Therefore, in some cases,

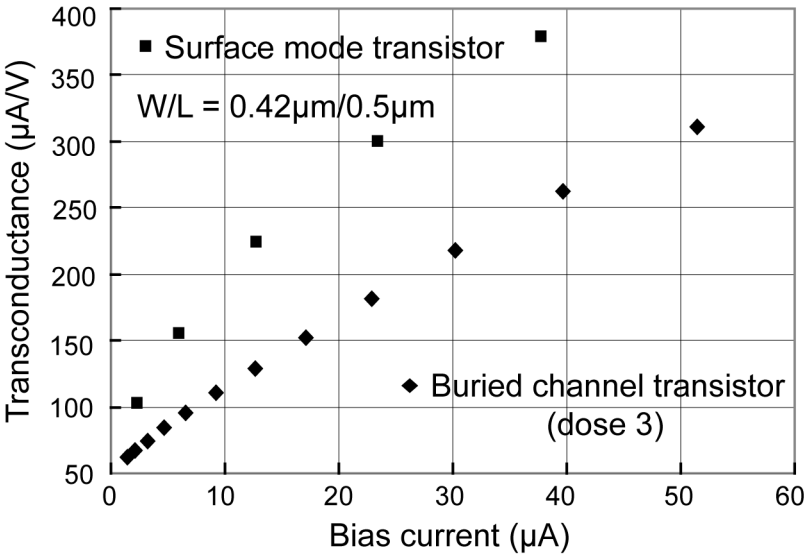


Figure 5-9:Measured transconductance of the surface-mode and buried-channel transistor.

e.g. as shown in Figure 5-10, a clear crossing point can be observed at around 10kHz.

The overall read noise due to 1/*f* noise or RTS noise in CMOS imagers is rather complex because of the CDS architecture in the analog read-out chain. As explained in Chapter 4, the CDS operation acts as a high-pass filter with a sinc transfer function [5.7][5.8], which is determined mainly by the CDS sampling interval. Figure 5-11 re-plots the PSD of the 1/*f* noise shown in Figure 5-10 with a typical CDS transfer function added. The overall read noise is actually determined by the intersection area of the PSD curve and the transfer function. In general, reducing the sampling interval, i.e. increasing the CDS frequency, suppresses the 1/*f* noise [5.9].

As shown, how much the imager read noise can benefit from the buried-channel transistor is determined by the area of the 'Gain' and the 'Lose' region indicated in the figure. The difference between these two regions completely depends on the location of the crossing point and the CDS transfer function itself.

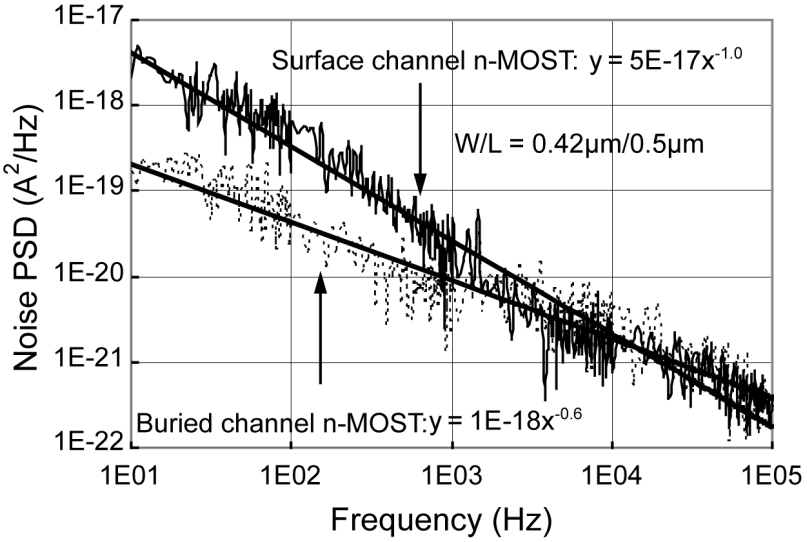


Figure 5-10: PSD of the 1/f noise for the surface-channel and buried-channel nMOST.

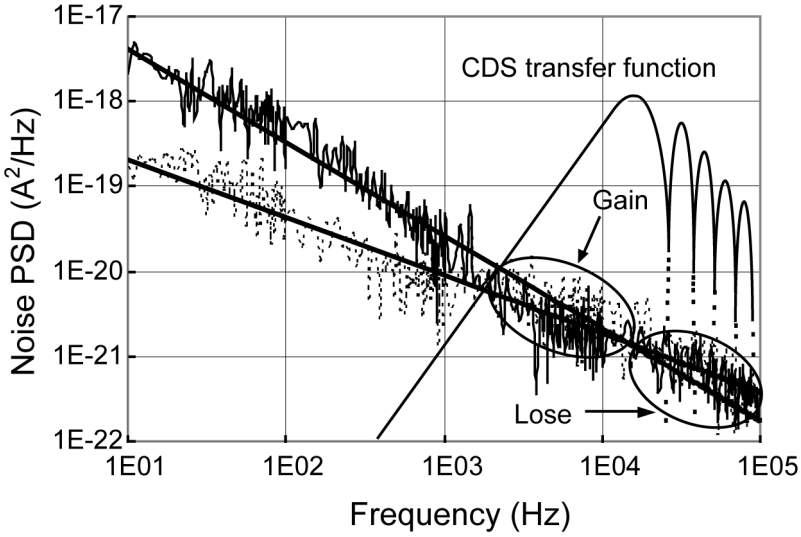


Figure 5-11: 1/f noise measurement with the CDS transfer function.

Clearly, in order to tell how much CMOS imagers may benefit from BSFs in terms of read-out noise, a large number of buried-channel nMOS transistors need to be measured to acquire the statistical data. Since there was both an absence of an actual CDS system in our setup and a lack of enough test samples, no definitive conclusion can be drawn regarding the noise reduction using BSF from this experiment. Thus it can only be proved from measuring actual sensors.

5.3.2 Pixel Output Swing Analysis

In order to evaluate the improvement of the maximum pixel output swing using the BSF, test pixels were fabricated with different source follower configurations.

The saturation level of CMOS imagers is determined by the photodiode full-well capacity, conversion gain, and the maximum pixel output swing. Nowadays, the limitation factor (the bottleneck) is the photodiode full-well capacity instead of the output swing for small pixel designs. This is because "analog" transistors with a thick gate oxide are being used in most of pixel designs. One of the advantages of using these thick oxide devices is the dramatically improved pixel output swing because from the higher supply voltage. However, they consume more space and power compared to digital (thin oxide) transistors.

If the BSF is used inside the pixel, the pixel output swing can be significantly improved, which therefore allow the possibility of using only digital transistors in the pixel design. The absolute pixel maximum output swing is explained in Figure 5-12, and can be determined using:

$$V_{swing} = V_{rst} - V_{FD} - V_{SFth} - V_{col} - V_{RS} \qquad (5\text{-}1)$$

where V_{rst} is the FD voltage after reset, which is one threshold drop from the pixel power supply regardless of the reset mode (for a soft reset, both RST and VDD are tied to the power supply, while for a hard reset, RST is tied to the power supply and VDD needs to be at least one threshold lower), V_{FD} stands for the voltage drop caused

by the FD dark current, V_{SFth} is the threshold drop across the SF, V_{col} is the minimum voltage required at the column to bias the current source and V_{RS} is the voltage drop across the row select transistor. Because the buried-channel nMOS transistors are depletion mode devices with a negative V_{SFth}, the output swing is therefore enhance.

The test pixel pitch is 6µm and all pixels are pinned photodiode designed. Six different test pixels were made under each different implantation solutions with different BSF dimensions. Because each pixel can be individually accessed, the row select transistors were not included in these pixels.

Figure 5-12 also shows the pixel output swing measurement results for different implantation doping and bias current. It can be seen that the output swing of the BSF pixel is nearly double that of the SSF pixel.

As shown, if the bias current is reduced while the implantation dose remains the same, the pixel output further approaches or even exceeds the line of $V_{FD}=V_{out,}$ which indicates that the channel is buried deeper into the silicon. If the bias current remains constant, increasing the implantation dose also pushes the channel deeper. Therefore, in principle, the pixel read-out noise level if dominated by the interface trap related noise will be smaller in the case of a

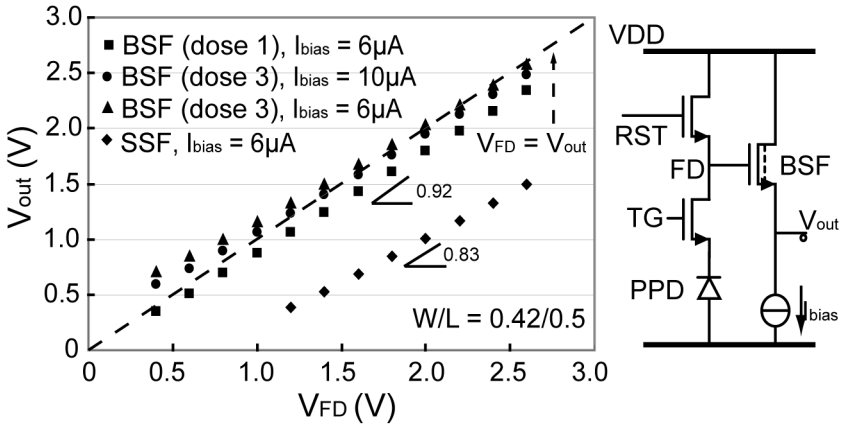


Figure 5-12: Pixel output swing measurement with different implantation dopings and bias currents.

smaller bias current or higher implantation dose of the buried-channel source follower. More importantly, it can be seen that regardless of bias current and implantation dose, all output swing curves tend to head toward $V_{out} > V_{FD}$ at a lower FD voltage, i.e. the source follower continues to operate the buried mode. Therefore, the pixel read-out noise is expected to be reduced as well.

