Foveated Sampling Architectures for CMOS Image Sensors

by

Fayçal Saffih

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2005

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Electronic imaging technologies are faced with the challenge of power consumption when transmitting large amounts of image data from the acquisition imager to the display or processing devices. This is especially a concern for portable applications, and becomes more prominent in increasingly high-resolution, high-frame rate imagers. Therefore, new sampling techniques are needed to minimize transmitted data, while maximizing the conveyed image information.

From this point of view, two approaches have been proposed and implemented in this thesis:

1) A system-level approach, in which the classical 1D row sampling CMOS imager is modified to a 2D ring sampling pyramidal architecture, using the same standard three transistor (3T) active pixel sensor (APS).

2) A device-level approach, in which the classical orthogonal architecture has been preserved while altering the APS device structure, to design an expandable multiresolution image sensor.

A new scanning scheme has been suggested for the pyramidal image sensor, resulting in an intrascene foveated dynamic range (FDR) similar in profile to that of the human eye. In this scheme, the inner rings of the imager have a higher dynamic range than the outer rings. The pyramidal imager transmits the sampled image through 8 parallel output channels, allowing higher frame rates. The human eye is known to have less sensitivity to oblique contrast. Using this fact on the typical oblique distribution of fixed pattern noise, we demonstrate

lower perception of this noise than the orthogonal FPN distribution of classical CMOS imagers.

The multiresolution image sensor principle is based on averaging regions of low interest from frame-sampled image kernels. One pixel is read from each kernel while keeping pixels in the region of interest at their high resolution. This significantly reduces the transferred data and increases the frame rate. Such architecture allows for programmability and expandability of multiresolution imaging applications.

Index terms: CMOS image sensor, Pyramidal image sensor, Multiresolution image sensor, 2D sampling, Foveated Dynamic Range (FDR), Videophone, Video communication, Remote imaging.

Acknowledgements

﴿ رَبِّ أَوْزِعْنِي أَنْ أَشْكُرَ نِعْمَتَكَ الَّتِي أَنْعَمْتَ عَلَيَّ وَعَلَى وَالِدَيَّ وَأَنْ أَعْمَلَ صَالِحًا تَرْضَاهُ وَأَدْخِلْنِي بِرَحْمَتِكَ فِي عِبَادِكَ الصَّالِحِينَ ﴾

من سورة النمل الآية 19

I would like to extend my sincere thanks to my supervisor Dr. Richard Hornsey for his support and patience all along my doctoral research program. His view of CMOS imaging has inspired me to endorse an interesting CMOS vision path I am very confident it will change the course of imaging technology in near and far future.

To my beloved wife Rima and my son Taha I dedicate my present thesis. Their support, encouragements and love were my driving force to write my thesis and carrying on my doctoral research.

I dedicate my thesis to my father Hocine, whom I am always missing …, who taught me the courage of carrying on my endeavours …, and to excel in science and encouraged me to follow up my studies. To my mother, I dedicate my present work for her tremendous effort in raising me since my early steps in this life.

To all my teachers and supervisors, I dedicate this thesis thanking them from the bottom of my heart.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction to CMOS Image Sensors

1.1 Imaging Evolution

It is interesting to follow the evolution of imaging technology from chemical imaging (using photographic film) to electronic imaging (using solid state sensors). Electronic imaging now encompasses two major imaging technologies: Charge-Coupled Device (CCD) imaging, which first appeared in the late 1960's [1], and Complementary Metal-Oxide-Semiconductor (CMOS) imaging, which has developed much attraction since the early 1990's [2]. CMOS imaging is an especially dynamic field of research and development, with advantages in development cost, ease of use and compatibility with surrounding technologies.

Historically, chemical, CCD and CMOS imaging technologies competed for dominance over imaging market share, without completely eliminating each other (as we can see still today film cameras on the market). This co-existence is mainly due to the different physical limitations (beside fabrication cost) of each imaging technology. CCDs, for example are well suited for low light imaging due to their superior low noise, whereas CMOS imagers perform better in high light imaging and integrated applications. These two major electronic imaging technologies will be further presented in the next section.

1.1.1 CCD and CMOS Image Sensors

Both imaging technologies use silicon for light transduction into electrical voltage signals. However, their main difference lies in their sampling architectures. While CCD imaging is based on transporting the integrated photo charges from their pixel sites to the output amplifier (where they are converted into voltages), CMOS imagers (based mostly on Active Pixel Sensor APS) make the charge-to-voltage conversion at the pixel site. This difference is the source of all advantages and disadvantages that divide imaging applications between the two technologies. For example, this architectural difference has given CCD imagers better noise figure, both temporally and spatially (uniformity), making it the technology of choice for imaging applications sensitive to noise, such as astronomical imaging. On the other hand, the architecture of CMOS imagers provides them with sampling flexibility, random accessibility and high integration.

1.1.2 Future Prospects and Trends

It is still unclear which of the two imaging technologies (if either) will dominate the market, although there are progressing signs that CMOS imagers are (and will in the near future) have the largest share due to their low power and architecture flexibility. However, CCD imaging will continue to find application due to its low noise, which makes it the best choice for high-quality imaging applications [3].

1.2 Thesis Objectives

The objectives of this thesis are presented by the following two statements, which comment on the unique performance or characteristic advantages that CMOS imaging technology has over CCDs, namely architectural flexibility and random accessibility. These features are used to address the two critical and related issues of sensor scanning and spatial sampling, as outlined in the philosophical statements below.

1.2.1 Thesis Philosophical Statements

1.2.2 Scanning Statement

Image Scanning should be more adapted to image sampling/acquisition rather than image display compatibility.

Since the appearance of CCD (and later CMOS imaging), raster scanning has become the adopted scanning technique for sampling integrated images at the imager's focal plane. This is partly due to the lack of random accessibility in CCD imagers' architecture, and the compatibility with the widely used raster scanning technique of contemporary display monitors. The flexibility of CMOS imagers has not been fully utilized to optimize scanning, nor was exploited to enhance system capabilities.

The second philosophical statement we explored in the present thesis was more relevant to the future development of the CMOS image sensor. An important feature in any imaging technology is the achieved image resolution. As the image resolution becomes higher and higher, the amount of data transmitted from the imager to a display or an image processor becomes larger and larger. This will impact power consumption and limit the transmission of video signals. In this view we developed our second philosophical statement.

1.2.3 Image Sampling Statement

As image resolutions increases and with it the amount of transmitted image data for display or processing, new architectures are needed for down-scaling the sampling resolution for regions of reduced interest. Innovative architectures are also needed to exploit the Human Visual System for transmitting only the most important regions in the acquired image.

In other words, as the amount of transmitted image data becomes higher, only selected data of interest, (regardless of resolution, dynamic range, or any other attribute), should be transmitted. This is similar to the strategy used in lossy compression algorithms, such as the Joint Photographic Experts Group (JPEG) format, that takes advantage of limitations in human vision to produce smaller image files [61].

1.3 Thesis Motivations and Goals

The main motivations behind the present thesis are to exploit the architectural flexibility of the CMOS imager, and to explore the possibilities it may offer. This intention is also

influenced by the image processing needs of future imagers, as well as the possibility of embedding smart acquisition sampling architectures. Indeed, future imagers are likely to have higher resolutions and higher frame rates, so smart architectures that can minimize the amount of sampled data while maximizing the amount of information would be desirable. Such vision systems are found in biological vision systems, such as human vision, where the spatial sampling is not uniform and yet the system has sharp vision and adaptability to various light intensities and spatial frequency patterns [4]. Finally, we note that, in order to be adopted for future CMOS imaging development, such novel architectures should conform as much as possible to certain principal features and not to be just *ad hoc* designs. These key characteristics are:

Expandability: refers to the lack of increase in complexity when increasing imager size. This means that the suggested architecture is not limited by the size of the imager.

Programmability: refers to the fact that the imager's functionality during sampling of the integrated image should be programmable (controllable) and flexible.

1.4 Contributions

In the context of the motivations mentioned above, a number of contributions have been achieved in this work in two principal themes:

Two-dimensional scanning: We have developed a new and practical architecture based on a standard array of pixels, in which the rows of the classical CMOS image sensor are replaced by rings, and orthogonal output buses by diagonals. The asymmetric information acquisition inherent in raster scanning is thereby replaced by a two-dimensional symmetric image

sampling. It is shown that distortion of moving objects is significantly reduced using this scheme. A new scanning scheme, called bouncing scanning, has also been demonstrated to realize the two integration time profiles for each ring. Fusing the two frames results in a foveated dynamic range (FDR) enhancement. Lastly, it is shown that the dominant form of fixed pattern noise in CMOS image sensors (columnar variations) are perceived less by human observers in the new sensor owing to the diagonal arrangement of the output buses.

Multiresolution CMOS imager: In this architecture the multiresolution approach is implemented on-pixel. Beside sampling and holding the integrated photo charge, each pixel can share charge with neighbouring pixels (given the proper control signals). This charge sharing allows a decrease in resolution on regions of less interest, while maintaining maximum resolution in regions of interest (ROI) where the charge sharing is disabled. Since the multiresolution mechanism (using the charge sharing concept) is implemented at the pixel level, the proposed scheme guarantees most of the requested characteristics discussed in the motivation section. This architecture is presented and discussed in Chapter 7.

1.5 Thesis Organization

CMOS imaging technology is reviewed in Chapter 2, with a focus on image acquisition processes beside our approach in implementing new imaging architectures and their fields of application. Subsequently, the system-level pyramidal architecture design is discussed in Chapter 3. The suggested scanning scheme, supported by the pyramidal architecture and known as the bouncing scanning, is also introduced in this chapter, together with its implications (particularly the foveated dynamic range). The theoretical framework of the

foveated dynamic range is analyzed mathematically in Chapter 4 before being experimentally verified in Chapter 5, which includes testing and characterization results. The discussion of the pyramidal CMOS imager concludes in Chapter 6 by verifying the reduced perception of the pyramidal imager's Fixed Pattern Noise (FPN) noise when viewed by human observer. The implementation and layout of the pixel-based Multiresolution CMOS imager are presented in Chapter 7 before the concluding remarks made in Chapter 8.

Chapter 2

Background: CMOS imagers and Spatial Sampling

In this chapter we will review some architectures used in CMOS imagers for spatial sampling and other objectives, such as dynamic range enhancement and high frame rate. This study will help us subsequently to locate where our suggested approach fits within the established framework, what it contributes to this research field and finally to be able to suggest the future trends for CMOS imager development. Before proceeding with the architectural aspect of APS CMOS image sensors, a short introduction of CMOS imaging technology along with some of its main characteristics is presented.

2.1 Introduction to CMOS Image Sensors

Fig 2.1 shows the construction of a typical CMOS image sensor, from the light sensing photodiode to the focal plane pixel active pixel sensor (APS) and finally the sampling architecture. Initially, the photodiode is reset to a relatively large reverse bias. Photo-electrons are generated both in the photodiode depletion region, from which they then drift towards the diffusion areas, and in the bulk, from which they diffuse towards the photodiode. The photo-generated electrons discharge the photodiode at a rate that is approximately proportional to the incident illumination. Pixels are later reset by row reset logic during the image sampling. Further details of this process are presented in Chapter 3.

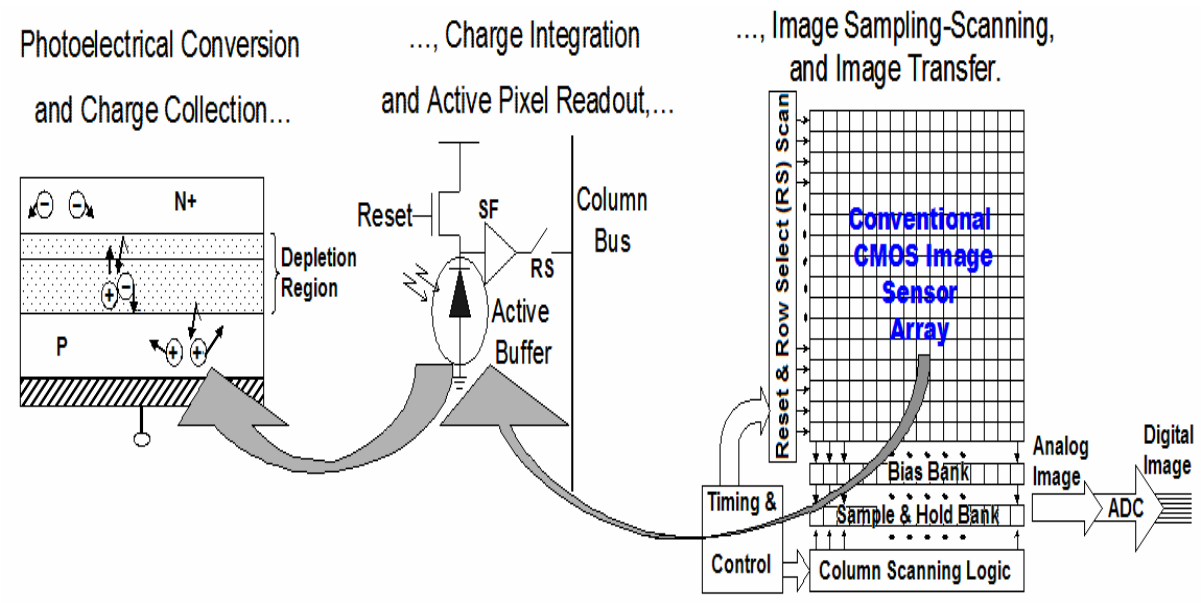


Fig 2.1 From photodiode to CMOS image sensor

The main three types of photodiode pixels widely used in CMOS image sensors technology are shown in Fig 2.2 and are discussed below.

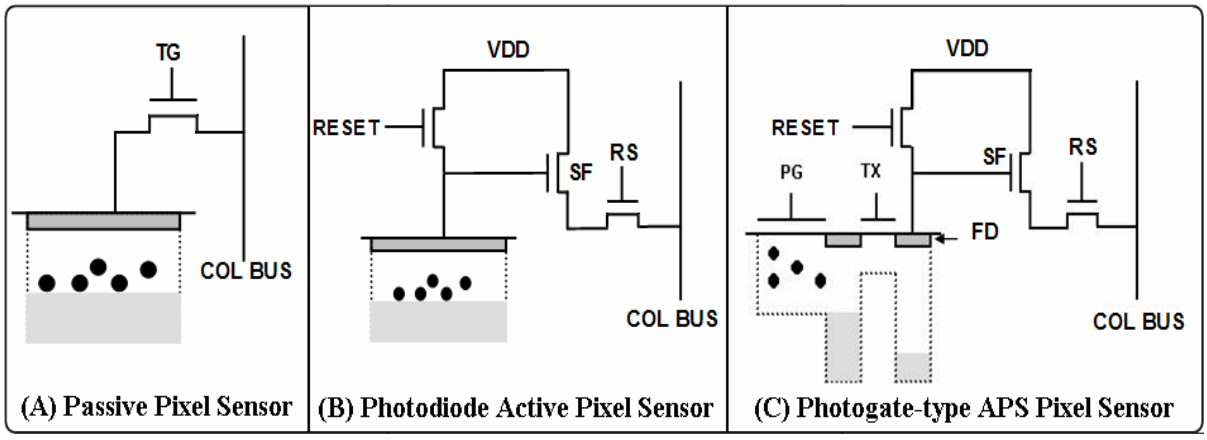


Fig 2.2 The different types of CMOS imaging pixels

2.1.1 Passive Pixel Sensor (PPS)

Fig 2.2.A shows the simplest and earliest (1967) MOS pixel type [5], built simply of a photodiode and a charge-transfer transistor used as a switch. Functionally simple, the passive pixel sensor integrates the light at the active region; the photo-generated charge is then transferred to the column bus *passively* through a transfer gate, simply a MOS switch. The column bus (COL BUS) is connected to a charge amplifier that keeps it at a constant voltage level. When the Transfer Gate (TG) is activated the charge of the photo-diode is transferred through the bus to the column charge amplifier which itself transduces it to the subsequent processing elements such as Analog-to-Digital Converter ADC or Correlated Double Sampling CDS.

Two main problems are attributed to this type of pixels: readout noise due the charge sharing between TG transistor and column bus, and the scalability due to the increased bus capacitance. While the Quantum efficiency of the pixel (ratio of collected electrons to incident photons) of the passive pixel sensor is much higher than that of the CCDs due to its high fill factor and the absence of overlaying layers, it has the read noise of 250 electrons rms compared with typical value of less than 20 electrons rms found in CCDs [6].

2.1.2 Photodiode Active Pixel Sensor

The Active pixel sensor as shown in Fig 2.2.B was firstly suggested by Noble [7] and later investigated further by Andoh at NHK, Japan [8]. The pixel is initially reset to VDD, then after some time, its output is read out through the source follower (SF) transistor in order to generate significant pixel output voltage with larger dynamic range. The readout is

performed through the row select (RS) transistor, which transfers the photo-generated voltage to the column bus. Although its fill factor (ratio of light sensitive to the pixel area) is less than that of the PPS sensor in the range of 20%-30%, it is the most used pixel in research as well as industry. This is due primarily to its low noise (less than 20 electrons rms), higher signal to noise ratio and scalability.

Finally, it is worthy to mention that the photodiode APS noise performance improves as the size of the pixel shrinks down since the reset noise scales as $C^{1/2}$, where C is the photodiode capacitance [9]. However, other counter effects emerge such as decreasing fill factor and sensitivity as the pixel sizes shrinks-down [10].

2.1.3 Photogate Active Pixel Sensor

The photogate APS was introduced first time by JPL in 1993 [11] for high-performance scientific imaging and low light applications. The principal structure of the photogate-active pixel sensor is based on the integration of the photo-generated carriers in the potential well, which is created by applying a large positive voltage to that gate in similar manner with CCD technology. Fig 2.2.C shows the structure of this kind of CCD-like active pixel photosensor. For readout, the output floating diffusion (FD) is reset and its resultant voltage is measured by the source-follower. The photo-generated charge is transferred to the output diffusion by pulsing the photogate. The new voltage is then sensed. The difference between the reset level and the signal level is the output of the sensor. This correlated double sampling suppresses reset noise, 1/*f* noise, and Fixed Pattern Noise (FPN) due to threshold voltage variations. Although this pixel structure has the lowest noise compared to the previous ones (13

electrons rms [12]), it suffers from the reduction of quantum efficiency, particularly in the blue due to the overlaying polysilicon photogate, beside its higher operational complexity and reduced fill factor.

2.1.4 Integration Time & Raster Scanning

In the classical CMOS image sensor architecture, raster scanning was adopted to sample the image at the focal plane. Raster scanning was originally implemented for displaying video images in cathode ray tubes (CRTs), and was later adopted to be the scanning scheme for early solid-state imagers [13] in order to maintain compatibility between acquisition and display systems. This reason is the cause of the statement put forward in section 1.2.2 to propose the new pyramidal sensor with 2-D ring scanning that will be discussed further in section 3.5.8. In raster scanning, the integrated image is sampled row by row, from the first top row until the last bottom row after which the scanning is back to the first row and so on. Because the pixel needs to be reset prior to light integration, every row has to be reset after being read out. The readout is made by sampling the row voltages to a sample and hold bank of capacitors, then the row is reset and sampled again but to another set of sample and hold bank of capacitors, then the same process is repeated for the next row and so on as shown in Fig 2.3. It should be noted that after sampling the row photo and reset signals into their corresponding sampling capacitor banks, the sampled voltages are serially buffered out before proceeding to the next row. Integration time is the time between two successive reset and readout events corresponding to the lapsed time to discharge the reverse-biased (reset) photodiode as a result of light integration. Because of the periodicity of raster scanning, pixel

integration time *Tint* is equal to the frame sampling time *Tframe*, unless other techniques such as the rolling shutter[1], are used. We note here that the raster scan results in a scene update rate that is faster in the column direction than it is in row direction, potentially resulting in spatial distortion in the image capture of moving objects.

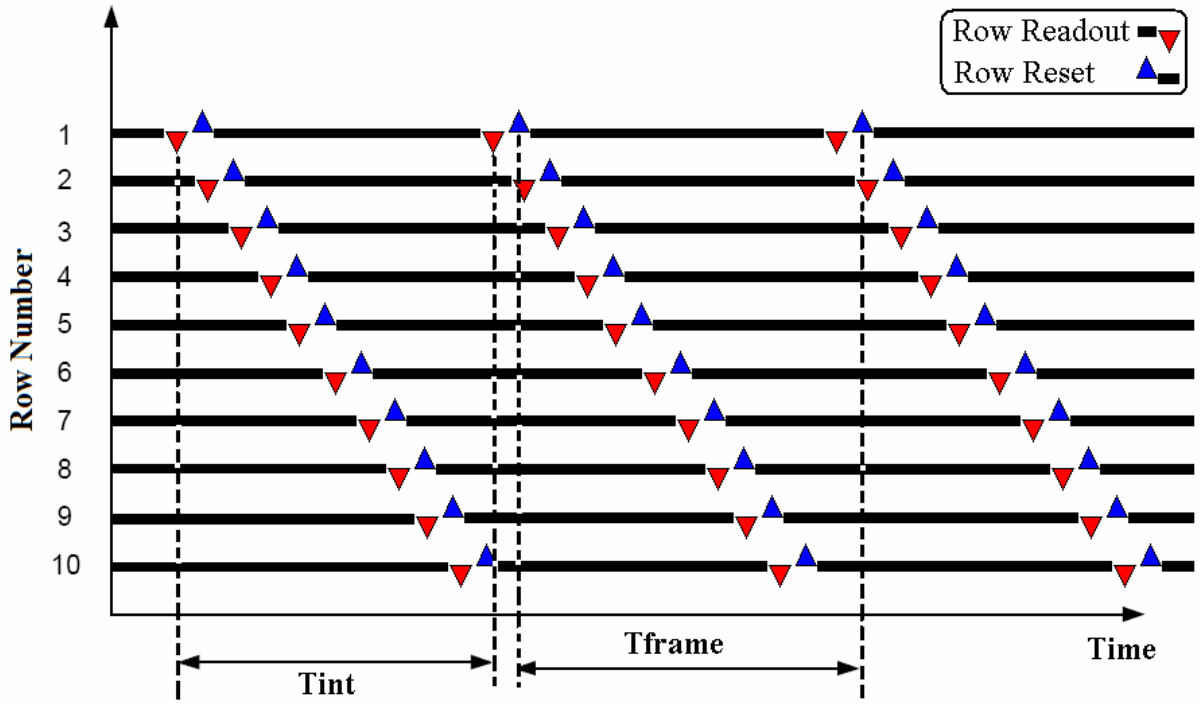


Fig 2.3 Pixel frame integration time in raster scanning

[1] In the rolling shutter scan, the readout (photo signal sampling) signal is applied ahead (by a number of rows) of the reset signal. The number of the separation rows defines a window that its integration time plays the same role of *Tframe* in the raster scan.

2.1.5 Correlated Double Sampling CDS

The reason behind sampling both photo signal and the reset signal of every pixel, is to reduce residual noise at every pixel by subtracting the photo signal from the reset signal. This technique called correlated double sampling (CDS) was originally invented to process CCD imager video signals to enhance the signal to noise ratio [14] by minimizing principally the reset noise. In general, all correlated spatial noise sources such as the spatial dark current source due to a crystal dislocation at the pixel level are cancelled out (or removed) using CDS as they are present in both the photo signal and the reset signal. The *KT*C noise related to the reset process (also called reset noise), however, is not removable by CDS in CMOS APS imager whereas it is completely removable in CCD imagers. The reason is the fact that in a CCD the output noise node is reset prior to photo signal transfer and hence both signals are 100% correlated which makes CDS the solution. In CMOS imagers, on the other hand, the photo signal is read first then the reset signal, and thus the reset noise present in the present photo signal is not correlated with the present noise present in the reset signal (but correlated with the previous frame's reset signal). That is why the CDS technique used in CMOS APS imagers is not a "true" CDS which explains why it is sometimes referred as just double sampling (DS) [15]. Furthermore, noise can also be added by this technique and some of the causes of this addition of noise are discussed in Chapter 6.

2.1.6 Fixed Pattern Noise (FPN)

Fixed pattern noise (FPN) is the spatial noise distribution with no illumination of the image sensor array that is explicitly time independent, and hence "fixed". This noise is due to two

types of mismatch between the photo-signals; photo-signal generation mismatch and the photo-signal transportation mismatch.

The photo signal mismatch is due to the variation of the photo-sensing areas related to process variation and photo-mask errors found also in CCD manufacturing [16], and to mismatched dark currents. The dark current is the accumulation of electrical charge in the photodiode from electron-hole pairs that are generated independent of the photo-detection process. The primary sources of this are impurities or lattice defects in the silicon substrate. Because these defects are localized, the dark current is different for each pixel, leading to a fixed pattern noise in the image resembling a starry sky or a dirty window in a dark image at long integration times [17].

Therefore, if we have to classify the sources of the FPN we can categorize them in two: (i) local impurities and silicon crystal imperfections, and (ii) the different paths taken by the signals during readout. CCD does not suffer so much from FPN partly because its high quality silicon substrate is built on for high transfer efficiency requirement as compared with commercially lower quality (and cheaper) CMOS substrate disks. The other principal reason of CCD imagers' relative immunity to FPN compared with CMOS counterparts is because the latter have charge-voltage conversion at every pixel whereas the former have a unique conversion node for all of its pixels. The topological distribution of the FPN noise is further analyzed in section 6.2.

2.2 Non Raster Scanning (Sampling) CMOS Imagers

It is rather interesting to examine the meaning development of the word “scan” [18]. The source of the word “scan” came from the Latin word “scandere” that means “to climb” and which is used in this sense in “scanning a verse of poetry”, because one could beat the rhythm by tapping one's foot. The Middle English verb “scannen” is, in this sense, derived from the Latin word “scandere”. Later in the 16th century other senses of “scan” have been developed towards the sense of “looking at searchingly” (first recoded in later 1798 [18]). With further developments, “scan” eventually broadened to include looking over a surface with or without close scrutiny of details reaching the modern usage of “scan” that means “look over quickly”. Therefore, “scan” includes both pattern and speed of the search in addition to resolution (scrutiny degree).

On the technical side, raster scanning was originally proposed for Cathode Ray Tube (CRT) displays using the electrical and magnetic fields influence on a finely focused accelerated electron beam “scanning” its fall on a phosphorescent screen (that emits light on the fall points) for information display purpose. The pattern of the raster scan looks exactly as the scan of poetry passages, namely line by line. The first cathode ray tube scanning device was invented by the German scientist Karl Ferdinand Braun in 1897. Braun introduced a CRT with a fluorescent screen, known as the cathode ray oscilloscope [19]. This method of display was later adopted to be the scanning scheme for early solid-state imagers [20] in order to maintain compatibility between acquisition and display systems in addition to its simplicity in storing and transmitting image data. The raster scan is also the scan of choice in image

readout at high speed rates since it minimizes discontinuities in reading the image rows, therefore increasing frame rate and thereby reducing motion artifacts such as motion blur. Recall that the integration time is the discharging period spent between the reset and the readout of pixel's photodiode. Consequently, and because high speed imaging is one of the merits of CMOS imagers, few attempts have been tried to deviate from this fundamental image scanning to a more complex scanning patterns. One of these attempts suggests scan patterns such as the Space Filling Curves (SFC) families [21] namely the Morton (Z) and Peano-Hilbert curves [22] for fast and efficient mean computation of 2x2, 4x4 kernels dedicated mainly for multiresolution imaging. The SFC scanning patterns are essentially fractal scan patterns that have the ability to scan the whole 2D image pixels (without missing one) and thus achieve the space filling property. Due to their inherently strong locality property[2], SFCs in general and Peano-Hilbert curves in particular are useful in exploiting the spatial coherence of nearby pixels which is very useful for image compression (lossless or lossy) [23][24] or pattern recognition [25]. To realize these applications (and others such as halftoning…etc), the image is scanned using an SFC scan generating a sequence of data which is transmitted through a communication channel before being processed by an application, as depicted in Fig 2.4. Using the same SFC scan map the resulting image is reconstructed.

[2] Due to the recursive nature of the SFC, neighbouring pixels are traversed before moving to more distant ones.

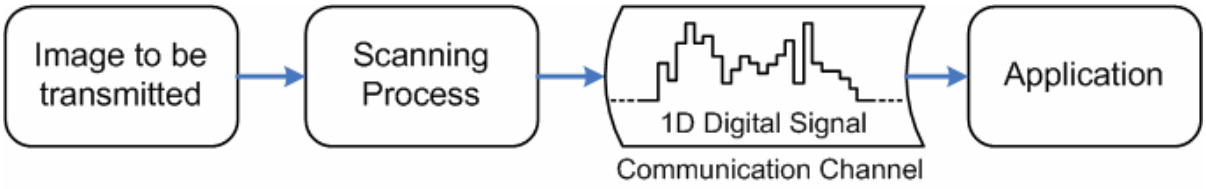


Fig 2.4 The scanning process framework

Thus far, SFC based scanning was considered mainly for the processing applications and the communication bandwidth needs, rather than image sampling requirements, which make these scanning techniques vulnerable to motion and lighting artefacts. Thus, the implementation in [22] will suffer greatly from these sampling conditions beside its intensive use of wiring which, in addition to its low fill factor (15%), will make the suggested CMOS imager a "light starved" design [26], mainly due to the vignetting phenomenon [27].

In reality, the initial attempt in the implementation of mechanisms allowing flexible scanning was suggested originally in [28] by using individual reset scheme as shown in Fig 2.5.