The measured voltage gain of the source follower is improved from 0.83 for the surface-mode devices to about 0.92 - 0.95 for the buried-channel transistor. To conclude, both the pixel output swing and the source follower voltage gain are improved by using the BSFs inside the pixel.

5.4 Sensor Design Overview

As mentioned in sub-section 5.3.1, no definitive conclusion could be drawn from the 1/*f* noise measurement of single transistors regarding the pixel read-out noise reduction using the BSF. Therefore, it is only possible to evaluate the noise improvement with a complete imager noise characterization.

The test sensors were fabricated in a 0.18μm CMOS process by TSMC. Figure 5-13 is the chip micrograph with several fundamental functional blocks of the test sensor. The sensor design is based on an existing prototype from DALSA. The pixel array is 300 x 240 with three different pixel pitches: 6μm, 7.4μm and 10μm. All the pixels are pinned photodiode 4T designs with both BSF and SSF pixels on the same sensor. In the pixel design, the gate signal of all transistors, i.e. the reset transistor, the charge transfer transistor and the row select transistor, can be supplied individually. The fill factors for 6μm and 7.4μm pixels are quite low in order to give flexibility to the source follower sizing.

The row and the column addressing circuitry were realized using a shift register structure. The CDS at each column is used to cancel out the offset, the reset (kT/C) noise, and the 1/*f* noise. The pixel output can be amplified ten times by the CDS amplifier to

already lift the signal and the noise floor from chip level in order to achieve good noise measurement accuracy.

The system clock frequency is 10MHz and the pixel frequency is 2.5MHz. The front-end read timing is supplied by an external FPGA. For the noise measurement, the CDS time interval is 1.5μs and the charge transfer period is 1μs. The outputs of the imager are analog signals that are converted into digital signals by an on-board off-chip image processor with a 12bit ADC.

Overall, the test imager is a conventional design meant to characterize the pixel noise level. All the analog circuitry was implemented as simply as possible to avoid any unexpected noise. The structure of the prototype sensor will not be explained because

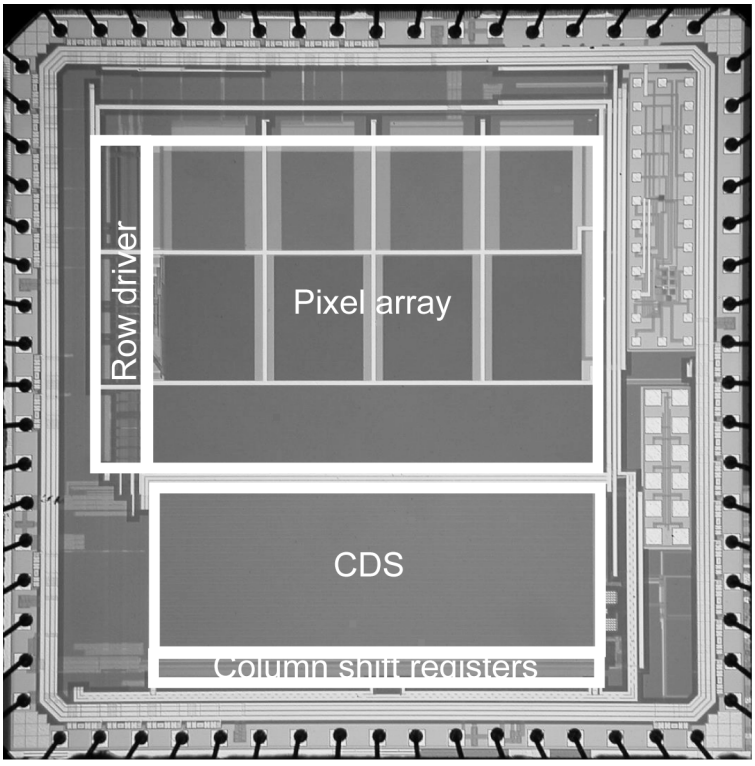


Figure 5-13: Chip micrograph of the test imager.

it is beyond the scope of this chapter. The exact analog signal processing chain of a CMOS imagers is well explained in [5.10].

5.5 Sensor Characterizations

The imager prototype was successfully fabricated and tested. Unlike most of the analog or digital circuitry measurements, which concentrate on either power or speed, the measurement of noise is less straightforward and therefore it involves many uncertainties. In fact CMOS imagers are possibly one of the most complex mixed-signal intergrated circuits on the market today, which routinely contains several million transistors. To narrow down the exact noise source to a single transistor is not easy and requires very extensive work.

Therefore, before jumping to the direct comparison of the 1/*f* noise of the source followers, many "routine" measurements need to take place in order to locate the noise source.

5.5.1 Locating the Noise Sources

As mentioned in chapter 4, the pixel random noise can be measured by calculating the standard deviation of each pixel output along multiframes. Verifying if the random noise is from the APS pixel or the analog processing chain is relatively simple. One common approach is to ground the gate of the row select transistor during operation, after which, the pixel output node can be considered to be floating. The measured random noise is thus the noise of the analog chain.

In practice, if a change in the in-pixel transistor/photodiode operation status during the imager operation (e.g. the integration time, the bias current of the SF, or the CDS period) leads to a significant influence on the measured dark random noise, it can be agreed upon that the measured noise is dominated from the pixel-level noise sources instead of the analog chain.

In order to exclude the contribution of the photon shot noise from the total noise floor, all noise measurements are taken in

complete darkness. The following "routine" procedures are used to determine/verify if the measured random noise is indeed from the in-pixel source follower.

Exposure time vs. Noise: Although the experiments are taken in darkness, an "absolute" dark environment is not always easy to achieve and therefore it needs to be verified to ensure that the photon shot noise does not contribute to the total noise floor. Similarly, even though the dark current generated from the pinned photodiode is extremely low [5.11], caution needs to be taken to ensure that the dark current shot noise does not play a role.

As mentioned in chapter 2, the rms value of the photon shot noise and the dark current shot noise are the square root of the actual video and the dark signal, which are in turn proportional to the exposure/integration time. Therefore, the contribution of these two shot noise sources to the overall noise floor can be measured by changing the exposure time.

In our experiment, when the integration time was increased from 300 to 3000 line times (each line time = 15μs), hardly any changes could be observed from the histogram of the dark random noise and the average noise rms value. These results indicate that the contribution of the photon shot noise and the dark current shot noise to the overall noise floor are negligible.

However, as mentioned in chapter 3, in PPD 4T APS made in a 0.18μm CMOS process, the majority of the dark current is contributed is from the TG instead of the PPD. Since it is independent of the integration time, this dark current source cannot be inspected with this approach.

Charge transfer on/off vs. Noise: The dark current caused by the transfer gate is generated and collected during the charge transfer period. Thus, its difference can be characterized simply by switching on or grounding the TG transistor during the pixel readout.

In our experiments, it was found that the measured noise level was not affected by the charge transmission operation, which suggests that although the dark current generated by the TG transistor may exists, its contribution to the overall read-out noise floor is too small and thus insignificant.

Of course, the junction leakage of the FD can be another dark current noise source. However, because of the very short integration time of the FD dark current, (i.e. the time interval between two CDS samplings), it has very little effect on the overall noise floor. This is also confirmed in [5.12] with some additional experiments by varying the FD integration time and the temperature. The reset (kTC) noise can be neglected in a PPD 4T pixel design with the help of a CDS architecture. Therefore, in principle these experiments are sufficient enough to confirm that the measured noise floor is dominated by the noise sources from the in-pixel source follower transistor.

5.5.2 Dark Random Noise for Surface-Channel and Buried-Channel Source-Follower Pixels

There are three types of noise sources from a source follower transistor: thermal noise, 1/*f* noise, and the RTS noise. Thermal noise in general is considerably smaller than the latter two noise sources [5.7]. Distinguishing between the 1/*f* and the RTS noise is simple because of the discrete character of the RTS noise. A detailed analysis and relationship between the 1/*f* noise and the RTS noise were given in chapter 4.