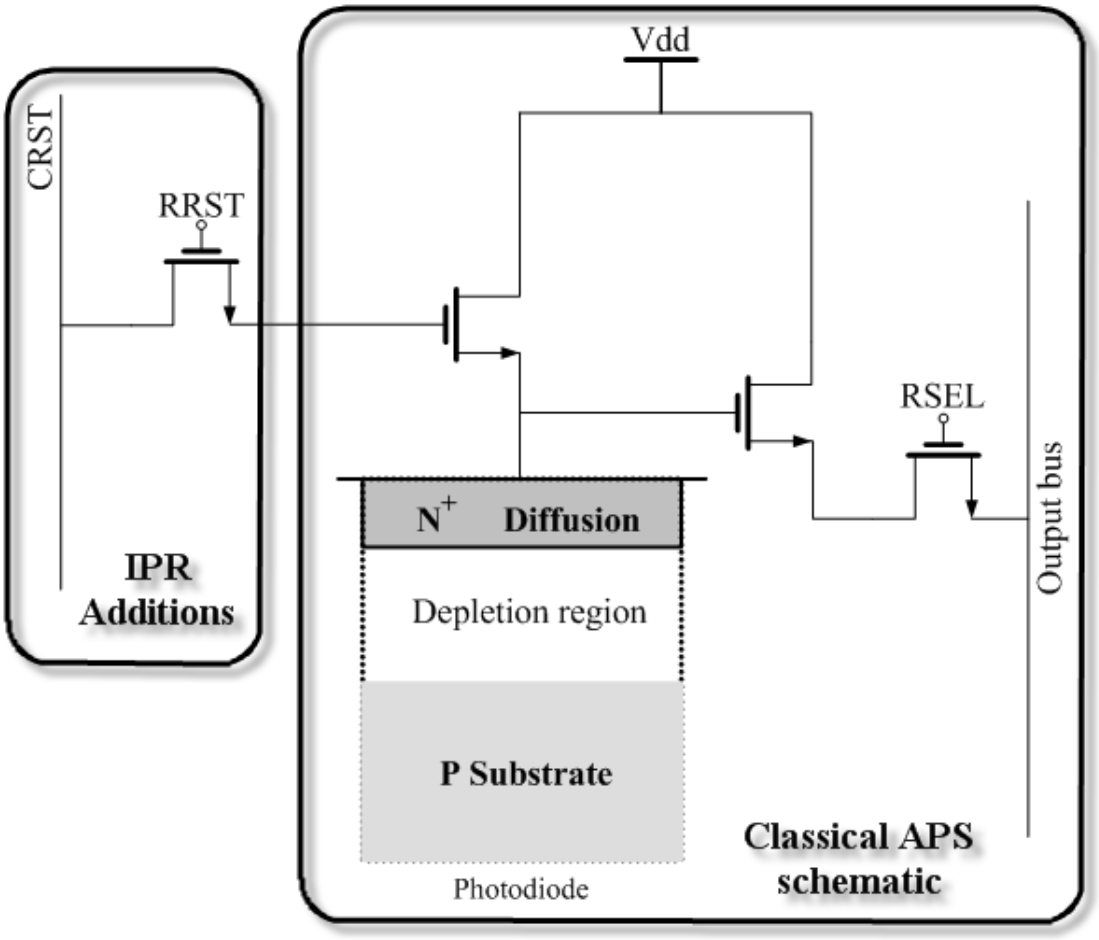


Fig 2.5 Individual pixel reset APS

Using vertical reset (CRST: column reset) and horizontal reset (RRST: row reset) signals the random reset is possible to implement. The main drawback of the approach is the V_{th} drop sensed at the gate of the APS pixel reset transistor, beside the fill factor decrease as a result of adding an extra transistor. The major drawback can be overcome by using hard reset technique in which the RRST gate voltage uses $V_{dd} + V_{th}$ instead of V_{dd} when enabled. One may suggest the use of PMOS reset transistor instead of NMOS. However, this technique

reduces the fill factor due to the n-well and larger size required for the reset PMOS transistor (due to lower hole mobility). Fully realising the random accessibility requires storing the sampled photo charge locally inside each pixel using a capacitor and a shutter (switch) between it and the photodiode [29]. Combining the two previous pixel architectures, any pixel can be reset individually as well as readout multiple times (without destroying the sampled image information). Again, the fill factor is an issue as the number of transistors inside each pixel increases. However, this may be solved in the future with CMOS technology scaling.

Another scanning approach suggested integrating motion detection circuitry beside ADC conversion at the pixel level in order to implement a quad-tree scan scheme. This avoids the redundant cycles[3] of raster scanning [30], thus minimizing energy consumption and increasing frame rate especially for high resolution imagers. Quad (means 4 members)-trees[4] [32] are most often used to partition a two dimensional space by recursively subdividing it into four quadrants, and it was primarily proposed for database searches [31]. This technique has been therefore borrowed from graph theory in order to replace raster scanning so that only active pixels (determined by the motion detection module inside the pixel) are read by jumping from one level to another searching for active nodes. The main drawbacks of this technique are the very large pixel pitch (96 μm) and the very low fill factor (3.2%). However,

---

[3] Redundancy in this sense refers to the sampling of pixels that do not change between frames (non active).

[4] Trees in this context are graph theory constructs and are graphs in which any two vertices are connected by exactly one path

the attempt target is definitely worth it in the future development of CMOS imagers especially with their inherent architectural flexibility advantage (over CCDs) especially with increasing image resolutions and high frame rate requirements.

2.3 Our Approach for Image Sampling

Visions systems found in mammals in general outperform by far silicon based image sensors, not only due to their analog yet stable and powerful signal processing cells at both the sensing site and the neuronal processing layers, but also due to their hierarchical organization. Many attempts have been made to mimic the image processing capabilities of the biological vision systems[5] [33], however, few attempts have been made to investigate the power and influence of sampling architectures on the vision performance. Furthermore, as CMOS imaging technology offers the ability to design complex architectures, as compared to CCDs, a small but increasing number of imagers are designed to test novel CMOS image sensor architectures. Our implementations described in this thesis are an example of this trend of CMOS imagers' development.

In designing novel CMOS imaging architectures, we focus on parallelism and symmetry in order to achieve high speed imaging without much expense in terms of hardware and power. Using this architectural view in image sampling we can implement our goal in building an optimized CMOS imaging solution that minimizes the sampled data while maximizing sampled image information. This is indeed one of the major goals for ever- increasing image

[5] Vision systems are imaging based applications encompassing image processing to mimic partly human vision.

resolution and high speed transmission needs for future CMOS imagers for a wide range of imaging applications from scientific remote imaging to military object tracking.

2.4 Biological Vision

In this section a qualitative description of the biological vision in general and human vision in particular will be presented. Our interest in this section is to highlight some of the powerful mechanisms characterising human vision making it a 'smart' vision. It is beyond the scope of this section to demonstrate all the vision capabilities of human visual system, but only selective properties will be shown as they may be mimicked in our CMOS imagers [34]. To narrow further the scope of human vision analysis, only the retina where the focal plan image is being spatially sampled is considered [35].

2.4.1 Spatial Sampling

The retina is the innermost layer of the human eye which contains light-sensitive photoreceptors and their associated neural tissue. There are two kinds of photoreceptors: rods (intensity-resolving) and cones (colour-resolving). Their distribution is further detailed in section 6.4.3 which principally shows the non-uniformity of the sampling architecture of the retina. Finally, it is worth mentioning the roughly circular distribution of the photoreceptors around the central region known as fovea which perfectly coincides with the optical axis of eye's lens [4].

The plexiform layer is an intermediate interconnection layer that connects between photocells and ganglion cells whose axons (neuronal communication extensions) make up the fibers of

the optic nerve before reaching the brain visual cortex. Without getting into details of this complex layer, this layer serve many visual functions especially visual resolution where it maps foveal cones (no rod cells are in the fovea) individually to ganglion cells whereas rods are grouped through specific neuron cells before being connected to ganglion cells. This explains why the fovea region of the retina has higher spatial resolution than its periphery.

Retinal *ganglion* cells are neurons responsible for converting chemical signals received from the photocells (cones and rods) into electrical firing pulses to be sent to the visual cortex at the brain visual cortex.

It is clearly visible from this short description that the human biological visual system, at least at the spatial sampling level, has great deal of architectural organization allowing it to maximize image sampling in one area (fovea) while keeping it low in other retinal areas. The second important point to notice is the interconnection between the sampling photocells through the *plexiform* [4] enabling the visual system to perform intelligent tasks such as extracting edges and reducing resolution. The first task is performed in the fovea region by the cone photocells and the second is realized by the peripheral rod photocells for motion detection [36]. Both of these two facts may have a great potential influence on future development of CMOS image sensors with their architectural flexibility. In fact in the present thesis the first observation mentioned herein was exploited to design a central foveated architecture called the pyramidal architecture while the second was mainly utilised in designing the multiresolution CMOS image sensor discussed further in Chapter 7.

2.5 Previous Foveated Vision CMOS imagers

Spatial resolution is directly related to the size of the pixel with respect to the system optics, which characterizes the detection of fine details of the sensed image, but often simply refers to the number of pixels of the acquisition or display device. However, the frame-scanning rate, data transfer bandwidth as well as power consumption and sensitivity are also key requirements for higher resolutions. Another factor, namely the limitation of the optical system, also plays an important role in determining the usefulness of high-resolution image sensors. It has been claimed that pixels of sizes much below 5μm pitch are not needed because of the diffraction limit of the camera lenses typically used in consumer cameras [37]. Furthermore, for a given technology (CMOS 0.35μm), it has been shown [38] that dynamic range and signal-to-noise ratio (SNR) degrades with the decrease of the pixel size in a square root form. On the other hand, smaller pixel sizes increase the spatial resolution by increasing the Nyquist-limit spatial frequency, defined as the half of reciprocal of the center-to-center pixel spacing. Besides, the Modulation Transfer Function MTF that characterizes the ability of the imaging system (CMOS imager in this case), to output the sharpest form of the acquired image[6], is higher for smaller pixel sizes. Consequently, the optimal pixel size is a trade-off between the above conflicting parameters and a given application requirement.

[6] The MTF function is defined as the normalized contrast amplitude response of the retransmission (imaging) system, as a function of the spatial frequency, below the Nyquist limit.

Other alternatives have been suggested to enhance the CMOS image sensor resolution by using different architectures while keeping the above-mentioned key-parameters in mind. These new architectures are developed in the following section.

2.5.1 Foveated CMOS Image Sensor

Biologically speaking, the fovea is the region of highest visual acuity in the human photo-transducing retina [4]. It contains no rod cells, which are more sensitive to low light intensities, but contains the highest concentration of cones, which detect colors. This explains why we cannot see very dim sources, such as weak starlight, when we look straight at them. The dim source only becomes visible when it is placed in the periphery and can be detected by the rods. Finally, it is worth mentioning that the dense representation of the foveal cones suggests that the spatial sampling of the cones must be an important aspect of the visual encoding [4]. This explains why humans vision declines (in spatial resolution power) away from the direction of gaze.

From the CMOS hardware implementation point-of-view, different strategies have been suggested in the literature [33]. Fig 2.6 shows some implemented CMOS image sensors adopting the spatial fovea-distributed pixel mapping such as linear-polar sensor (a), log-polar sensor (b) and log-polar with Cartesian center CMOS image sensor (c).

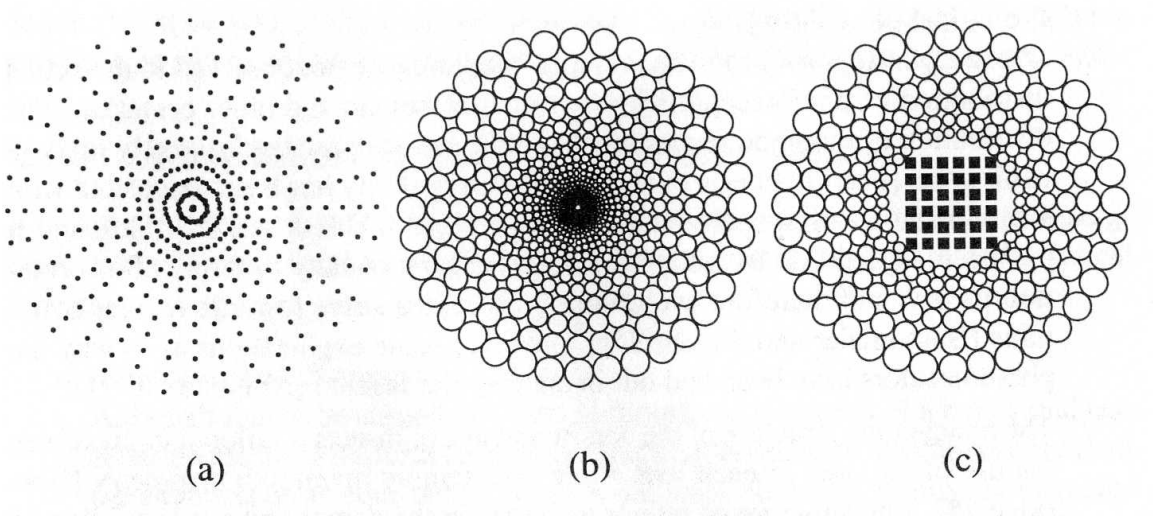


Fig 2.6 Polar foveated CMOS image sensors [33]

The above-mentioned image sensors are able to map the Cartesian two-dimensional image to polar (circular) (a, b), or mixed coordination (c), facilitating therefore some image processing operations such as scaling and rotation invariance assuming the origin at the center of the fovea. These processing operations are simply implemented with a simple shift operation in the angular axis or the radial axis respectively. Foveated imaging is achieved through the high sensor density near the central part of the image sensor, although in practice the number of pixels in fovea region rings is usually fewer than that of the outer rings due to the finite physical size limits. A significant problem with this type of foveated sensor was the difficulty of forming an acceptable colour image at the low pixel densities in the periphery. Another approach using rectilinear pixels for the implementation of foveated CMOS image sensor was reported recently [39] and is shown in Fig 2.7. It uses the standard CMOS process and

adopts the concept of photo-charge normalization in order to use the same charge amplifier with the different pixels sizes and therefore displays a wide dynamic range response to the incoming light. A drawback of this architecture is the complexity of the synchronization circuitry needed to scan the different image sensor rings.

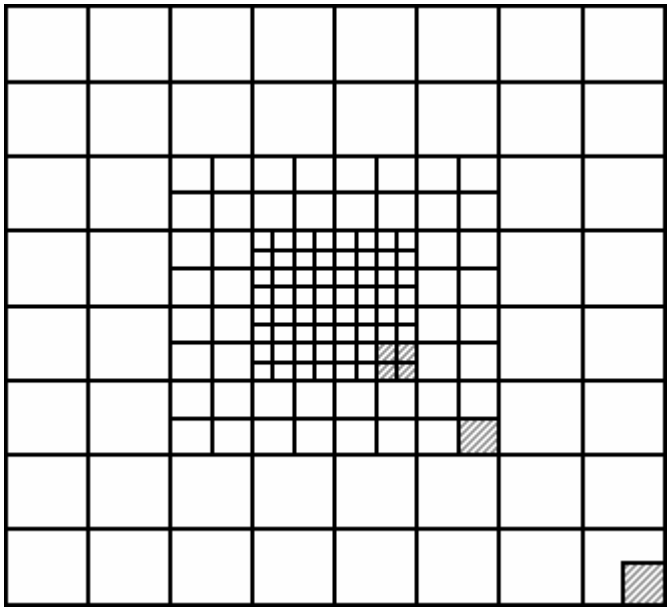


Fig 2.7 All CMOS rectilinear foveal image sensor

Interestingly enough, important properties of such biologically inspired image sensor, namely the fast frame (scanning), wide field of view (FOV) and high resolution continue to drive interest on this kind of image sensors in low-vision enhancement, communication and target tracking applications. The product of these parameters defines the so-called metric of visual

information acquisition power. In a foveal system, the metric is computed as the product of the peripheral field of view, the spatial resolution at the fovea, and the overall frame rate.

2.5.2 Multi-Resolution CMOS Image sensors

Resolution was one of the most important competing factors between CCD technology and CMOS technology for the development of high definition image sensors. Although the CCD image sensor has attained higher resolutions, the present and future development of CMOS technology is leading to smaller and smaller pixel sizes, potentially overcoming the CCD lead. Furthermore, the CMOS ability to add processing elements at the pixel level, column level and at chip level, is adding an impetus in the race to higher imaging performances. Benefiting from this capability, CMOS image sensor can obtain multiple resolutions, and together with their addressability, an electronic zoom capability is achieved [40]. The suggested pixel neighbourhood averaging scheme in [40] (in row and column) is realised at the column level. This averaging mechanism is implemented by using switch capacitors and by using shift registers to control the required resolution i.e. number or averaged rows and columns as shown by Fig 2.8. A dual approach to the previous multiresolution scheme is suggested but, instead of voltage averaging, a current averaging strategy has been adopted [41] and a new scheme for current-mode Correlated Double Sampling (CDS) circuitry is proposed. The applications of multiple resolution image sensors are very wide especially in real world systems that impose some constraints such as format choices, processing speed, and bandwidth such as data reduction, robotics and target tracking.

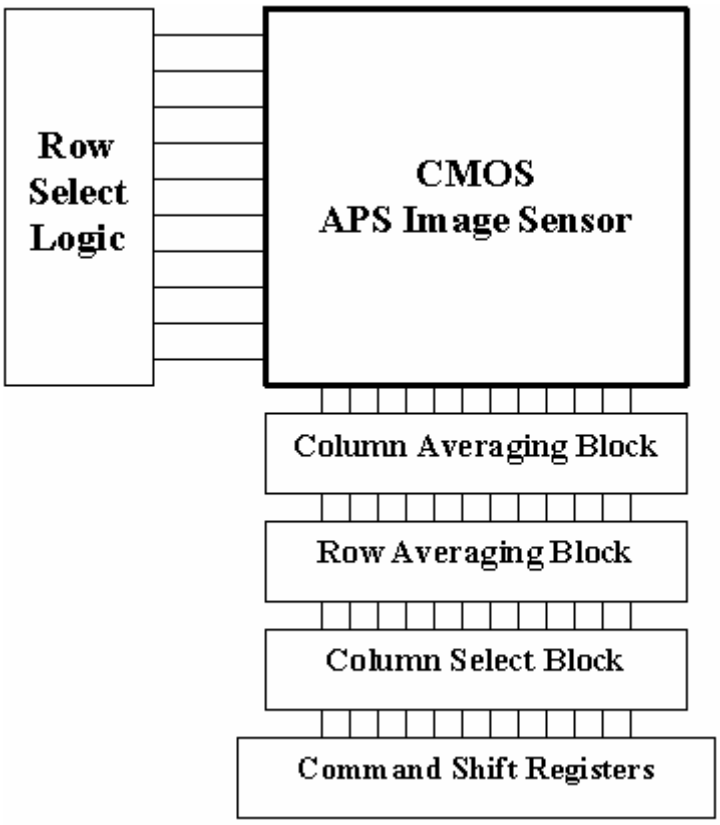


Fig 2.8 Programmable multiresolution CMOS active pixel sensor architecture

2.6 Dynamic Range Enhancement Techniques

Dynamic range (DR) is a term widely used in many fields to describe the ratio of the largest measurable or detectable value (of interest) to the smallest value. In imaging systems, the physical measure of interest is obviously light intensity of the sampled image. In CMOS or CCD image sensors, dynamic range is one of their most significant characteristics because it describes their ability to sample bright and dim areas of a sampled scene within one frame. The minimum detectable light intensity is the intensity that would create an electrical signal

enough to equal the noise floor of the imaging device. Another definition of dynamic range is ratio of the voltage saturation level to the readout noise. Because it uses electrical parameters it is referred as the electrical dynamic range and given that the (optical) dynamic range is measured in the linear response of the CMOS image sensors, the electrical dynamic range is equal to the optical dynamic range (although this distinction is important for non-linear devices, such as the logarithmic pixel mentioned in section 2.6.1).

While natural scenes have a very wide dynamic range from illuminations of 10^{-3} lux[7] for night vision to illuminations of 10^{5} lux for bright sunlight (dynamic range of about 10^{8} or 160dB), typical CMOS or CCD image sensors have dynamic range of 65-75dB. Two alternatives are possible to improve DR, either by reducing the noise level of the imager and expanding its dynamic range towards darker illuminations, or by extending the saturation level of the incident light and thus improving the DR towards brighter scenes. With the exception of very few applications such as astronomy, most frequent imaging situations correspond to bright light imaging such as automotive night vision or objects tracking in an uncontrolled lighting condition such as star tracking and aircraft landing. This explains why most of the DR enhancement techniques are mainly for bright light imaging applications. Another reason for this trend of DR enhancement could be due to the fact the noise floor of the imager is an ultimate limit whereas the saturation level could be expanded to a accommodate non-limited bright lighting imaging conditions. Many DR enhancement

[7] Lux is an SI illumination unit that is equal to one lumen per square meter. Lumen is an SI unit for luminous flux equal to the light emitted in a unit solid angle by a uniform point source of one candle intensity.

approaches have been suggested to expand the limited electrical dynamic range of CMOS (or CCD) to reach the large optical dynamic range (of natural scenes) and at several levels of the imager design and functionality. An extensive review of these attempts is provided in [46]. We will concentrate on CMOS imagers dynamic range enhancement techniques which can be clustered into two categories:

DR enhancement by compressing or transforming the pixel photo-response.

DR enhancement by manipulating globally or locally pixel integration times. Even further development of these approaches includes external or internal (autonomous) control of the integration time.

2.6.1 Dynamic Range Enhancement by Photo-Response Compression

In the first category, the pixel response is changed from integration mode to continuous mode and from linear to logarithmic [42] [43] in a technique originally used in CCDs [44]. This is realised by connecting an active resistor (diode connected MOSFET), that is biased to behave like a continuously–conducting logarithmic resistor, with a photodiode as shown in Fig 2.9.

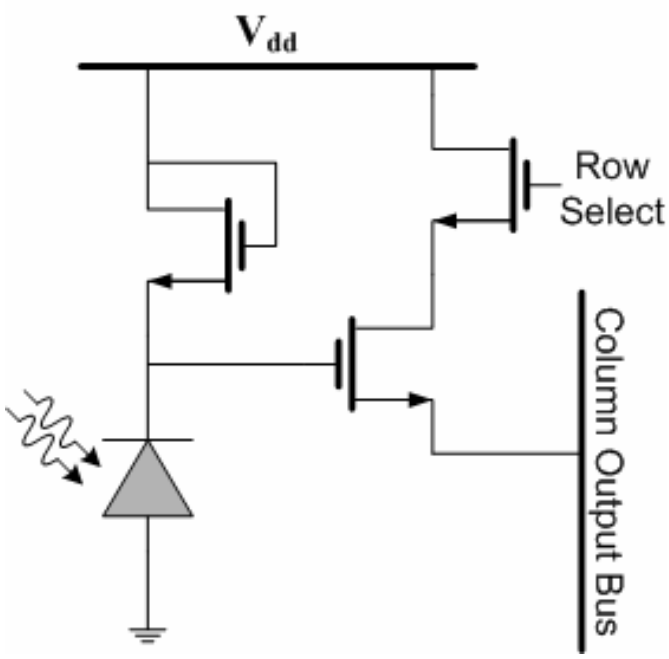


Fig 2.9 Logarithmic active pixel sensor

Because the NMOS resistor is biased in weak inversion, the NMOS resistor will be biased in the subthreshold regime that exhibits a logarithmic behaviour between its source current (photo current) and the photodiode voltage as shown by the following equation.

$$V = V_{dark} \ln\left(I_{sat} / I_{ph}\right) \ldots (2.1)$$

where V_{dark} is the photodiode dark voltage, I_{sat} is the leakage current and I_{ph} the photocurrent. Although the main reason in using MOSFET in resistor configuration was to have a large resistor between the photodiode and Vdd due to the small photocurrent (fA to nA), the logarithmic behaviour of the pixel looks very similar to human vision response to incident light [45]. The logarithmic pixel suffers mainly from the reduced image contrast.

Another biologically analogous alternative is to convert the photo-signal of the pixel in modulated pulses [48] [49] similar to ganglion cells [4] neural (spike signals) response. The basis of this light-modulated pulse generation is the conversion of the linear photo discharging mechanism of the APS photodiode into digital pulses which carry out the light intensity information in their pulse width. This means converting the light intensity into time domain pulse signal [8] [47] and in order to realise this functionality, a digital feedback is realised between the photo diode discharging node and the reset transistor as shown in Fig 2.10. The pixel scheme in Fig 2.10.a is used in [48] while Fig 2.10.b is used by [49]. Using either an external or an internal threshold voltage, a comparison is performed in the digital feedback block and if the photodiode voltage drops below the threshold a reset signal is applied on the reset MOSFET. Therefore, the higher the light intensity, the faster the photodiode discharge will be and thus the shorter the generated pulse is going to be, which demonstrates the light modulation of generated pulse. Despite the high dynamic range achieved by this technique (120dB [49]), its main disadvantage is the multiplicative noise caused by the feedback path. This can be solved by implementing in-pixel noise cancellation circuitry, further reducing the pixel fill factor. This is probably the reason why there is no feedback path between ganglion cells and retinal photocells [4] in the retina.

---

[8] Neural information processing is mainly in time domain because neural responses are discrete in their amplitudes but analog in time. Thus, time plays an essential role in the flow and transformation of information in biological systems.

In conclusion of this description of these different alternatives of dynamic range enhancement, it is apparent that manipulating the pixel photo-response to enhance DR suffers mainly from the reduced fill factor, degradation of sampled image contrast and the non-linear increase of fixed pattern noise due to the feedback loop with the photodiode [46]. The nonlinearity of FPN noise implies the non applicability of CDS unless it is implemented in pixel before light-to-pulse conversion which worsen further fill factor. These factors explain why integration time techniques are the dominant players in DR enhancement of CMOS imaging.

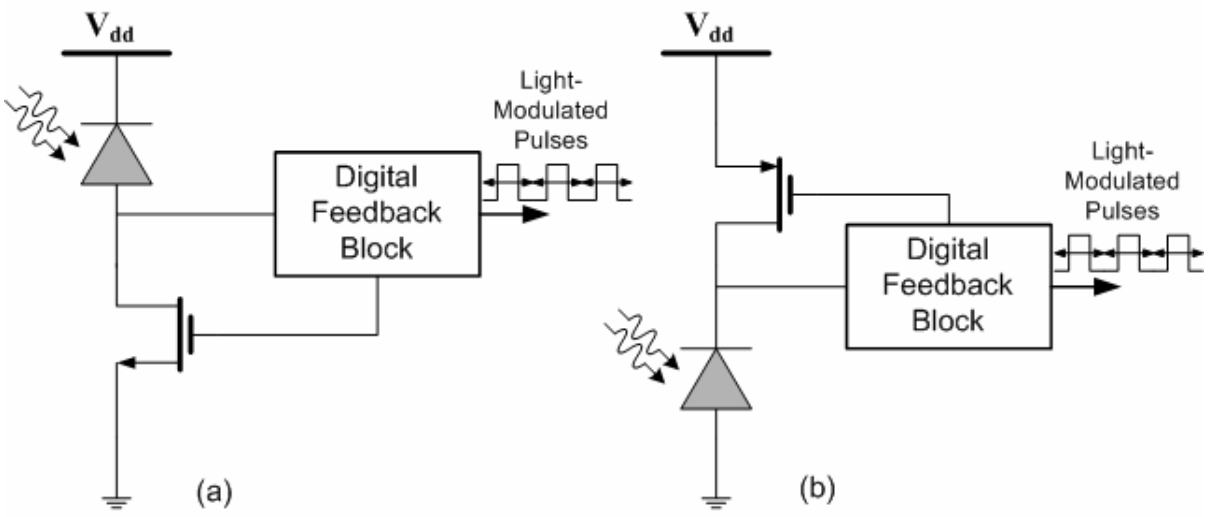


Fig 2.10 Light to pulses conversion forms

2.6.2 Dynamic Range Enhancement using Integration Time Control

In the second category of DR enhancement, the integration time of the imager is programmed globally or locally in order to resolve the scene's bright and dark spots. The importance of using the integration time as a tool to enhance the optical dynamic range, in integration APS sensors, come from its equal effect with light intensity on the photo signal as expressed in equation (2.2).

$$V_{photo-signal} = Ss\ L_{inc}\ T^{\text{int}} \ldots (2.2)$$

where Ss is the pixel constant sensitivity, L_{inc} is the light intensity on the pixel and T^{int} is its integration time. Therefore, to get the same electrical photo signal, in the APS linear regime, we need have the same product $L_{inc}\ T^{int}$ and thus short integration time is needed for high light intensities and long integration time is needed for low light intensities. This simple demonstration explains the use of integration time in extending the optical DR.

Integration time control has been initially implemented through global control of the frame time using an electronic shutter replacing the mechanical iris used to control incident light intensity in old film imaging cameras [50][51]. A schematic of a typical electronic shutter, which will also be used (with minor changes) in our multiresolution CMOS imager in Chapter 7, is shown in Fig 2.11. The only difference between APS with electronic shutter and the standard APS sensor is the addition of a shutter transistor between the photodiode and source follower (SF) and the extra reset transistor (reset). The capacitance of the drain diffusions from the reset and the shutter transistor beside the gate capacitance of SF constitutes a storage node (that needs to be shielded from light) where the sampled signal

charge will be temporarily stored. The global shutter signal is shared for all the CMOS imager pixels and thus comes the global attribute of the shutter. Finally, reset and timer transistors are used to reset the storage node and the photodiode respectively.

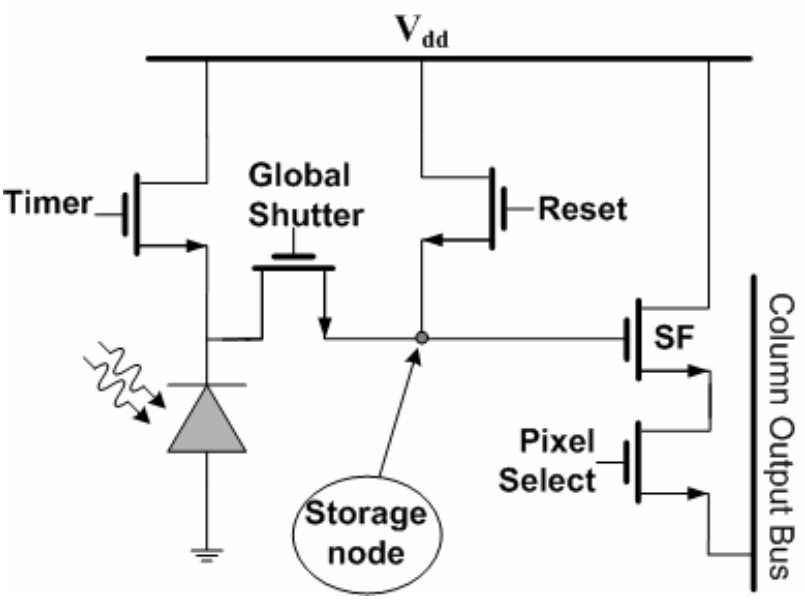


Fig 2.11 Global shutter general schematic

The main disadvantage of using of using the global shutter to control integration time emerges in high intrascene dynamic range situations where only bright regions can be visualised with shorter integration times and the darker ones using longer integration times but not both at once. Therefore, the dynamic range of the imager is limited by the electrical dynamic range (~60-75dB) which is substantially less than the optical dynamic range of many natural scenes. The most important application of this structure is found in fast imaging

applications with intermediate (controlled) light intensities [52] because of its minimization of motion blur [53].