Figure 5-14 depicts the histogram of the dark random noise measurement for both BSF and SSF pixels. Both the BSFs and SSFs are biased with 6uA current. The FD reset voltage of the BSF pixels is 1.8V, while it is 2.6V for SSF pixels. The transistor dimension for both BSFs and SSFs are the same, being W/L = 0.42μm/0.5μm. The CDS period is 2μs. As shown, the pixel random noise floor of the SSF pixels is around 500μV at the sensor internal gain of 10, which is in the same range achieved in recently reported [5.13]. The averaged dark random noise of the BSF pixels is reduced by more than 50%, and the noise histogram of the BSF pixels closely approximates a true Gaussian distribution with significantly reduced noise spread.

Roughly 0.5 percent of the total SSF pixels are detected as RTS pixels. In a 300 x 120 BSF pixel array, no hot pixel (high 1/*f* noise) or blinking pixel (RTS pixel) was found.

Figure 5-15 shows a test image measured in darkness at 30fps with 10 times analog sensor gain using a 12-bit board level ADC. The upper part shows the raw data, while the lower part shows the data after a digital column FPN cancelation. As seen from the image, the white spots are hardly visible for the BSF pixels.

As explained in the sub-section of the device simulation, the noise reduction efficiency relates to the actual depth and the "potential distance" of the buried-channel, which are determined by the implantation doses, the bias currents, and the gate (FD) voltage. The dark random noise dependency on these factors will be analyzed in detail in the next sub-section.

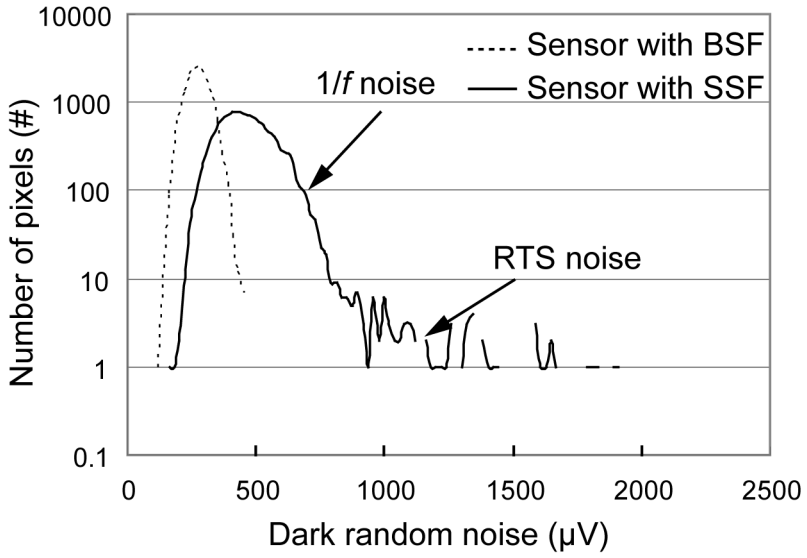


Figure 5-14:Histograms of the dark random noise for BSF and SSF pixels.

5.5.3 Dark Random Noise Dependency of Buried-Channel Source-Follower Pixels

Noise vs. Dose: As mentioned at the end of process simulation, a total of five different implantation solutions were selected for the fabrication of the BSF sensor, with varying implantation energies and total doses. In our experiments, it was found that the implantation solution with the highest energy introduces a large leakage and was therefore not suitable for the source follower application. The other implantation solutions were able to properly produce an image.

Figure 5-16 shows the measured dark random noise (in a linear scale) of the BSF pixels with the highest and the lowest implantation doses that have the same implantation energy under the same bias current of the source follower. The histogram is plotted in a linear scale to highlight the majority of the pixels. As shown, the average of the noise of BSF sensor with the highest dose (319.2μV) was slightly lower than the one of the sensor with the

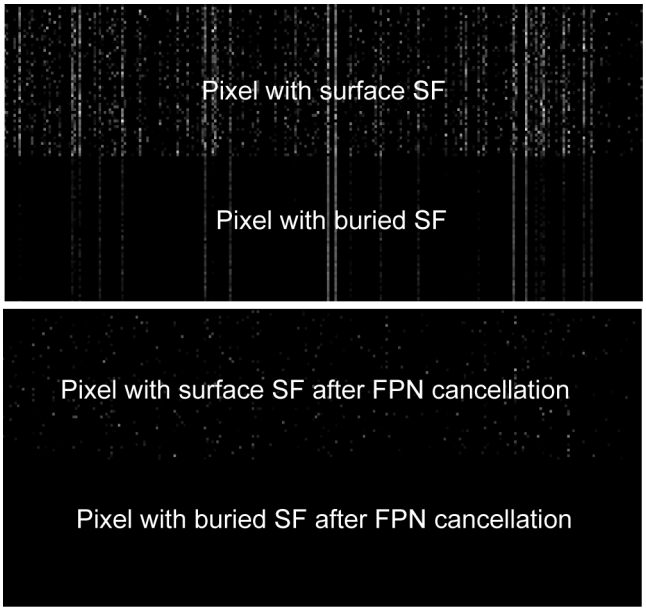


Figure 5-15:Test imagers of the BSF and SSF pixels.

lowest dose (322.7μV). The noise sigma for both sensors was close to 55.7μV. However, the difference (1%) is too small to reveal any systematic relation between the noise and the implantation dose considering the senor-to-sensor noise spread.

In theory, the channel is buried deeper with a higher implantation dose, thus, the random noise is expected to be smaller. However, such dependency is rather weak according to our measurement results. The reason for this may be explained with Figure 5-17, which illustrates how the channel position changes with implantation doses. As shown in the figure, the channel is buried deeper into the bulk silicon with a higher dose. However, as shown in Figure 5-7, for all doping solutions, the effective "potential distance" near the drain side is much greater than the required threshold value (kT/q). A further increase in the "potential distance" in this region cannot bring much benefit to noise reduction. Only the slightly deeper channel at the source side may contribute to a further noise reduction, under the conditions that the

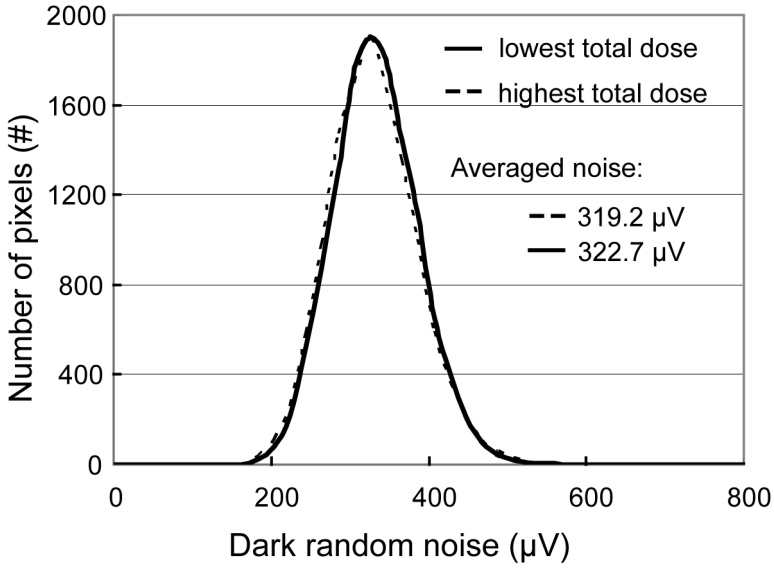


Figure 5-16:Dark random noise of the highest and lowest implantation dosed BSF pixels.

"potential distance" in this region is initially smaller than kT/q and that the interface defects do exist in this region.

In other words, the existence of the buried layer efficiently pushes the channel into the bulk silicon near the drain side because of a high drain voltage (VDD in a normal case). Thus, a significant noise reduction can be achieved. As long as the total dose exceeds the threshold, which ensures the "potential distance" along the full gate length higher than kT/q, a further increase in the dose does not bring extra benefit with regard to noise.

Noise vs. Bias current: As explained previously, reducing the bias current helps to bury the channel deeper. Meanwhile, the possibility of minority carriers being trapped reduces because of the smaller current density. Figure 5-18 shows the measured noise histogram with different bias currents. The experiment was conducted with the same BSF sensor and with the same FD reset voltage. As can be seen from the figure, the random noise was slightly reduced with a smaller bias current, with an average 318.2μV at 6μA bias current and a 322.7μV for the 12μA bias current. The noise sigmas for both situations were similar: around

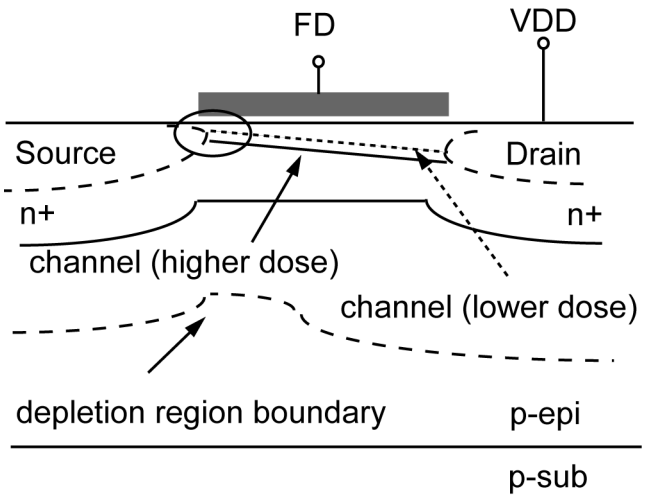


Figure 5-17:BSF device cross-section figure with different implantation dose.

50.64 μV. This matches the theoretical expectation. However, such an improvement is not significant.

Noise vs. FD voltage: Figure 5-19 shows the simulation results of the depletion region and the channel location change of a BSF pixel with varying FD voltages. As can be seen, the "potential distance" and the depletion region at the Si-SiO_2 interface are very sensitive to the gate bias, i.e. the FD reset voltage. Reducing the gate bias helps to push the channel deeper, which, moreover, extends the depletion region at the interface further toward the source side. This means that a larger portion of the total channel length converts from the surface mode into the buried mode operation. Compared to the noise dependency of the total dose and the bias current, the noise dependency of the FD voltage is expected to be more sensitive and systematic.

Figure 5-20 shows the measurement result of the average dark random noise as a function of different FD voltages. As expected, the noise level reduces with the decrease in the FD voltage.

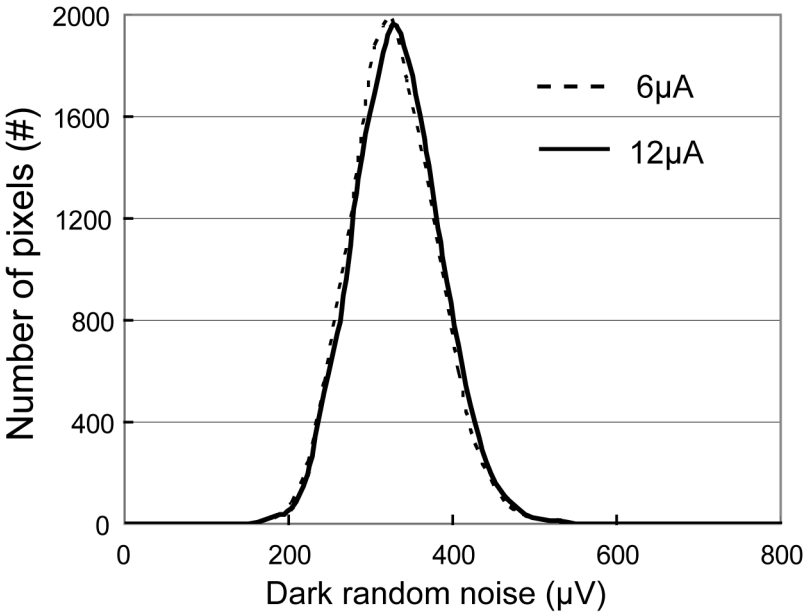


Figure 5-18:Dark random noise with different bias currents.

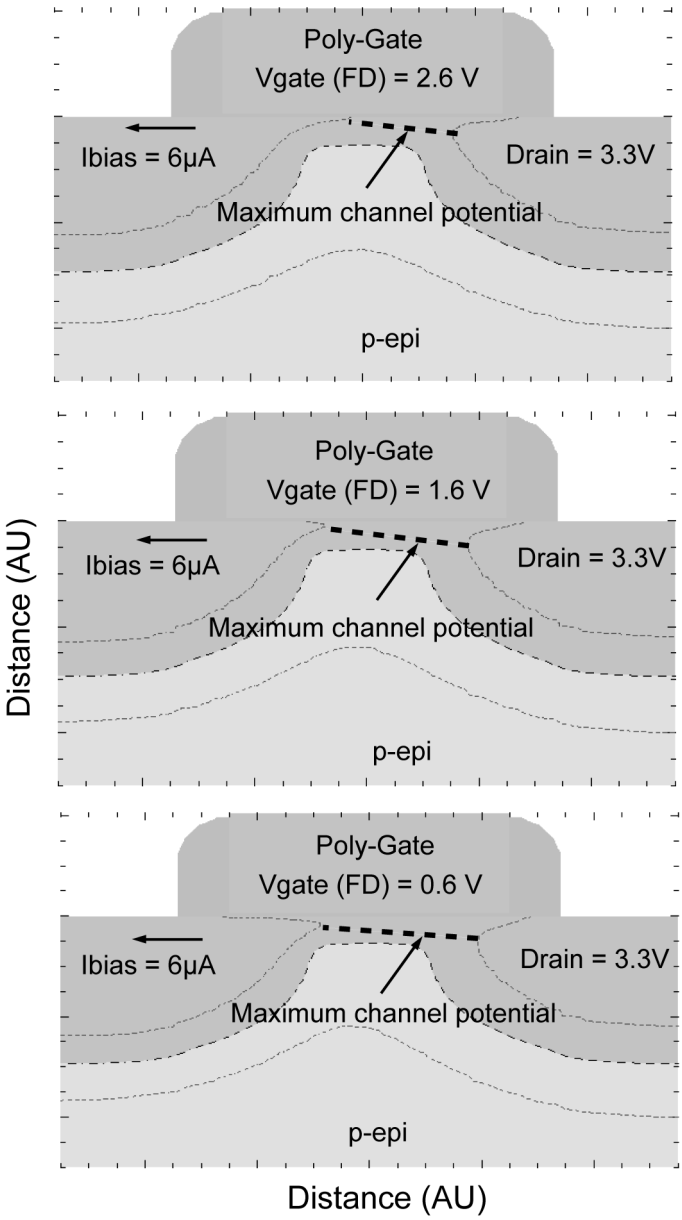


Figure 5-19:Simulation of the depletion region and the channel location of a BSF with different gate biases.

Although a clear dependency was observed, the noise reduction was weak among the shown measurement points: around 15μV with 0.4V voltage difference. In principle, such dependency should be much stronger if the FD voltage increases above 2.5V. However, in practice, such bias conditions cannot be reached. The reason for this will be explained in detail in the next sub-section.

5.6 Trade-Off of BSFs in Pixel Operation

The noise improvement achieved by using BSFs inside the pixels is significant. However, it is important to note that the DC point of the pixel operation changes because of a shifted V_t. Such a change introduces a fundamental trade-off between the maximum pixel output and the image lag.

As explained above, the maximum pixel output swing is significantly improved because of the negative V_t of the BSF transistor. However, such an improvement is limited by the row

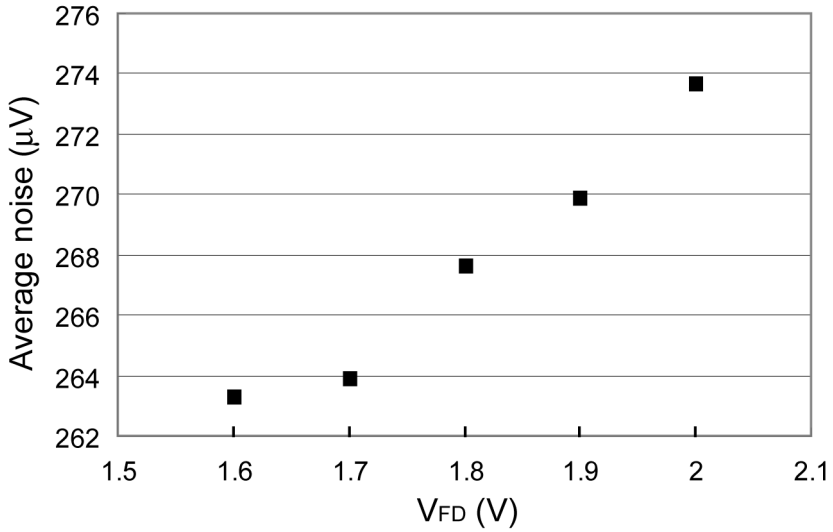


Figure 5-20:Dark random noise measurement with different FD voltages.

Operation condition	Advantage (plus)	Constrains/Risks (minus)
Normal reset (FD_{reset}=Vdd-V_{th})	1: Proper charge transfer (no lag, charge sharing) 2: Higher swing	1: Higher temporal noise 2: Output cut off by RS
Low level reset (FD_{reset} < Vdd - V_{th}	1: Lower noise 2: No output cut off	1:Incomplete charge transfer/lag 2: Reduced swing

Table 5-3.Pixel operation trade-off using BCSF.

select switch, which is normally realized as a standard nMOS transistor. The maximum voltage that can pass through the row select switch is determined by the gate voltage and threshold voltage of this transistor. In other words, in Figure 5-12 for example, although the output of the BSF pixel is able to reach above 2.5V, in reality, it cannot exceed 2.3V with a 3.3V on-voltage of the row select transistor. Thus, the FD reset voltage is expected to be low in order to ensure that the video signal can be properly read-out. Furthermore, as mentioned in the previous sub-section, a small FD voltage is also preferred in relation to the random noise.

However, reducing FD reset voltage brings the potential risk of incomplete charge transfer from the photodiode to the FD region, thus introducing image lag. Table 5-3 is a summary of the pixel operation condition trade-offs explained above.