The local integration time technique is based on estimating the integration time independently at every pixel in the image sensor [54] [55]. This technique, called multiple integration time (MIT), is implemented through the automatic selection of integration time (from among a predetermined set of values) corresponding to the closest level to saturation. This technique is sometimes referred as the time-to-saturation technique. The MIT pixel shown in Fig 2.12 is composed of 2 identical photodiodes, a transparent latch, an inverter, 2 NMOS, 5 PMOS transistors and two capacitors.

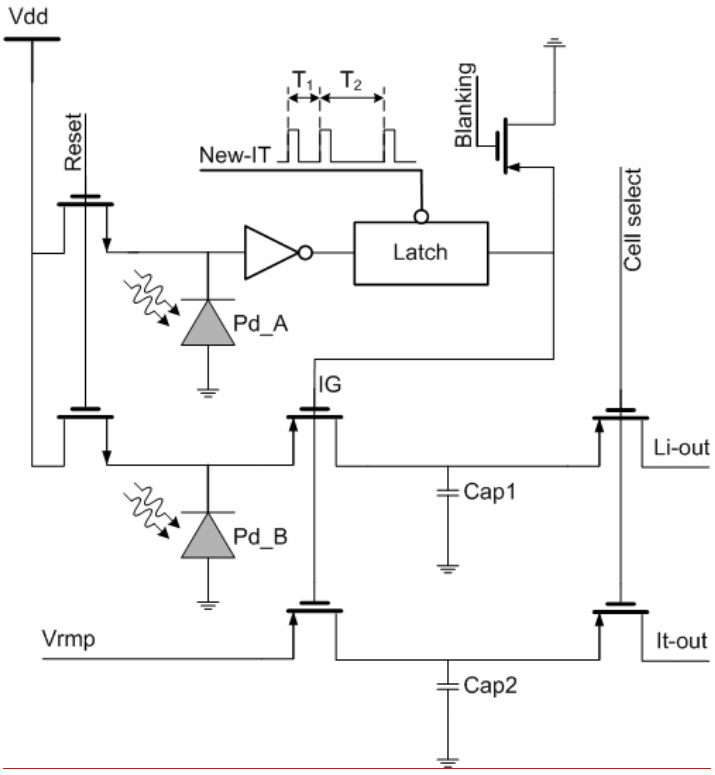


Fig 2.12 Pixel schematic of MIT photoreceptor

Photodiodes Pd_A and Pd_B are initially reset concurrently and left discharged by the impinging light intensity. Pd_A output discharges continuously and is compared to the inverter's threshold whose output will be latched out at specific time intervals corresponding to the predetermined integration times T_1, T_2... T_N. Once the photodiode output crosses the threshold the inverter output is latched out and the integration gate (IG) is closed sampling the photo charge in Cap1 and the corresponding incremental voltage Vrmp, which indicates which integration time interval is being used, at Cap2. These two stored voltages will be output to Li-out and It-out that carry photo signal and the corresponding integration time interval respectively. Although this technique can achieve DR of 120dB, its large pixel size of 110 μm pitch reduces the acquisition of high spatial frequency details (resolution) with a reduced fill factor. Autonomous control is mentioned in [55] but again, it suffers from a low fill factor and a large pixel area.

2.6.3 Dynamic Range Enhancement using Variable Light Exposures

One may alleviate the integration of the control (of integration time) from the pixel site to its neighbourhood so that instead of having one single-sensitivity pixel, four (or more) pixels of different sensitivities are integrated. Therefore, the pixel of highest sensitivity would integrate better low light intensities and the pixel of the lowest sensitivity would integrate better the high light intensities. However, the sensitivity is a technology constant parameter thus one way to mimic this characteristic is by controlling the exposure [57] of the pixels and to keep all pixels similar. The exposure is varied over a cluster (of pixels) and this cluster structure is mapped over the whole imager as shown in Fig 2.13, where the cluster's pixels

have their exposures so that $Ex_0 < Ex_1 < Ex_2 < Ex_3$. In [57], the exposures have been assumed to be uniformly increasing values such that $Ex_0 = K\,Ex_1 = K^2\,Ex_2 = K^3\,Ex_3$, and K was set to $K = 4$. The exposure control is realised by masking the pixel with cells of different optical transparencies or directly etching over the pixel on a solid-state imager (CCD or CMOS imager). The idea of using exposure control to extend the dynamic range comes from film imaging technology by using the important concept known as the reciprocity principle [58]. The reciprocity law states that within the normal range[9] (of intensity and time for the film) different choices of aperture and shutter speed that result in identical exposure are equivalent. By definition, exposure = light Intensity x integration time, which explains that for the same integration time, opacity is reversely proportional to exposure.

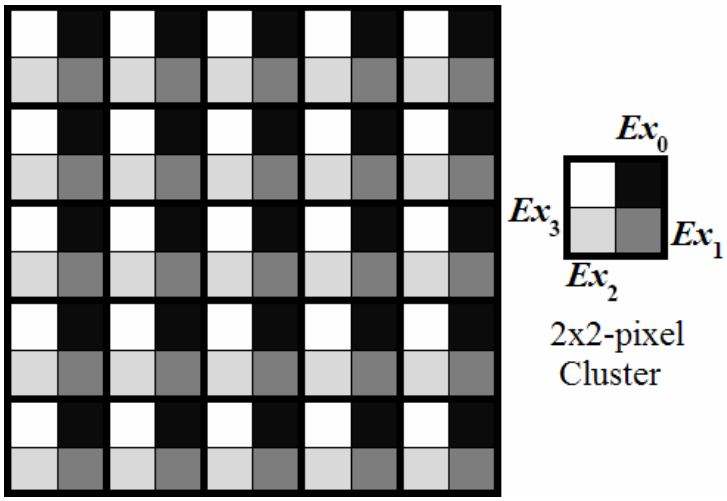


Fig 2.13 Spatially varying exposure time technique for DR enhancement

[9] Normal range refers to the linear range of the film photo-response.

By using similar integration time for all pixels to sample an image, a different exposure value will result in each pixel of the cluster, hence resulting in this technique being known as the spatially varying exposure technique (SVE) [59]. The DR expansion emerges from the assumption that a multiple-exposure cluster shares the same light intensity[10] and thus a low light intensity would be better imaged with a higher exposure pixel and a higher light intensity is well sampled with a lower exposure pixel, thus expanding the original pixel optical DR. The importance of the SVE technique in enhancing the imager DR resides in manipulating the exposure of the imager pixels either by using different microlenses on the array, or using different integration times for different pixels, or embedding pixels of different potential well apertures. All these alternatives (we will review some of them subsequently) fall under the spatially varying exposures technique for DR enhancement. The key feature of the present SVE technique is the simultaneity of spatial dimensions sampling and exposure dimensions sampling (through 4 or more exposures per cluster). Two important questions arise: how much DR enhancement is achieved, and how is the final image reconstructed, using the SVE technique?

The DR definition is

$$DR = 20 \log\left(\frac{I_{\max}}{I_{\min}}\right) \ldots (2.3)$$

where I_{max} and I_{min} correspond to the maximum and minimum gray levels respectively. I_{max} and I_{min} correspond respectively to the full-well capacity (saturation) and the minimum signal

[10] True only in the case of an image with low spatial frequency *vis-à-vis* the imager spatial resolution.

(or read noise) detectable by the imager [60]. This correspondence is made by the adjustment of the gain of the analog-to-digital converter ADC and the number of required image gray levels, which has been found to be 8 levels [62] for false-contour-free images [61]. In general the DR of an imager is equal to number of gray levels necessary to encode the imager analog output [16]. Therefore, an 8-bit CCD or CMOS imager has DR equal to

$$DR_{8-bit\ imager} = 20\log(255) = 48.13\ dB \ldots (2.4)$$

where, I_{min} is set to 1 as level I_{min}=0 represents a non meaningful information.

Now, using the spatially varying exposure scheme in image sampling will maintain the lowest level at 1 but will extend the maximum detectable light intensity level (after ADC conversion) to I_{max} (Ex_{max}/Ex_{min}) consequently expanding the DR of the resulting image to

$$DR = 20\log\left(\frac{I_{\max}}{I_{\min}}\frac{Ex_{\max}}{Ex_{\min}}\right) \ldots (2.5)$$

The extension of I_{max} is due to the fact that the saturation of the lowest exposure pixel (the most opaque) requires higher saturating light intensity equal to I_{max} (Ex_{max}/Ex_{min}). This is clarified by the following the mathematical proof where *Ss* is the pixel sensitivity, *Vs* the pixel photo signal and $Ex = Tr\ L_{inc}\ T^{int}$ its exposure, where *Tr* is the transmission factor of the pixel cell mask, L_{inc} is the incident light intensity and T^{int} the pixel integration time. Recall also that all pixels are using the same T^{int} and the pixels cluster is postulated to have the same L_{inc}.

$$Vs = Ss\ Ex \text{ or } Vs = Ss\ Tr\ L_{inc}\ T^{\text{int}} \ldots (2.6)$$

At the saturation level the saturation photo signal *Vs_sat* will be

$$Vs\_sat = Ss\ Tr_{\max}\ L_{inc\_\max}\ T^{\text{int}} \ldots (2.7)$$

where Tr_{max} correspond to the most transparent mask transmission rate, and L_{inc_max} is the maximum light intensity causing saturation. However, the product of $Tr_{max}\ L_{inc\_max}$ can be achieved, in similar way through reciprocity principle, by the product $Tr_{min}\ L_{inc\_max\_ext}$ where $L_{inc\_max\_ext}$ is the new (extended) saturation light intensity and Tr_{min} is minimal transmission rate corresponding to the most opaque mask. Therefore, we get

$$L_{inc\_\max\_ext} = L_{inc\_\max}\ \frac{Tr_{\max}}{Tr_{\min}} \text{ or } L_{inc\_\max\_ext} = L_{inc\_\max}\ \frac{Ex_{\max}}{Ex_{\min}} \ldots (2.8)$$

It is assumed that the cluster pixels share the same incident light intensity and integration time. This means the maximum light intensity (saturating the most transparent pixel) has been extended by a factor of Ex_{max}/Ex_{min} (saturating the most opaque pixel). Therefore, and assuming the minimum light intensity remains the same (read noise), the resulting DR becomes

$$DR = 20\log\left(\frac{I_{\max}}{I_{\min}}\ \frac{Ex_{\max}}{Ex_{\min}}\right) \text{or } DR = 20\log\left(\frac{I_{\max}}{I_{\min}}\right) + 20\log\left(\frac{Ex_{\max}}{Ex_{\min}}\right) \ldots (2.9)$$

Hence, using variable exposure technique enhances the uniform exposure imager DR by a factor of $20\log\left(\frac{Ex_{\max}}{Ex_{\min}}\right)$ and with the present case where $Ex_{max} = K^3\ Ex_{min}$, the final dynamic range enhancement becomes

$$DR_{enhancement} = 20\log\left(K^3\right) \text{ or simpler } DR_{enhancement} = 60\log\left(K\right) \ldots (2.10)$$

Now, the question of how to reconstruct the multi-expose (or SVE) sampled image arises. To answer this question let us examine the quantization of the SVE discussed above through the examination of Fig 2.14 which represents the quantization of the different exposures.

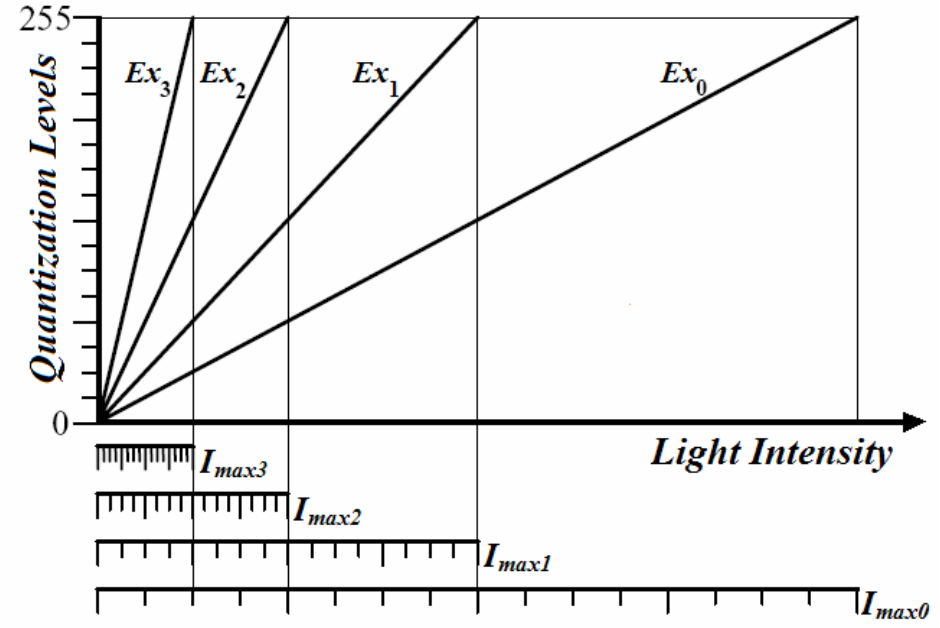


Fig 2.14 The quantization in SVE imager

Uniformly quantizing the different exposures of the SVE cluster pixel (composed of 4 different exposures) to the same ADC resolution leads to non-uniform quantization of the scene radiance. This non-uniform quantization of the scene radiance, apparent from the difference in the increasing quantization radiance steps of the different exposures at the X-axis, is advantageous as it represents a judicious allocation of resources, namely data bits [63]. On the other hand, the contrast of the image is not going to change (in the resulting

image) because its definition as the difference of brightness over its average in the Michelson definition [64]. The total number of gray levels (*GL*) in the case of a uniform scene quantization is $q\ Ex_{max}/Ex_{min}$, where q is the number of quantization levels in the pixel the highest exposure (e_0), which is the maximum number of quantization levels coinciding with that of the pixel of the lowest exposure (e_3). The total number *GL* of gray levels from the above uniform quantization is less than $q\ Ex_{max}/Ex_{min}$ (due to the overlapping levels) and can be evaluated from the following equation assuming K is the number of exposures in a cluster[11].

$$GL = q + \sum_{k=1}^{K-1} Round\left[(q-1)\left(1 - \frac{Ex_k}{Ex_{k-1}}\right)\right] \quad \text{... (2.11)}$$

Round is a function that rounds-off its argument to the closest integer. Applying the above formula on the example of Fig 2.13 (4 exposures with 8-bit quantization) will result 829 gray levels far from the original 256 original levels but also far less than the uniform scene quantization of 256 x 64 or 16384 gray levels.