Therefore, a trade-off exists between the noise reduction and improvement of the output swing (the possibility of introducing image lag). In our experiment, since all noise measurements were taken in darkness, the actual image lag issue was not a concern. However, in real imaging applications, especially when a wide dynamic range is expected, further optimization is necessary to lift the voltage pass level of the row select transistor. To solve this issue, the most straighforward method is to lower the threshold voltage of the nMOS row select transistor, or to implement a transmission gate. However, these methods involve changes in the processing technology, the combination of both pMOS and nMOS

transistors inside the pixel, which, as explained, hurts the pixel fill-factor.

5.7 References

[5.1] M.F. Snoeij et al., "The Effect of Switched Biasing on 1/*f* Noise in CMOS Imager Front-Ends", IEEE Workshop on CCDs and Advanced Image Sensors, pp. 68-71, June 2005.

[5.2] M. Cohen et al., "Fully Optimized Cu Based Process with Dedicated Cavity Etch for 1.75µm and 1.45µm Pixel Pitch CMOS Image Sensors", *International Electron Device Meeting,* pp. 127-130, Dec 2006.

[5.3] C. Leyris et al., "Evolution of RTS Source Activities in Saturation Range in N-MOSFETs for Different Oxidation Temperatures", *18th International Conference on Noise in Physical Systems and 1/f Fluctuations*, pp. 213-216, 2005.

[5.4] R.H. Walden et al., "The Buried-Channel Charge Coupled Device", *Bell Systems Technical Journal*, pp. 1635-1640, Vol. 51, 1972.

[5.5] A.J.P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishers, Dordrecht, 1995.

[5.6] A. Hossinger, *Simulation of Ion Implantation of ULSI Technology*, PhD thesis, 1969.

[5.7] J. Janesick et al., "Fundamental Performance Difference Between CMOS and CCD Imagers" *Proceeding of the SPIE*, Vol. 6276, 2006.

[5.8] H.M. Wey et al., "Noise Transfer Characteristics of Correlated Double Sampling Circuit" *IEEE Transactions on Circuits and Systems*, No.10, Oct 1986.

[5.9] K. Findlater et al., “SXGA Pinned Photodiode CMOS Image Sensor in 0.35μm Technology”, *ISSCC Dig.Tech*, pp. 218-219, Feb 2003.

[5.10] M. Snoeij, *Analog Signal Processing for CMOS Image Sensors*, PhD thesis, 2007.

[5.11] X. Wang et al., “Fixed-Pattern Noise Induced by Transmission Gate in Pinned 4T CMOS Image Sensor Pixels”, *European Solid-State Device Research Conference*, pp. 331-335, Sep 2006.

[5.12] J. Kim et al., “Characterization and Improvement of Random Noise in 1/3.2” UXGA CMOS Image Sensor with 2.8μm Pixel Using 0.13μm-Technology”, *IEEE Workshop on CCDs and Advanced Image Sensors*, pp. 149-152, June 2005.

[5.13] C. Leyris et al., “Modeling of the Temporal Pixel to Pixel Noise of CMOS Image Sensors”, *International Image Sensor Workshop*, pp. 219-222, June 2007.

Chapter 6

Summary and Future Work

In this thesis, both fixed-pattern noise (FPN) and temporal noise in CMOS image sensors were investigated. A new pixel structure using a novel source follower transistor to reduce the pixel dark noise level was proposed. This final chapter summarizes the whole thesis and provides an overview of possible future work.

6.1 Summary

Regarding the pixel fixed-pattern noise:

- The primary offset-FPN in CMOS imager sensors is due to the dark current generated inside pixels. Although the dark current nowadays is so small that it may not noticeable during normal pixel operation, it becomes a salient factor in applications which require long integration time and low illumination.
- The total dark count from a CMOS imager pixel depends on many factors, e.g. sensor fabrication technology, the pixel pitch, and photodiode structure and its biasing condition. But generally speaking, the dominant dark current mechanism in

a 3T pixel is the thermal generation in the photodiode, while for a pinned-photodiode (PPD) 4T active pixel sensor (APS), it is the diffusion dark current that dominates.

- In PPD 4T pixels, a significant amount of dark electrons is generated from the overlap between the transfer gate and the p+ pinning layer. The fundamental mechanism of this dark current is the surface generation, which, however, may also involve the hot carrier (HC) effect. Lowering the floating diffusion (FD) voltage helps to reduce the total dark count. However, doing so also reduces the pixel output swing and may introduce image lag.
- Because of the improved fabrication technology and further optimized pixel structure available nowadays, the dark current of CMOS imagers has been reduced significantly. Consequently, for most imaging applications, the dark current shot noise becomes so insignificant that it can be even ignored with respect to the total pixel dark temporal noise floor. However, the shrinking of the pixel size in scaling processes may introduces new effects that increase dark current, e.g. if the doping concentration of the substrate is increased.

*Regarding the pixel temporal noise and the RTS noise***:**

- Conventionally, 1/*f* noise is believed to dominate the pixel random noise floor in a PPD 4T APS. However, when the process scales down, instead of the well-known 1/*f* noise, it is actually random telegraph signal (RTS) noise that is exhibited. The RTS noise is induced by a single trap located in or near the Si-SiO_2 interface of the in-pixel source follower transistor.
- The RTS noise introduces "blinking pixels" from which the output exhibits three discrete levels. It has been found that the RTS noise composes the majority of the tail noise in the pixel temporal noise histogram acquired from PPD 4T CMOS imagers made in a 0.18μm CMOS process.

- The two most important parameters to describe the RTS trap behaviour are the trap capture time (τ_c) and emission time (τ_e). The RTS noise behaviour can be modeled using the probability of trap occupancy (PTO) parameter, which is determined by τ_c and τ_e, and can be extracted in experiments.
- The CDS period has a significant influence on the RTS trap time constant in a non-equilibrium state, and thus on the overall pixel read-out noise as well. The RTS noise decreases with the reduction of the CDS period, i.e. during faster CDS operation. However, the increase of the CDS frequency can only suppress the RTS noise instead of eliminating this noise completely.
- There is a turning point in the temperature dependency of the RTS noise. If the sensor operation temperature increases before this point, the overall transient PTO decreases and the noise increases. After the turning temperature, the thermal excitation mechanism becomes too strong and the trap is likely to be filled by the electrons excited from the valence band, and thus the overall PTO increases and the temporal noise reduces.
- The energy level of the interface traps that cause the RTS noise can be extracted with noise measurements taken at different operation temperatures. The experiment proves that the RTS-trap energy level is very close to the silicon conduction band.
- The RTS noise amplitude in CMOS imagers cannot be fully explained using classical RTS theories. Moreover, the McWhorter $1/f$-RTS noise model is challenged according to our experiments. The exact mechanism of the RTS noise, or in other words, how the trapping and de-trapping process of a single carrier manipulates the transistor conducting current, is still unknown and deserves further studies.

Regarding CMOS imager sensors with buried-channel source follower (BSF) transistors:

- The main motivation to use a buried-channel transistor as the in-pixel source follower is to take the conducting carriers away from the Si-SiO_2 interface so that interface traps-induced noises can be suppressed.
- Process simulation proves the possibility of creating buried-channel transistors using a deep sub-micron CMOS process with only one additional implantation step while the following fabrication steps remain unchanged.
- Device simulation proves that the "potential distance" of a buried-channel transistor that is biased in the source follower operation mode is well beyond the desired value. Thus, a significant noise reduction can be expected from using such a device as the in-pixel amplifier.
- Experiments show that the averaged dark random noise of BSF pixels is reduced by more than 50% compared to surface-mode source follower (SSF) pixels, and that the noise histogram of BSF pixels closely approximates a true Gaussian distribution with significantly reduced noise spread. Using BSF pixels also significantly reduces the amount of blinking pixels due to the RTS noise.
- As shown in Figure 5-19, the conducting channel at the drain side of a BSF transistor is efficiently pushed down into the bulk silicon because of a strong positive drain voltage. Thus, a significant reduction of the pixel dark temporal noise can be easily achieved using a BSF. A further increase in the buried layer dose cannot bring further benefit from the noise perspective.
- The temporal noise of the pixel with a BSF is rather sensitive to the gate bias voltage of the source follower transistor. Reducing this voltage helps push the channel deeper and extend the depletion region at the interface further towards the source side. Thus, the overall noise decreases.

6.2 Future Work

As previously explained, the scaling of CMOS imager fabrication processes introduces new challenges in terms of the sensor performance, particularly to noise behavior. The following topics would be worthwhile to investigate in future work to gain a better understanding of noise in CMOS imagers and to create pixels with a lower noise level.

- Further investigation of the generation mechanism of the dark current from the overlap between the transfer gate (TG) and the PPD. The exact influence of the strong electric field on the total dark count is still unclear. Dark current measurement of pixels with different TG and PPD overlap distances could be performed to determine the relation between the dark current generation rate and the electric field. When fabrication processes scale down, the electric field across the overlap region may become stronger due to the smaller device size. Also, the influence from the edge of the TG may become more significant because of a smaller TG size, thus creating a 3-d effect. Such influence further complicates the dark current generation mechanism. Thus, a device solution that minimizes this dark current component without compromising the transfer efficiency becomes very appealing, particularly for very small pixels made using more advanced technologies.
- Further investigation regarding the physical mechanism of the RTS noise from in-pixel source follower transistors. Although the RTS noise has been widely observed in CMOS imagers made in deep sub-micron processes, it is not entirely understood how a single electron can indeed introduce noise up to several milli-volts. Moreover, the design factors which can be adapted to compromise this dominant noise source are unknown. Sensor noise measurement with varying source follower transistor dimensions and RTS-trap properties or densities (e.g. through radiation or different gate biasing) will provide valuable information on the noise nature.

- The appealing circuitry solutions to remove or suppress the RTS noise. Until now, most of the effective approaches to reduce the RTS noise are either based on the device [6.1] or the process solutions [6.2] which normally require extensive technology support. This constraint limits the application of these methods. Alternatively, circuitry solutions can be easily adapted to be more or less independent of the fabrication technology, which thus becomes a very attractive option. The first approach could start by investigating techniques whose efficiency of reducing the $1/f$ noise is proven to be effective in reducing the RTS noise. Because the RTS noise is caused by a single interface trap and those traps from different transistors share similar energy levels, techniques like large-signal excitation [6.3] may produce impressive results if their sampling frequencies are optimized for these particular traps. Another approach that could benefit from the fact that the RTS noise amplitude is symmetric in the time domain, e.g. adding a median filter based on multi-sampled pixel outputs. Thus, as long as the samples contain the full RTS pattern (3 levels), this median filter will be able to capture the actual pixel output precisely with the RTS noise free.
- Further optimization of CMOS image sensors with BSF. The prototypes described in chapter 5 focus solely on noise behavior. Thus, further optimization regarding the overall image quality is necessary mainly in the following aspects:
  - Since the output voltage level of the BSF pixel is lifted up by the reduced threshold voltage of the BSF transistor, the row-select switch transistor inside each pixel and CDS sample switches in every column need to be adjusted to be able to accommodate sufficiently high voltage levels.
  - In order for the produced-image to be able to benefit from an improved pixel output voltage swing and to acquire a wider dynamic range, the pixel full-well capacity needs to be improved. For small pixels, the charge capacity is limited by the PD and not the source-follower. To do so, the pixel pitch can be increased while the pixel fill factor has to be opti-

mized.
- To achieve the minimized pixel dark noise level, the FD voltage is expected to be as small as possible. However, a small FD voltage introduces the chance of incomplete charge transfer. To find the optimal operation condition, image lag experiments with respect to different FD reset voltage can be performed.

6.3 References

[6.1] X. Wang et al., “A CMOS Image Sensor with a Buried-Channel Source Follower”, Technical Digest ISSCC, pp. 62-63, San Francisco, US, Feb. 2008.

[6.2] C. Leyris et al., “Evolution of RTS Source Activities in Saturation Range in N-MOSFETs for Different Oxidation Temperatures”, *18th International Conference on Noise in Physical Systems and 1/f Fluctuations*, pp. 213-216, 2005.

[6.3] M. Snoeij, *Analog Signal Processing for CMOS Image Sensors*, PhD thesis, pp. 57-62, 2007.

Summary

CMOS image sensors are devices that convert illumination signals (light intensity) into electronic signals. The goal of this thesis has been to analyze dominate noise sources in CMOS imagers and to improve the image quality by reducing the noise generated in the CMOS image sensor pixels.

In Chapter 1, a brief introduction was given of the historical background and modern development of image sensors. Nowadays, two types of semiconductor image sensor technologies are used in digital cameras, namely the charge-coupled device (CCD) and the complementary metal-oxide-semiconductor (CMOS) image sensor. CCDs are used mainly in applications that require high image quality, while CMOS imagers dominates in the applications which require complex functionality, low power consumption, and low cost. One big advantage of using the CMOS imager is that it is able to benefit from the ever-shrinking CMOS processes. However, making CMOS imagers with extremely high resolution or small pixel pitch does involve many technical challenges, which inspired the motivation for this thesis.

Chapter 2 gave an overview of the architecture and performance of CMOS image sensor pixels. The purpose was to briefly introduce the advantages and disadvantages of CMOS imagers with different pixel structures. Some crucial criteria on evaluating imaging performance, e.g. the pixel quantum efficiency, dynamic range, full-well capacity, and signal-to-noise ratio were explained as well.

An overview of the physical origin of the fixed-pattern noise (FPN) and random (temporal) noise in CMOS image sensors was also provided.

In Chapter 3, the physical mechanisms of various types of dark current were explained. Two different dark current mechanisms were discussed in detail, namely the thermal generation current from inside the depletion region of the photodiode, and the diffusion current that is collected from outside the depletion region. By theoretically modeling the dark current density, their generation dependencies were shown. From their different temperature dependencies, the approach used to distinguish these two types of dark current mechanisms were discussed. An important original contribution in this chapter is the discovery and characterization of the dark current contribution from the transfer gate (TG) of a pinned 4T pixel CMOS image sensor. Because of the strong electrical field in the overlap between the TG and the photodiode, the dark current generation mechanism is a combination of thermal generation (surface effect) and impact ionization. Its dependency on the charge transfer period, and floating diffusion (FD) voltage were measured and discussed.

In Chapter 4, the random telegraph signal (RTS) noise of CMOS imagers was characterized. One original contribution in this chapter is the discovery of this type of noise and proving of its dominance in the sensor temporal noise floor. Experiments showed how the single-interface-defect-induced RTS noise composes the tail noise in the sensor random noise histogram. Because of the trapping and de-trapping process of a conducting electron by a single interface defect in the gate oxide of the in-pixel source follower transistor, the pixel output exhibits two different levels. The correlated double sampling (CDS) operation thus produces three discrete levels at the sensor output, which results in blinking pixels in the produced image. A theoretically model is built to explain and predict the noise behavior under different sensor operation conditions. Using this model, an experimental approach was developed to extract the RTS

trap properties using only existing sensors during normal operation without the need of additional test structures.

In Chapter 5, a new type of device was introduced: a buried-channel nMOS transistor, which is used inside the pixels in order to reduce both the 1/*f* and RTS noise. All simulation and characterization results presented in this chapter are original. Buried-channel transistors are the transistors whose majority conducting carriers flow far beneath the gate Si-SiO2 interface during operation. The basic principle of the buried-channel transistor and its applications in CCDs were explained first. Then based on the knowledge gained from its applications in CCDs, the targeting buried-channel device parameters were proposed. Finally results of process simulations that study how such devices can be made in a modern CMOS process were shown. The key challenge from technology point of view is the control of the buried layer implantation energy and dose. For this application, it is important to create a buried-channel device which has as low leakage current as possible that meanwhile is still able to obtain as high a potential distance as possible. The following device simulations confirm the buried-channel working principle through the transistor threshold voltages, potential distances, channel depths simulations. It is evident that although the potential distance is not uniform along the transistor channel because of the biasing condition of a source follower, its value is still high enough to exclude the interface effects from the conducting current. The architecture of the prototype sensor was briefly described in this chapter and the noise measurement results of both test structures and prototype sensors were shown. An average temporal noise reduction of 50% is obtained. The amount of blinking pixels (RTS noise) is reduced drastically as well. Also, because of the negative threshold voltage of the buried-channel source follower, the pixel output swing is improved by almost 100%. According to the theoretical and experimental results presented in this chapter, the buried-channel device is very suitable to be used as the in-pixel source follower for imaging applications that require low temporal noise and a high dynamic range.

Samenvatting

CMOS beeldsensoren zetten licht signalen (licht intensiteit) om in elektrische signalen. Deze thesis heeft als doel de dominante ruisbronnen in CMOS beeldsensoren te analyseren en de beeldkwaliteit te verbeteren door vermindering van de ruis die in de CMOS beeldsensor pixels wordt gegenereerd.

Hoofdstuk 1 geeft een beknopte inleiding in de geschiedenis, de achtergrond en de recente ontwikkelingen in beeldsensoren. Tegenwoordig worden er twee types halfgeleider beeldsensoren gebruikt in camera's, namelijk: charge-coupled devices (CCD) en complementaire metaal-oxide-halgeleider (CMOS) beeldsensoren. CCD's worden hoofdzakelijk gebruikt in toepassingen waar een hoge beeldkwaliteit vereist is, terwijl CMOS beeldsensoren dominant zijn voor toepassingen met een hoge complexiteit, een laag vermogen verbruik en een lage kost. Een groot voordeel van CMOS beeldsensoren is dat zij mee genieten van steeds verkleinende CMOS proces technologie. Het ontwikkelen van CMOS beeldsensoren met extreem hoge resolutie of zeer kleine pixel afmeting vormt echter een zware technische uitdaging.

Hoofdstuk 2 geeft een overzicht van de architecturen en prestatie van CMOS beeldsensor pixels. De voor- en nadelen van de verschillende pixel structuren worden besproken. De belangrijkste criteria om de prestaties van een pixel te evalueren worden uitgelegd: de quantum efficientie, het dynamisch bereik, de full-well lading en signaal-ruis verhouding. Er wordt ook een

overzicht gegeven van de fysische oorzaken van de vaste patroon ruis (FPN) en de toevallige (tijdsdomein) ruis in CMOS beeldsensoren.