Constructing the image from a SVE image sensor needs a mathematical computation to normalize or to interpolate the clusters of pixels and generate a uniform high DR image. Towards this end two approaches can be used. One approach involves aggregating the pixels of each cluster by averaging their response, the second involves ignoring the saturated and noisy response of low intensity pixels and interpolating the remaining pixels to estimate the resulting image. We will discuss the aggregation method because of it simplicity (despite its

[11] Which interestingly coincides in the present case with ratio between subsequent exposure (ex. $Ex0 = K\ Ex1$)

main limitation of decreasing the spatial frequency of the integrated image). The second method is further examined in [59].

In the aggregation method, the resulting image is constructed by convolving the captured image with a 2x2 box filter, which yields the average response of the cluster of four pixels. Because the assumption of uniform radiation impinging the cluster of pixel adopted, therefore the cluster average can be assigned to a single pixel in the resulting image. This will lead to a halving of the width and height of the generated image compared to the sampled original image. If, instead of averaging, the convolution is made, then the image size will be preserved but as the convolving kernel (box filter) passes over the sampled image, it will always find all exposures at every computation. However, this will reduce the contrast at the edges within the image. Finally, recall that averaging computation involves the summation of the cluster's pixels responses (which were assumed to share the same light intensity), the cluster response will be linear piece-wise function of this summation. The cluster resulting response function is shown in Fig 2.16. The form of this function is very similar to a gamma correction function with gamma greater than 1.

We have further analysed the piece-wise function of Fig 2.16 and we have demonstrated that in fact the constructed function is an exact gamma correction function with $\gamma = K$ (the ratio of between two adjacent exposures Ex_{k-1}/Ex_k).

To prove this result, note first that, assuming $GC(x)$ the gamma correction function;

$$GC(x) = \alpha\, x^{1/\gamma} \text{ , thus } \frac{\partial GC(x)}{\partial x} = \frac{GC(x)}{\gamma\, x} \quad \text{... (2.12)}$$

Normalizing the exposures maximum intensities by the smallest one (of the highest exposure) and calculating the slopes of the generated piece-wise function we get the slopes as shown in Fig 2.16. In other words, by assuming that I_{max3}=1 the slopes shown in Fig 2.16 were calculated at the horizontal indices x_1=1 (at I_{max3}), x_2=K (at I_{max2}) and x_3=K^2 (at I_{max1}) whereas the vertical indices were recalled from the maximum gray value of 255 as shown in Fig 2.14.

Table 2-1 shows the different slopes calculated graphically and their corresponding values extracted from equation 2.12 assuming $\gamma = K$.

Table 2-1 Graphical and functional slopes of SVE gamma correction aspect

HORIZONTAL INDEX (*X*-AXIS)	GRAPHICAL SLOPES	GAMMA CORRECTION FUNCTION SLOPES
1 (at I_{max3})	$255/K$	α/K
K (at I_{max2})	$255/K^2$	$\alpha K^{1/K}/K^2$
K^2 (at I_{max1})	**$255/K^3$**	$\alpha K^{2/K}/K^3$

The graphical slopes are in perfect fit in their denominators as their function-calculated counter parts. However, they differ in their nominator values. In order to match the graphical and function-calculated nominators, the value of α has to be set equal to 255 (at I_{max3}), $255/K^{1/K}$ (at I_{max2}) and to $255/K^{2/K}$ (at I_{max1}). The practical way would be the reverse where the maximum quantization level would be set to be equal to 255 (at I_{max3}), 255 $K^{1/K}$ (at I_{max2}) and

to finally to 255 $K^{2/K}$ (at I_{max1}). Thus applying the derivative of the gamma correction function shown earlier give us the same slopes shown in Fig 2.16 with the conclusion that the resulting piece-wise function is that of a gamma correction function with $\gamma = K$. Fig 2.15 shows the change of the maximum quantization levels versus K.

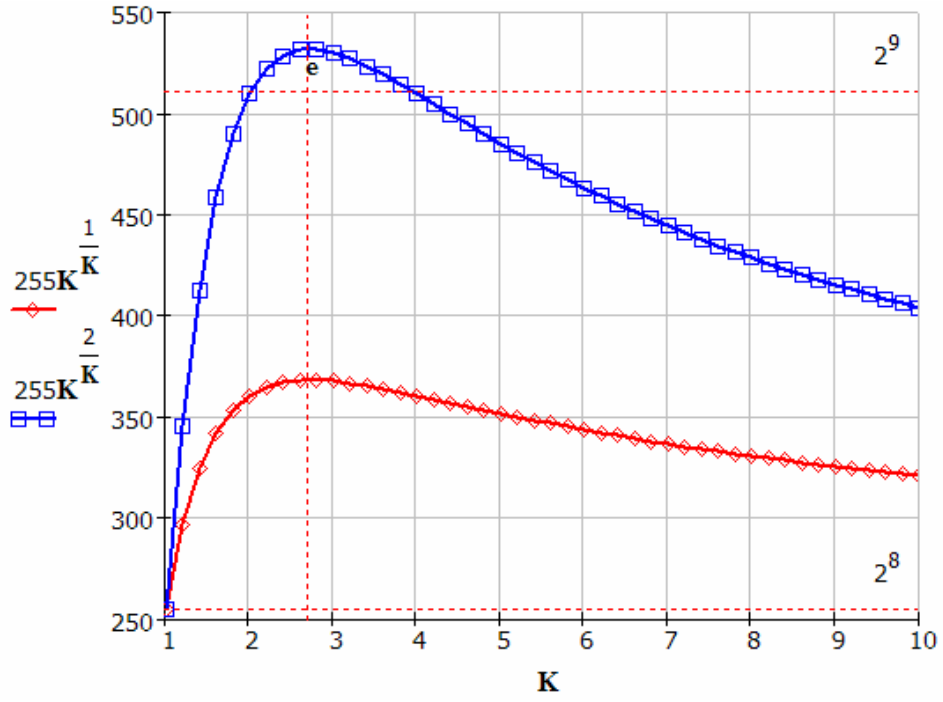


Fig 2.15 Quantization level adjustment for the SVE exact Gamma correction

Consequently, only one extra bit is necessary in the ADC converter to correct the Gamma correction aspect of the SVE technique except when $2 \leq K \leq 4$.

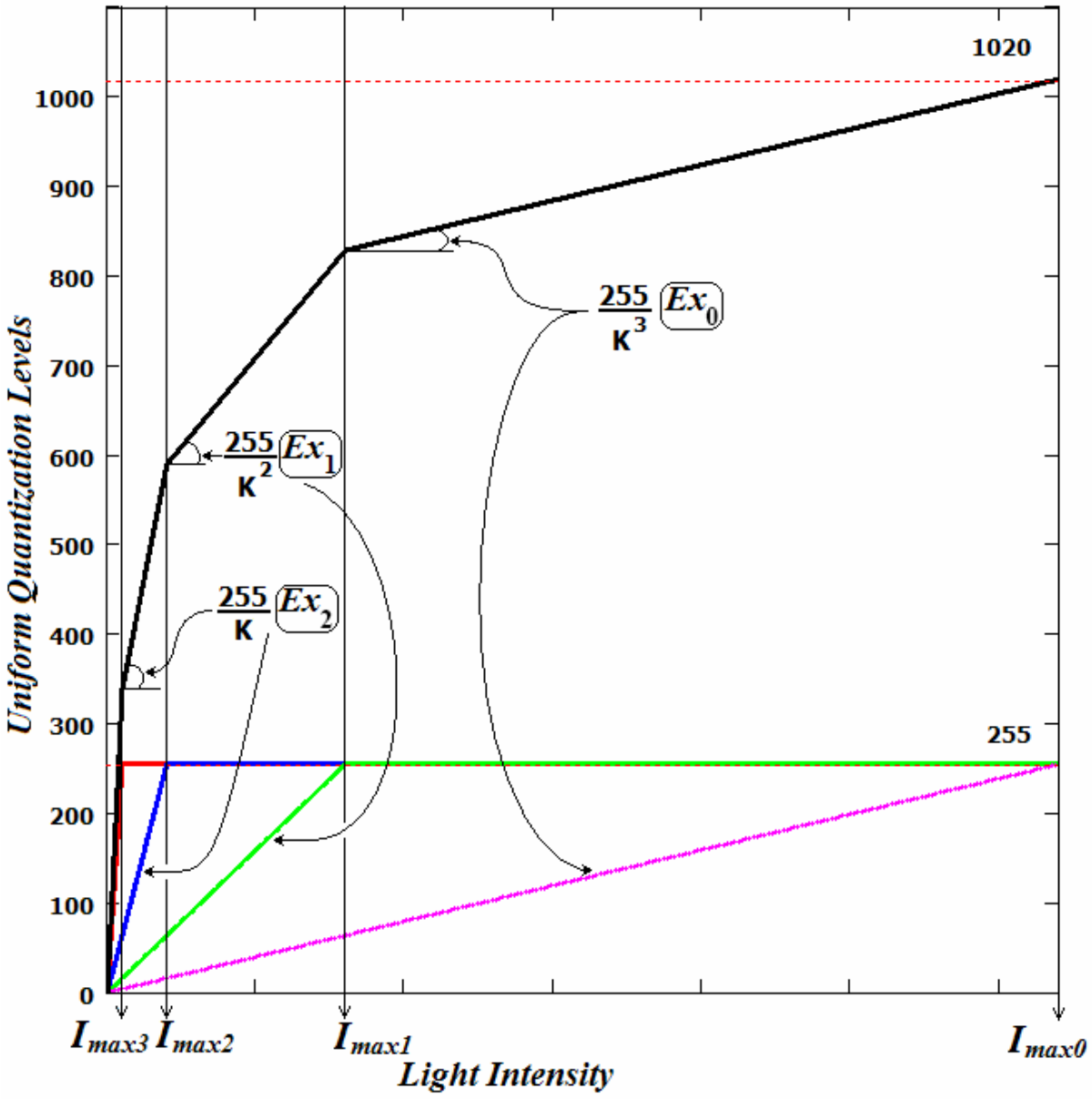


Fig 2.16 SVE cluster aggregation photo-response of the local brightness

We have found indeed an important result that implies the following. It is known that monitors and printers (and some scanners) suffer from the non-linear transduction of the

analog (or digital) video (or still image) signal into luminous (monitors) signal due to physical limitations [65]. This deviation from linearity was found to follow the profile of a gamma function and hence to linearize this response a gamma correction is needed [66]. This is done by reversing the power of the gamma function and by applying the correction to the video signal prior to displaying it thereby achieving or at least approaching linearity. The correction usually is applied at the video signal source to avoid processing bottleneck at the reception site. Therefore, what we found out is that we had achieved two goals with the SVE technique. We achieved the enhancement in the dynamic range that follows logarithmically the value of K (that is >1) and simultaneously we gamma-correct the constructed image to a power of K. In fact the value of gamma (or K) should be dependant on the display environment.

This demonstrates the importance of the exposure ratio K which plays simultaneously two roles; first, the logarithmic increase in the dynamic range enhancement and second, the gamma correction of the constructed image. The final point is even further important because it translates the gamma-correction from off the imager right to the image sampling chip increasing further the functional integration capability that is one of main advantages of CMOS imagers over CCD counterparts.

2.6.4 Dynamic Range Enhancement using Multiple Sampling

To conclude this review of dynamic range enhancement techniques, two more approaches need to be mentioned. The first approach is called dual sampling technique [68] and the second is the multiple capture [69]. Both techniques use varying exposures to extend the DR

of CMOS imagers but instead of using light, they use time, which plays a similar role in the exposure formulation. Only the dual sampling technique will be discusses in detail subsequently.

The dual sampling technique uses standard CMOS technology with the standard APS pixel structure and its architecture schematic is shown in Fig 2.17.

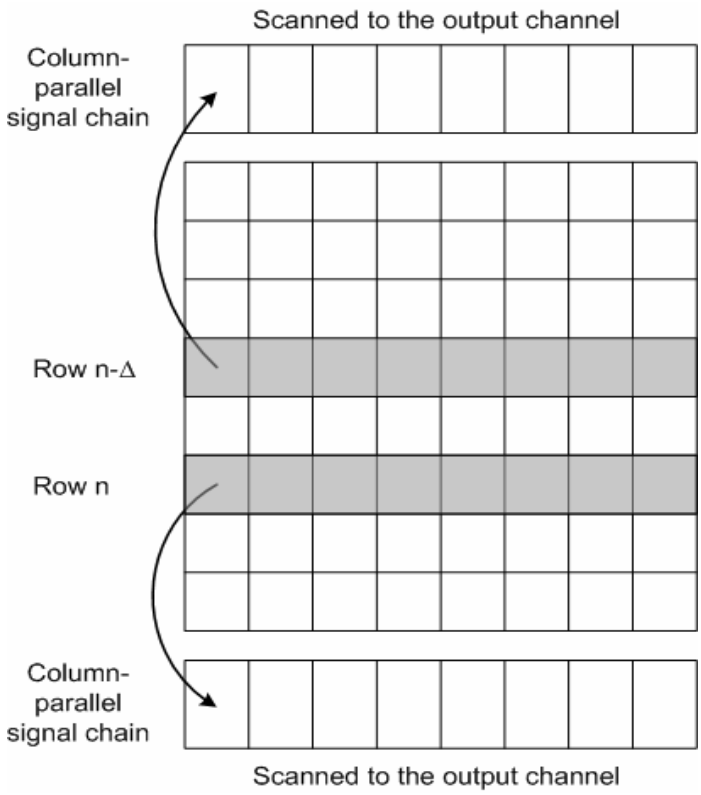


Fig 2.17 Architecture schematic of the dual sampling technique

The image sampling is made by rolling two busses of read and reset signals over the imager from the top of the imager to its bottom and back thus sampling two rows simultaneously.