Hoofdstuk 3 legt de fysische werkingsmechanismen van de verschillende soorten donkerstroom uit. Twee mechanismen worden in detail besproken: de thermische generatie stroom binnen de ruimteladingslaag van de fotodiode en de diffusie stroom die word verzameld van buiten de ruimteladingslaag. Een theoretisch model van de donkerstroom dichtheid toont de afhankelijkheden van de generatie snelheid. Door hun verschillende temperatuursafhankelijkheid kunnen deze twee types donkerstroom onderscheiden worden. Een belangrijke originele bijdrage van dit hoofdstuk is de karakterisatie van de donkerstroom bijdrage van de transfer gate (TG) van een pinned 4T pixel in een CMOS beeldsensor. Door het sterke elektrische veld in de overlapping tussen de TG en de foto diode is het donkerstroom generatie mechanisme een combinatie van thermische generatie en impact ionisatie. De afhankelijkheden van de ladingstransfer periode en vlottende diffusie spanning worden gemeten en besproken.

Hoofdstuk 4 karakteriseert en bespreekt de “toevallig telegraaf signaal” (RTS) ruis in CMOS beeldsensoren. Een originele bijdrage van dit hoofdstuk is het vinden van dit type ruis en bewijzen dat het dominant is in de tijdsdomein ruis vloer van de sensor. Door middel van experimenten wordt aangetoond hoe de staart van de distributie van de toevallige ruis van de sensor bepaald wordt door de RTS ruis veroorzaakt door enkelvoudige oppervlakte defecten. Door het vangen en loslaten van een geleidingselectron door een enkel oppervlakte defect in het gate-oxide van de in-pixel source-volger heeft de uitgang van het pixel twee verschillende niveaus. De gecorreleerde dubbele bemonsteringsoperatie (CDS) creeert hieruit drie verschillende niveaus aan de sensor uitgang, met als resultaat knipperende pixels in het beeld. Een theoretisch model is opgesteld om het ruis gedrag te verklaren en te voorspellen onder verschillende sensor werkingsomstandigheden. Op basis van dit model is er een experimentele aanpak ontwikkeld om de

eigenschappen van de RTS vallen te bepalen, enkel gebruik makend van sensoren in normale werking, dus zonder bijkomende teststructuren.

Hoofdstuk 5 introduceert een nieuw soort component: de begraven kanaal nMOS transistor voor gebruik in het pixel om zowel de 1/f als de RTS ruis te verminderen. Alle simulatie en karakterisatie resultaten die gepresenteerd worden in dit hoofdstuk zijn origineel werk. De term begraven kanaal transistor duidt op een transistor waarvan tijdens de werking de meerderheid van de ladingsdragers diep onder het gate silicium-siliciumdioxide raakvlak vloeien. Het basisprincipe van een begraven kanaal transistor en zijn toepassing in CCD's worden eerst uitgelegd. Gebaseerd op de kennis uit de toepassing in CCD's worden begraven kanaal component specificaties voorgesteld. We tonen de resultaten van proces simulaties die onderzoeken hoe zulk een component kan gemaakt worden in een hedendaags CMOS proces. De belangrijkste uitdaging vanuit technologie standpunt is de controle over de begraven laag implantatie energie en dosis. Voor deze toepassing is het belangrijk dat de begraven kanaal component een zo laag mogelijke lektroom heeft en tegelijkertijd toch een zo hoog mogelijke potentiaal afstand bereikt. De component simulaties bevestigen de begraven kanaal werking door transistor drempel spanning, potentiaal afstand en kanaal diepte simulaties. Het is aangetoond dat, hoewel de potentiaal afstand niet uniform is langs het transistor kanaal door de werking als source-volger, de waarde hoog genoeg is om oppervlakte effecten op de stroom uit te sluiten. De architectuur van de prototype sensor wordt kort beschreven in dit hoofdstuk en de ruis metingen op zowel teststructuren als prototype sensoren worden getoond. Een gemiddelde verlaging van de tijdsdomein ruis met 50% wordt bereikt. Het aantal knipperende pixels (RTS ruis) wordt eveneens drastisch verminderd. Bovendien wordt door de negatieve drempelspanning van de begraven-kanaal source-volger de pixel uitgaanszwaai met bijna 100% verbeterd. Volgens zowel de theoretische als de experimentele resultaten die in dit hoofdstuk worden voorgesteld is de begraven kanaal component erg geschikt voor gebruik als in-pixel source-volger voor

beeldsensor toepassingen waar nood is aan lage ruis en een hoog dynamisch bereik.

(Translated by Manuel Innocent)

简述

基于 CMOS 工艺的图像传感器是一种用于将光强度信息转化成电子信号的半导体器件。这篇论文的主要目标是对 CMOS 图像传感器的像素噪音组成进行分析，通过降低其最为显著的噪音源从而提高生成的图像质量。

第一章对各类图像传感器的历史背景和现代发展进行了简单的介绍。当今的数码相机主要使用两种不的半导体图像传感技术：电荷耦合器件 (CCD) 以及基于互补金属氧化物半导体 (CMOS) 技术的图像传感器。CCD 主要用于需要高图像质量的各类应用，而 CMOS 图像传感器主要用在需要复杂功能，低功耗，低造价的领域。CMOS 传感器的一个重大优势在于它得宜于不间断的半导体技术的革新。但是，制造极高分辨率或者极小像素尺寸的 CMOS 传感器事实上涉及到非常多的技术挑战。而正是这些挑战构成了这篇工作的最原始的驱动力。

第二章简单概括了现代 CMOS 图像传感器的像素结构极其各项性能，其目的是为了分析他们各自的优势和不足。为了后文的图像传感器的性能比较，这一章同时定义了一些关键的用于评估其像素性能的参数，比如传感器的量子效率 (QE)，动态范围 (DR)，全阱容量 (FWC) 和信噪比。而在这些参数之中，CMOS 图像传感器的静态 / 固定模式噪音以及起动态 / 瞬时噪音更是尤为重要。因此，这一章也对各种已知的静态或者动态噪音源进行了比较详细的阐述。

暗电流，做为一种图象传感器中的噪音源，无论是对 CCD 还是 CMOS 器件的性能都起至关重要的影响。因此，论文的第三章对各种不同的暗电流的物理机制进行了详细的分析，尤其针对着两种最为显著的暗电流生成机理：热生成电流以及扩散电流。热生成电流主要生成于像素的光电二机管的耗尽区，而扩散电流主要从耗尽区外收集而得到。因为这两种暗电流的物理

生成机制的不同，他门在不同像素结构中的比重也不尽相同。通过分析他们的各自的电流密度模型，尤其是其不同的温度变化率，我们可以相对容易的判断他们在各类不同 CMOS 图象传感器中的影响程度。在第三章，一个重要的独创性贡献在于发现了在 4T CMOS 传感器中由传输门 (TG) 而引发的特异的暗电流的存在。由于在像素中传输门和光电二极管的重叠区存在着极强的电场，由这部分生成的暗电流存在着其与其他种类暗电流不同的特性。本质上来说，这类暗电流类似于普通热生成电流，是由硅晶体（或者硅 / 氧化硅交界面）的杂质而产生，但是，生成的暗电流电子被周围的强电场加速，从而导致了可能的击穿效应。结合实验结果，这一章对这种特殊的暗电流特性，参数，以及其在不同像素工作模式下的不同表现进行了详细的分析和论证。

第四章分析了 CMOS 图像传感器的随机电报信号 (RTS) 噪音。在这章中，主要的独创性的工作在于发现了这种噪音源的存在，并证实了这种噪音事实上已经成为整个传感器的背景噪音中最为显著的组成部分。从物理机理上说，RTS 噪音由单一的交界面杂质的导致：假设在像素内源点跟随器的门极氧化硅中存在一个杂质（阱），当晶体管器件导通时，这个阱就会不断的捕捉并释放导通电流中的电子，从而使电流发生波动。这种波动导致了像素的电压输出在两个离散值之间进行随即跳跃。因为随后的相关双次采样 (CDS) 操作，最终的传感器输出将会在三个不同的离散值随机变化，也就是所谓的 RTS 噪音。如果这种效应体现在最终生成的图象上，这个特殊的 RTS 像素将成为一个闪点，显而易见，大量闪点的存在将大大降低整体图象质量。本章的另外一个重要的独创性贡献在于它对 RTS 噪音进行的理论建模，并根据此模型分析预测了这种噪音源的温度变化率，和其在不同像素操作模式下不同的表现特性。应用这个模型，我们发展了一套可以直接用来提取 RTS 杂质特性的实验方法。应用这种方法，所有的杂质信息，比如杂质激活能级等，可以直接从已存在的 CMOS 传感器中提取而不需要额外的特殊测试器件。这一章的最后对 RTS 噪音和传统 $1/f$ 噪音之间的关系进行了阐述，并在此提出了对经典 RTS 噪音理论的挑战。正因为对这两种噪音生成机制还不是完全了解，如何同时降低其噪音值成为了一个相对困难的命题。