The control bus (encompassing read and reset) signal at row (n) samples the photo voltages of the row into sample and hold capacitors at the bottom of the imager before resetting it and moving to the next row (n+1). The second control bus at row n-Δ samples the photo voltages of this row into the sample and hold bank of capacitors at the top of the imager before resetting it and moving to row (n-Δ+1). From the definition of the integration time, namely that it is the time difference between consecutive pixels's reset and sampling readouts, two images of different integration time will result. Therefore, the image constructed from the imager's lower sample and hold capacitor bank will have integration time equal to T_{lower_spl} whereas the image constructed from the upper sample and hold capacitor bank will have integration time equal to T_{upper_spl} shown in equations 3.3 and 3.4 respectively.

$$T_{lower\_spl} = (N - \Delta)\, T_{row} \quad \text{... (2.13)}$$

$$T_{uperr\_spl} = \Delta\, T_{row} \quad \text{... (2.14)}$$

N is the imager's number of rows, Δ the number of rows between the consecutive sampling control bus pair as shown in Fig 2.17 and T_{row} is the time needed to sample one row. In the classical photodiode-type CMOS APS [67] operating in normal mode, image sampling is realized by reading row by row (raster scan). The sensor data from the selected row is sampled simultaneously for all columns onto a sampling capacitor bank at the bottom of the columns. The pixels in the row are then reset and read (all together) a second time, and a new integration is started. The capacitor bank is then scanned sequentially for readout. This scan completes the readout for the selected row. The next row is then selected and the procedure repeated. Thus, the readout of one row is composed of two steps; the row sampling step that

takes a period of time T_{spl} in the range of 1µs to 10µs, and the scanning step to readout serially the sampled pixel values T_s, which is in the range of 0.1µs to 1µs [68]. Assuming the imager has M pixels per row, T_{row} can be estimated through the following equation 2.15

$$T_{row} = 2T_{spl} + M\,T_s \text{ ... (2.15)}$$

Now, two images of two different integration times for the same impinging scene light intensity are sampled. Therefore, we have here two exposures of the same scene and thus for the DR enhancement estimation we can use the formula in equation 2.9 after image fusion (either non linearly by summation or linearly by bit concatenation [68] or else) as shown by equation 2.16 assuming $\Delta < N/2$.

$$DR = 20\log\left(\frac{I_{\max}}{I_{\min}}\right) + 20\log\left(\frac{Ex_{\max}}{Ex_{\min}}\right) \text{ or } DR_{enh} = 20\log\left(\frac{T_{lower\_spl}}{T_{upper\_spl}}\right) \text{ ... (2.16)}$$

The assumption of $\Delta < N/2$ is made just to make sure that $T_{lower\_spl} > T_{upper\_spl}$. Developing further equation 2.16 will lead to;

$$DR_{enh} = 20\log\left(\frac{N}{\Delta} - 1\right) \text{ ... (2.17)}$$

The main drawback of the dual sampling technique in particular and using multi-exposure in expanding the DR through the integration time in general is its vulnerability to temporal changes in light intensities such as in the case of moving objects. Whereas, using the multi-exposure through light intensity sense may degrade the spatial resolution of the imager. Any solution to overcome these limitations either by integrating multiple (colorless) light filters such as the one used for in-pixel color filtering in the Foveon™ X3 pixel [70], or by the

integration of more circuitry inside the pixel (to minimize blur effect) such as the one used in the multiple capture technique [69], is in the right direction to extend further the DR of CMOS imagers. We suggested in the present thesis a spatial-varying DR enhancement called foveated dynamic range (FDR) based on the same principle of dual sampling but extending the sampling dimension from the 1D row (and 2 output channels) sampling to a 2D ring (and 8 output channels) sampling. The proposed FDR approach is intended to achieve two goals; minimizing the blur effect inherent to 1D sampling and realising non-uniform (foveated) DR. The spatial variance of the FDR (higher DR enhancement at the imager center area and decreasing outwards) is aimed at minimizing the image data throughput while maximizing the amount of transmitted image information (through DR), mimicking biological vision.

2.7 Why Foveated Dynamic Range?

Before answering the above question let us first consider the importance of foveated imaging and why its emergence is getting more and more interest [71][72][73][74]. This interest arises from the fact that most of the human visual system characteristics have been exploited to reduce the video communication system requirements (cost) when it is designed for information consumption by human observers. First, the temporal contrast sensitivity of the human visual system declines at high frequencies creating a temporal resolution cut-off of approximately 60Hz. Second, the spatial contrast sensitivity of the human visual system declines at high frequencies creating a spatial resolution cut-off of approximately 50 cycles per degree (cpd). Third, chromatic information is encoded in the human visual system by only three broad-band photoreceptors, with peak sensitivities at 440, 540 and 570 nm.

Fourth, the chromatic spatial resolution of the human visual system is lower than the luminance spatial resolution by a factor of approximately two. However, it is well known that the human visual sampling is not uniform and in fact the spatial resolution of the human visual system reduces to 50% at 2.5º away from the direction of gaze (fixation point) and to 10% at 20º [4]. This latter human vision limitation has not been exploited until recently [74] to minimize the bandwidth cost when transmitting video information. This cost reduction will be better appreciated with increasing video image resolution with the advent and emergence of High Definition TV (HDTV) technology and higher frame rate and image resolution video phones. In this regard, suggesting a foveated sampling system would be of greater interest as it will reduce the cost of coding/decoding for image transmission. Using standard CMOS technology to design a spatial-variant foveated imager is on the other side not practical either especially in terms of yield. Therefore, using another foveated imaging characteristic is of interest and thus the idea of foveated dynamic range (higher DR for central pixels than peripherals) emerges.

Chapter 3

Foveated CMOS Image Sensors Design

3.1 Brief description of CMOS 0.18μm technology

Standard complementary metal-oxide-semiconductor (CMOS) 0.18μm technology, which has been used to design the CMOS image sensors discussed in this chapter and in Chapter 7, is briefly introduced in this section.

CMOS 0.18μm technology using 0.18μm as the minimum feature size[12] is an N-well process technology that uses a P doped substrate. This very large scale integration (VLSI) technology is a 6 metal 1 polysilicon (6M1P) Salicide technology that has the capability to use up to 6 metal layers and only 1 single polycrystalline layer. Salicide refers to the Self-Aligned siLICIDE process in which "Self-Aligned" refers to the technique of making the source and drain of MOS transistors not extending below the gate when their diffusion junctions are formed and "Silicide" refers to the silicon-metal alloys (e.g. $TiSi_2$ (Titanium Silicide) $CoSi_2$ (Cobalt Silicide) or $TaSi_2$ Tantalum Silicide)) [75]. Silicides are being used mainly for their electrical benefit, by lowering the electrical resistively of polysilicon used in MOSFET gate and metal contacts, beside their mechanical strength supporting the dry etching in plasma reactors (needed during chip fabrication). These metal alloys are also know for their immunity of the electro-migration in polysilicon contact [76]. It is very important to note that in the SALICIDE process the silicide alloys are not only deposited on the MOSFET gates

[12] Feature size refers to the minimum MOSFET channel length of fabricated MOS transistors.

(and in this case the process is called *Policide* approach [75]) but also on the source and drain beside building diffusion contacts, all for the sake of minimizing their resistivities. In the active areas (source and drains), which are used in photodiode structure (N+ diffusing over P substrate for example) the silicide is opaque to the incident light (because of the silicide's metal reaction (opacity) to electromagnetic waves). To avoid this drop of quantum efficiency (due to the low e-h pairs generation), which will translate into photo-signal drop, a Resist Protection Oxide (RPO) layer is used over the photo-sensing area of the pixel photodiode in order to avoid depositing the silicide over it during fabrication. This operation has also been used for the same reason in [77]. Finally, the CMOS 0.18µm technology used in designing CMOS imager described in this thesis is a dual voltage technology supporting both 1.8V and 3.3V power supplies. This means, additionally, that it supports the fabrication of both, thin oxide MOSFETs of 0.18µm feature size with power supply of 1.8V and thick oxide MOSFETs of 0.35µm feature size with power supply of 3.3V power supply. This is a very useful feature to be used in designing CMOS imagers so that certain circuitry will be using thin MOSFETs and certain ones will be using thick oxide MOSFETs depending on individual needs that will be clarified below. The fabrication and design were realised thanks to the support of Canadian Microelectronic Corp (CMC) and thus detailed proprietary information regarding CMOS 0.18µm technology will not be discussed. However, the reader is referred to [75] for further general information about CMOS technology. All designed chips were fabricated at Taiwan Semiconductor Manufacturing Corporation (TSMC).

3.2 Foveated Architecture Motivations

One of the major characteristic of human vision is the foveation of the photocells distribution and interconnection (with ganglion cells) in the retina. This gives more emphasis to the central area (corresponding to the optical axis) sampling allocating more neuronal resources to the vision of this area. Adding to this vision aspect the eye saccadic movements (gazes) the human eye can build (over some period of time and saccadic steps) a high resolution image over a large spatial field of view. This architectural aspect of human vision is indeed important to future high resolution high frame rate CMOS imagers in a verity of applications. One of the most important applications of foveated CMOS imagers is video phones which are currently available in the market but with slower frame rates. In this application the video phone is directed specifically to the area of interest to sample it and transmit it to the other end of the communication in a similar way of the human eye vision. The video phone motion to track an event or a moving object is therefore equivalent to the eye saccadic movements. Yet, the sampling architectures of the current video phone still acquire images with the uniform classical imaging architectures suffering therefore from the large amount of transmitted vision data. This is why more of mobile videophones use sample video information and save it internally before send it to the other communication end. This is obviously limited by the relatively small and cost memory impacting the limited recording time. Another application that is attracting increasing interest is head-mount cameras needed for low vision enhancement application. This application as it mimics the human vision to enhance the impaired vision people is a potential application of foveated CMOS imagers.

3.3 Design Motivation

It well known in designing Application-Specific-Integrated-Circuits (ASIC) circuits [78], that the designer may choose one of three methodologies for implementing any microelectronic system on a chip (SOC), based on the complexity of the design. These approaches are;

Top-Down approach: in which the designer starts working at the system or architectural level of the microelectronic system without including the lower level building blocks. Subsequently, each block increases in details by partitioning it until a structural representation is realized.

Bottom-Up approach: in which the designer starts building the microelectronic system from a structural representation and constructing the compositional blocks of the system up to the highest level (chip level). This is rather a cumbersome methodology especially with larger ASIC designs (larger than 10.000 gates) [78].

Bottom-Up and Top-Down approach: in which the designed uses the previously mentioned methods by swapping between them in order to meet certain constraints to meet such as power, silicon area and speed.

In the present chapter, the system level design approach has been chosen because a new architecture is proposed for sampling the integrated image. Based on the higher level (or behavioural level) schematic architecture, the designed chip will be built block by block accordingly. While the design level selection for a designer is usually made upon the degree of complexity of the design itself and the level of abstraction the designer wants to use, in our case the reason is different. The reason behind choosing the system level approach in

designing a new architecture is to come up with a scheme that can provide us with some key features needed in a CMOS image sensor dedicated primarily to minimize the amount of sampled data and maximizing the information carried. This will be further examined in the subsequent sections. The suggested architecture, called pyramidal architecture, is based on keeping the building block pixel structure the same as the classical active pixel sensor (shown in Fig 3.4), while changing the sampling architecture. In the following sections, the design constituents will be presented in more detail until the whole pyramidal CMOS image sensor chip is constructed.

The bottom-up approach is an important approach in designing dedicated CMOS image sensors when the pixel architecture is the driving force of innovative CMOS imagers. It is through this philosophy that the Multiresolution CMOS imager was designed in which the classical orthogonal CMOS imager architecture has been adopted. In this case, however, the APS architecture has been altered to support more functionality. A summary of this architecture is mentioned in section 0 and further details of the multiresolution CMOS imager implementation is demonstrated in Chapter 7.

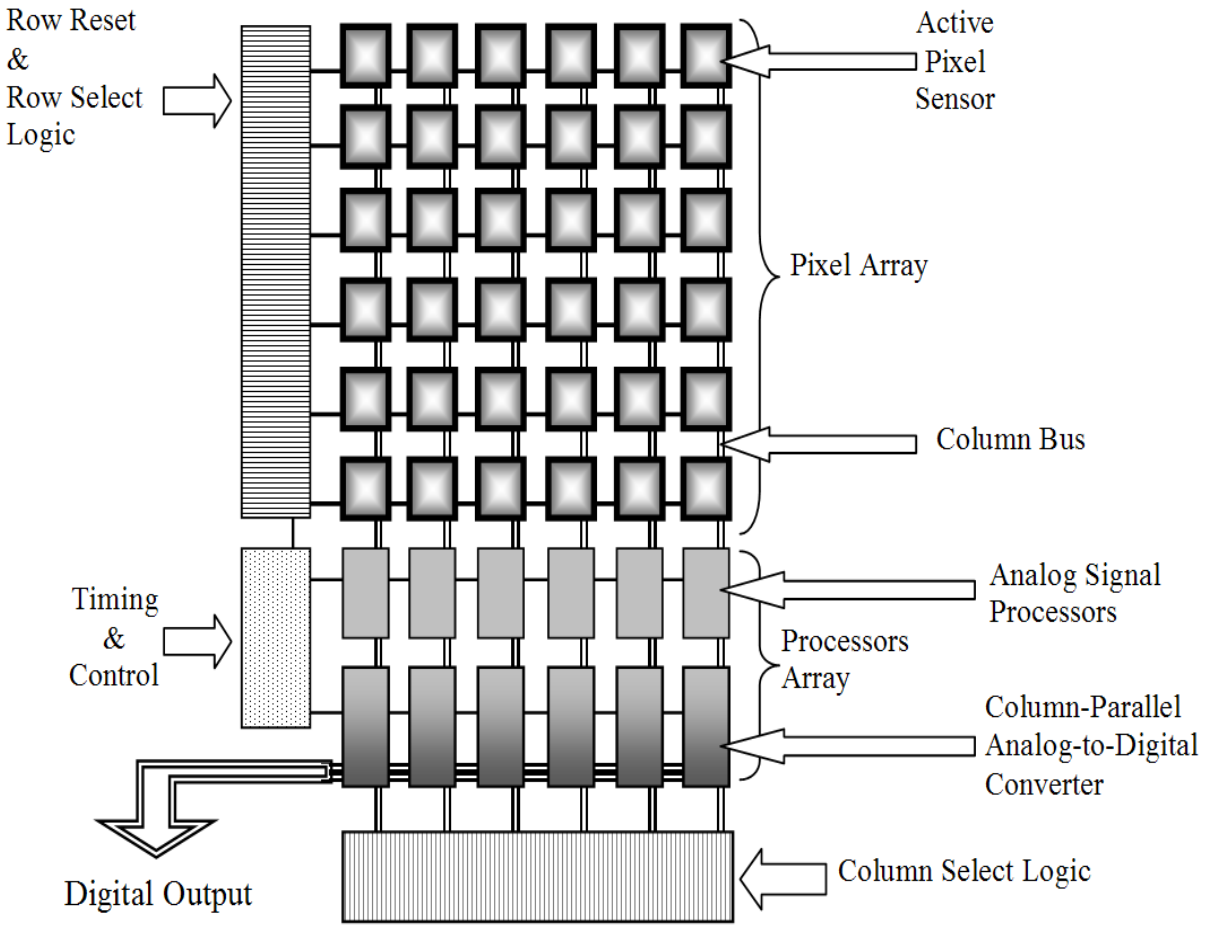


Fig 3.1 Classical sampling architecture of CMOS image sensor

3.4 Pyramidal CMOS Imager Design Tools

The main tool used to design the pyramidal standard CMOS imager (and the multiresolution CMOS imager) is a UNIX-based chip design tool package known as Cadence® supported by the Canadian Microelectronics Corporation (CMC). The main tools used in this package are Virtuoso® schematic editor and Virtuoso® layout editor. The former tool was used to

simulate the schematic view of the different blocks of the imager and the latter used to layout the different mask layers used to physically fabricate the whole chip. Virtuoso® layout editor was used beside design rule check (DRC) using Dracula® or DivaDRC®, was used to extract the physical parameters of the different devices of the blocks beside their parasitic capacitors and resistors. All simulations were carried out using Cadence® Analog design environment which uses the model files of CMOS 0.18µm process of TSMC foundries where our CMOS imagers were fabricated.

Another tool has also been used to design the decoders called ICCraftman® by translating their schematic view and auto place and route then into the layout view. This tool was mainly used in the large multiresolution decoders. It never finishes the whole routing so we did complete the remaining routes manually.

After successfully testing each block in the schematic and layout view, all the imager blocks were brought together and connected to construct the imager chip. Simulating the whole imager chip was not possible because of the difficulty to integrate the optical signal interaction with the active silicon areas beside the larger array of pixels. Thus, visual inspection as well as Mark-Net command (in the layout editor) testing all the interconnection between the blocks was only way to verify the correctness of imager chip. The last verification of the design rules was made by streaming out the imager chip and uploading it to CMC server were further DRC were carried out. This phase includes (beside basic DRC check) the antenna violations caused by long wires which are vulnerable to electric discharges and burn out (cut). This problem is remedied by placing diodes on the long wires

paths. The final step before submitting the imager for tapeout (fabrication) was to make sure the polysilicon and metal filling densities meet the requirements of TSMC foundries.