第五章提出了一种新型的器件，其目的在于同时降低 CMOS 图像传感器中的 RTS 和 $1/f$ 两种噪音源。其基本概念在于：通过在使用埋沟道晶体管取代表面沟道器件作为像素源点跟随器来降低其门极氧化硅交界面杂质对导通电流的影响。理论上，因为 RTS 噪音和 $1/f$ 都是由交界面杂质产生，通过使用这种埋沟道器件，两种噪音都会被大大降低。而从现实角度，最大的挑战在于如何在使用现代的半导体工艺技术来制作这类器件，在降低噪音的同时而又不会损害其他重要的性能参数。

本章首先阐述了埋沟道器件的基本理论和其在CCD领域中的应用，并根据其在 CCD 使用中的经验推断出其需要达到的器件参数标准。 其后的工艺仿真详细论证了在现代 CMOS 工艺下制作这类器件的可行性。主要的工作在于如何优化埋沟道层的离子注入操作。不同的离子注入剂量，步骤，能量将改变对应生成器件的沟道深度，势能差，漏电流，晶体管阀值电压，跨导等各项参数。所以随后的器件仿真结合了工艺仿真结果对上述所有参数进行更为详细的分析。结果显示在现代 CMOS 工艺下，虽然埋沟道器件的实际物理沟道深度难以达到类似BCCD的范围，但是其势能差仍旧可以足够大，从而在理论上忽略表面杂质的影响。为了验证理论概念和仿真结果的正确性，我们设计了一个完整的 CMOS 图像传感器测试芯片。芯片的结构和所选用的像素类型也在这一章进行了简单的描述。文章最后介绍了测试芯片和测试器件的测量结果。实验显示：通过使用埋沟道源点跟随器，测试图像传感器的动态噪音被降低了 50%，同时，在生成的图象中，闪点的数量也被大大降低。因为埋沟道器件的阀值电压接近负值，所以像素的模拟信号输出范围也提高了近 100%。

Acknowledgements

The first piece of work in this thesis started in January 2004, when I just graduated from my M.Sc program and had absolutely no idea what CMOS image sensor is. Now, more than four years later, all these small pieces of works finally make into a full thesis. I would like to thank the people who made all of these possible, and also they made this four years of mine really enjoyable.

First of all, I would like to thank my promotor, Prof. Albert Theuwissen for giving me the opportunity to work on such an exciting topic. Being my promotor, supervisor and mentor, he practically lead me go through all difficulties I used had. I really appreciate for all his support, encouragement, patient and trust in me along all the years.

Secondly, I would like to thank all my colleagues at the Electronic Instrumentation Laboratory for making the Lab a big warm family. Specially I want to thank to our image sensor group members: Rao, Bongki, Ning, Yue, Gayathri for their support on my work and all interesting discussions. In particular I want to thank Martijn Snoeij for his assist for my first tape out, and proof-reading my thesis.

Also, I would like to thank Prof. Erik Bodegom, during his one year of staying in Delft, we sited together for countless hours to debug test system, analyse measurement data. He showed me the way how research is performed.

Certainly, this work will never been accomplished without the help from experienced experts from industry. Here, I want to first thank the people at DALSA for their supporting on design test vehicle, all the technique discussions and their assistance on measurement, Edwin Roks, Jan Bosiers, Inge Peters, Willy Maes, Daniel Verbugt. In particular, I want to thank Adri Mierop for all his help in almost every subjects, from design to characterization. Also,

I want to thank the CMOS image sensor R&D group at TSMC for their support on our tape out. Changing fabrication process isn't always easy, it is their warmly help that makes it possible, S.G.. Wuu, Ming Li, Walter, specially thanks to Vincent Hsu for his personal supervision on our split lots.

I would like to thank Manuel Innocent from Cypress Semiconductor for proof-reading my thesis, and for his help on translating the summary and propositions into dutch. I also want to thank Dan Mcgrath from Eastman Kodak for giving very detailed technique comments on my thesis. It helps a lot to improve the overall quality. And certainly, I also would like to thank STW for supporting this project.

Last but definitely not the least, I want to thank my families. For me, not able to live with my parents and take care of them is a big regret in my life. I do appreciate their understanding and being supportive. And, I want to thank my wife for always being there for me. Living in a strange country is tough, it is her who makes this cold, rainy place a home where I can always turn for comfort, support, encouragement, and most of all, love.

Xinyang Wang
Delft, September 2008

Publications

Key Publications

[1] **Xinyang Wang**, Martin. F. Snoeij, Padmakumar R. Rao, Adri Mierop, Albert J.P. Theuwissen, "A CMOS Image Sensor with a Buried-Channel Source Follower", pp. 62-63, *ISSCC Dig. Tech.* Feb 2008.

[2] **Xinyang Wang**, Padmakumar R. Rao, and Albert J.P. Theuwissen, "Characterization of Buried Channel n-MOST Source Followers in CMOS Image Sensors", pp. 223-225, *International Image Sensor Workshop*, June 2007.

[3] **Xinyang Wang**, Padmakumar R. Rao, Adri Mierop, and Albert J.P. Theuwissen, "Random Telegraph Signal in CMOS Image Sensor Pixels", *International Electron Device Meeting*, pp. 115-119, Dec 2006.

[4] **Xinyang Wang**, Padmakumar R. Rao, and Albert J.P. Theuwissen, "Fixed-Pattern Noise Induced by Transmission Gate in Pinned 4T CMOS Image Sensor Pixels", *European Solid-State Device Research Conference*, pp. 331-335, Sep 2006.

Other Publications

[5] Albert J.P. Theuwissen, Martijn F Snoeij, **Xinyang Wang,** Padmakumar R. Rao and Erik Bodegom, CMOS Image Sensors for Ambient Intelligence, AmIware Hardware Technology Drivers of Ambient Intelligence, Philips research book series, SpringerLink, 2006.

[6] Padmakumar R. Rao, **Xinyang Wang**, Albert J.P. Theuwissen, "Radiation Hardness of Deep-Submicron CMOS Technology Evaluated by Gated-Diode Measurements", *poster at Sense of Contact*, April 4th, 2007, Zeist (Netherlands).

[7] Padmakumar R. Rao, **Xinyang Wang**, Adri J. Mierop and Albert J.P. Theuwissen, "Gamma-Ray Irradiation Effects on CMOS Image Sensors in Deep Sub-Micron Technology", *2007 International Image Sensor Workshop*, June 7-10, 2007, Ogunquit (Maine).

[8] Padmakumar R. Rao, **Xinyang Wang**, and Albert Theuwissen, "Degradation of spectral response and dark current of CMOS Image Sensors in deep sub-micron technology due to gamma//-irradiation", *ESSDERC*, Munich, Sept. 11-13, 2007.

[9] **Xinyang Wang**, Padmakumar R. Rao, and Albert J.P. Theuwissen, "Noise in Deep Sub-Micron CMOS Image Sensors", *poster at Sense of Contact*, April 4th, 2007, Zeist (Netherlands).

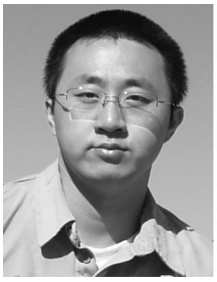
[10] Johan Leijtens, Albert Theuwisen, Padmakumar R. Rao, **Xinyang Wang**, Ning Xie, "Active pixel sensors: the sensor of choice for future space applications", *SPIE Europe Remote Sensing Conf*., Florence (Italy), Sept. 17-20, 2007.

[11] Padmakumar R. Rao, **Xinyang Wang**, and Albert J.P. Theuwissen, "Degradation of CMOS Image Sensors in Deep Sub-micron Technology due to Gamma-irradiation", pp. 1407-1413, Volume 52, Issue 9, *Solid-State Electronics,* September 2008.

[12] Padmakumar R. Rao, **Xinyang Wang**, and Albert J.P. Theuwissen, "CCD Structures Implemented in Standard 0.18-micrometer CMOS Technology", Vol. 44, No. 8, *Electronic letters*, April 2008.

[13] Ning Xie, Albert J.P. Theuwissen, **Xinyang Wang**, Johan Leijtens, Henk Hakkesteegt, Henk Jansen, "A CMOS Image Sensor with Row and Column Profiling Means", *IEEE Sensors 2008*, Oct 2008.

About the Author



Xinyang Wang was born in Harbin, P.R.China, on Feb 6, 1980. He studied at Zhejiang University, Hangzhou, from 1998 till 2002 where he received his B.Sc degree in Electrical Engineering. In 2004, he received his M.Sc degree from the University of Southampton, Southampton, UK on Microelectronics System Design.

In January 2002, he joined the Electronic Instrumentation Laboratory and worked as a Ph.D candidate focusing on CMOS image sensors. His research mainly concentrated on pixel-level noise of CMOS image sensors that are made in deep sub-micron CMOS processes.

Since May 2008, he is working as a pixel designer for Custom CMOS image sensors at Cypress Semiconductor in Mechelen, Belgium. His main research interest is in the design and optimization of CMOS imager pixels and their process development.