3.5 Pyramidal Architecture and Its Building Blocks

Before further examination of the pyramidal CMOS image sensor design, the architecture along with its constituent blocks is first presented. Then, the strategy of how the whole design is going to be laid out using the available standard CMOS 0.18μm technology is developed.

3.5.1 Architecture Description

Before describing the new pyramidal architecture, we first briefly review the operation of the conventional active pixel sensor.

The classical CMOS image sensor shown in Fig 3.1 is usually read out using raster scanning. The incident light on every pixel is integrated over a period of time, called the integration time, which starts from the (current frame) reset signal until the (next frame) select signal. The controlling signals (Reset and Select), which are shared for every row's pixels in the matrix of pixels, are generated from the row reset and select logic block that is made typically of shift registers or of decoders. The structure and physical functionality of the pixel will be presented in the next section. After the row selection signal is pulsed the whole row' pixels voltages are dumped, through column buses, into the sample and hold (S&H) bank capacitors shown in Fig 3.1 via the analog signal processors. Finally, the sampled voltages in S&H capacitor bank are serially buffered out either after being digitized, through the column

parallel analog-to-digital converters (ADC), or just directly buffered out in their analog form through an analog buffer.

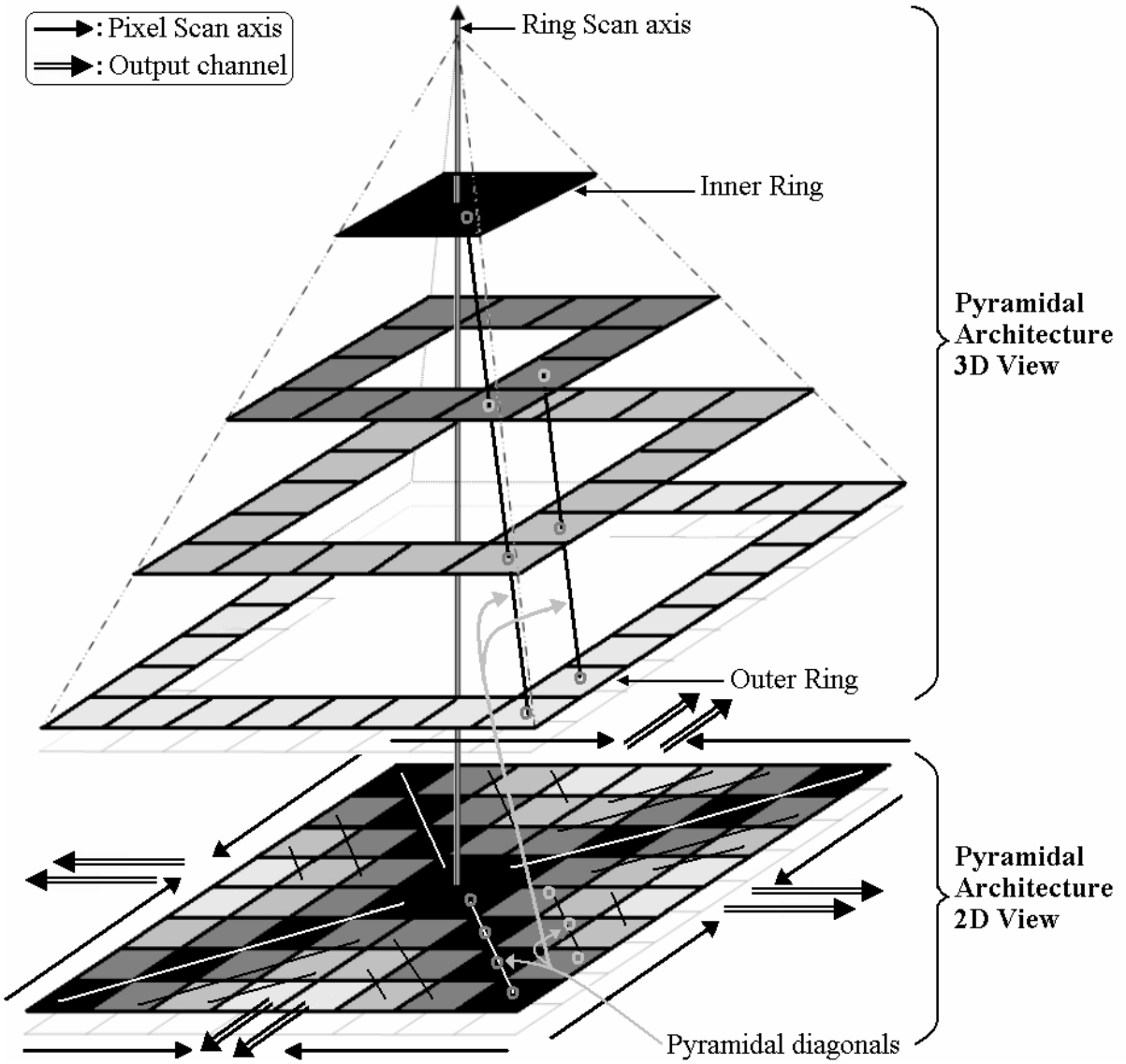
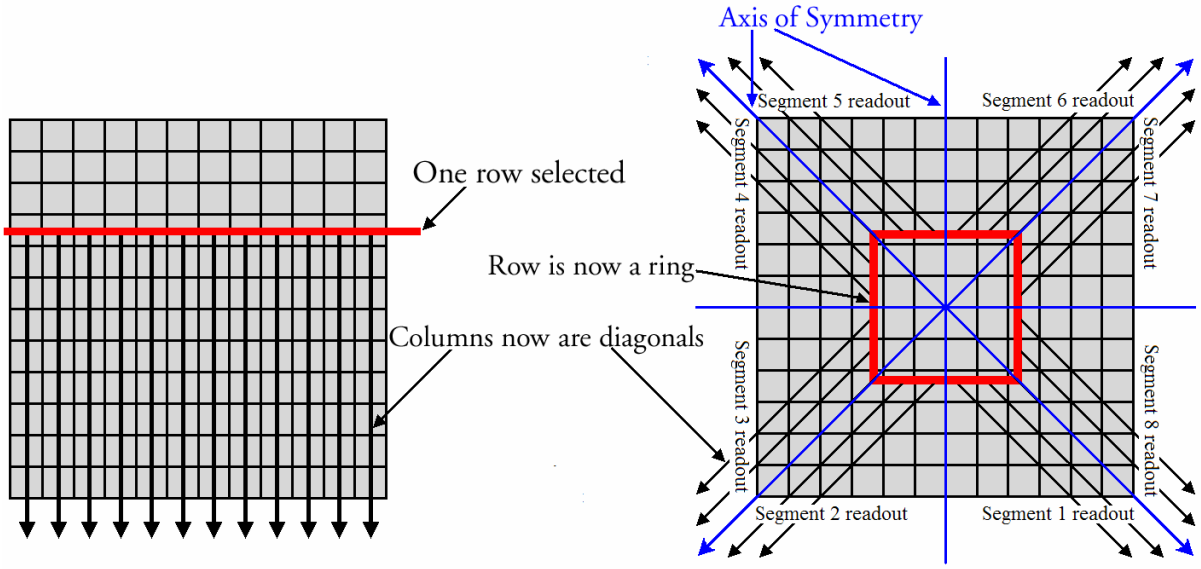


Fig 3.2 Pyramidal architecture schematic views

In the pyramidal CMOS imager architecture, the sampling unit is a 2D ring instead of the 1D row sampling in the classical architecture, as shown in Fig 3.2 and Fig 3.3. Therefore, all the pixels belonging to a *ring* in the pyramidal CMOS imager share the same reset and select signals that are consequently "2D" control signals. The reset rings are connected to a reset decoder and the select rings are connected to a select decoder. After the select signal is pulsed, the sampled photo-voltages are dumped into the S&H capacitors (not shown in Fig 3.2) situated all around the pyramidal imager right after the last ring (outer ring). The sampled photo-voltages are transported to the S&H capacitors through diagonal (at 45º and 135º) busses, as shown in Fig 3.2. The rest of the photo-signal path is similar to the classical CMOS imager architecture. Fig 3.3 recapitulates the major differences between the classical architecture and pyramidal architecture of CMOS image sensors.



(a) Classical architecture sampling (b) Pyramidal architecture sampling

Fig 3.3 Variations between the classical and pyramidal architectures of CMOS imagers

3.5.2 APS Pixel

The picture sampling unit cell, known as a pixel, used in the pyramidal architecture is the same as that used in the classical CMOS imagers, and in particular the active pixel sensor that is shown Fig 3.4.

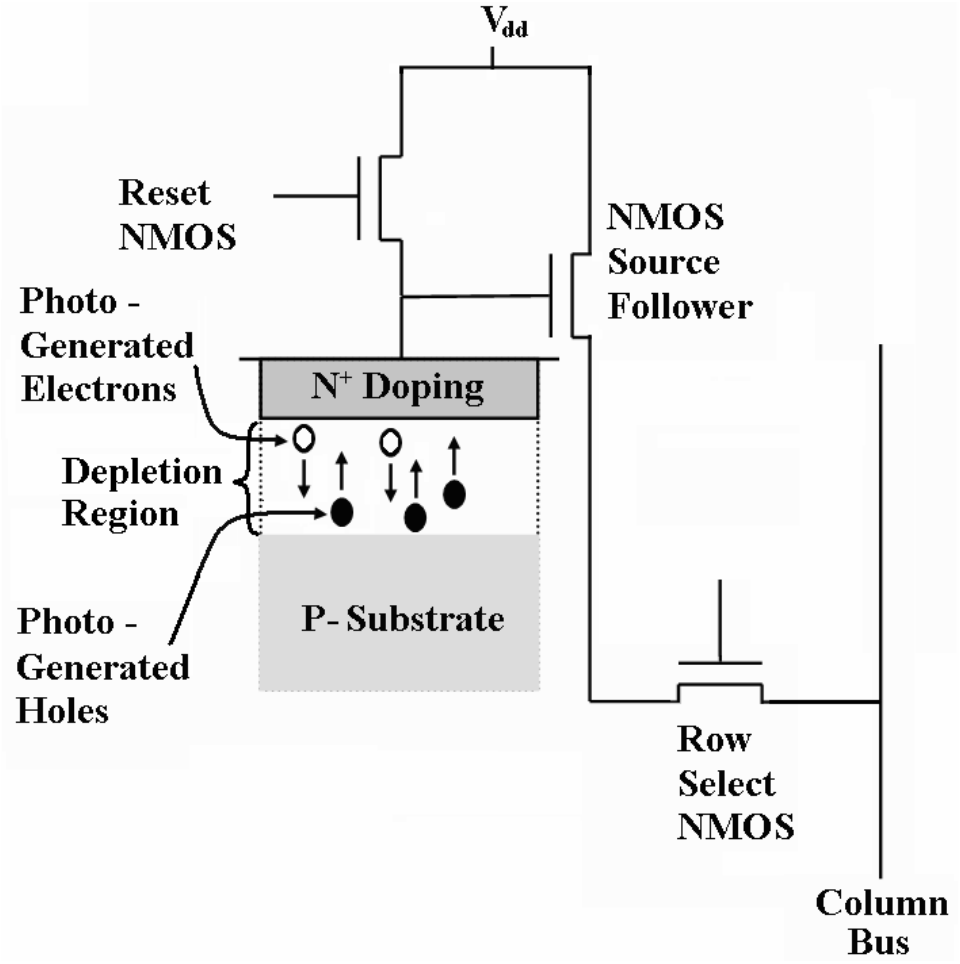


Fig 3.4 The classical structure of active pixel sensor with N^+/P_{sub} photodiode

The APS used in the design of pyramidal CMOS imager uses an N^+/P_{sub} photodiode for integrating incident light. The photodiode is first reverse-biased through the reset NMOS transistor by PRST signal. The incident light starts discharging the photodiode (PD) right after the reset transistor is turned off. At the end of the integration time, the pixel is read out (to POUT) though the select transistor activated by the PSEL signal. The pyramidal CMOS imager pixel shown in Fig 3.5 is a 16µm by 16 µm pixel and its sensing (or active) area, composed of N^+ diffusion over the P-type substrate P_{sub} (not shown in the figure) is of 199.66 μm^2.

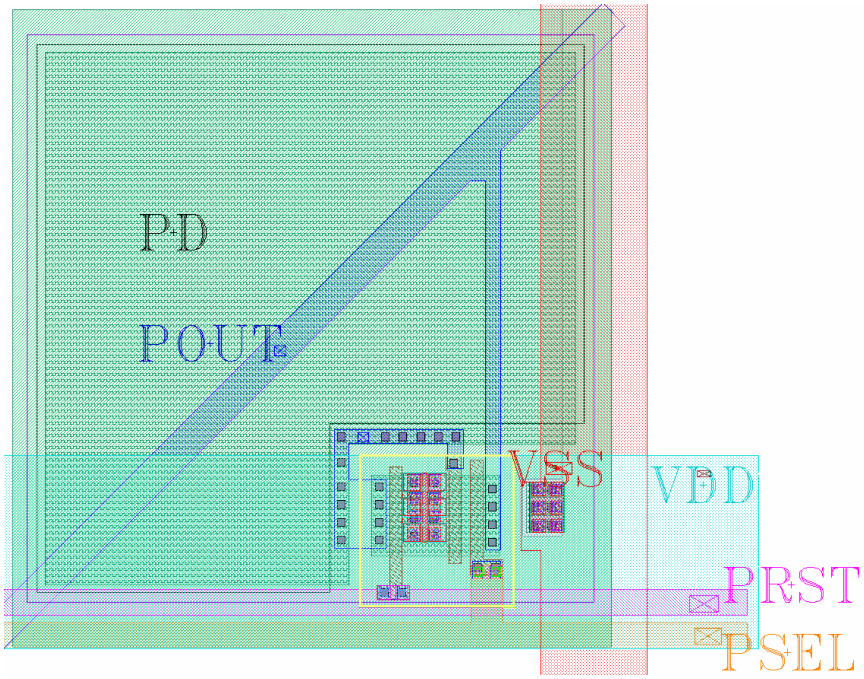


Fig 3.5 Layout of the pyramidal CMOS imager pixel

The fill factor that is the ratio of the sensing area over that of the whole pixel is about 78%. Note that the pixel output bus (POUT) is diagonally laid out (using metal1) by 45º (others by 135º), as shown in Fig 3.5, in order to construct the pyramidal imager's diagonal output buses.

Interestingly, the construction of the focal plane pixels' array in the pyramidal CMOS imager is different from that of the classical CMOS imager architecture because of the difference of symmetry between these two architectures. While the classical CMOS imager architecture pixel array matrix has the vertical and horizontal axis of symmetry the pyramidal architecture has (besides the vertical and horizontal axis) the diagonal axis (45º and 135º) as extra axes of symmetry, as shown in Fig 3.3. This symmetry difference between the two architectures has an impact on the process of laying down the pixels array. In classical imager architecture, translating the pixel (by copying and pasting it) horizontally and vertically, the whole pixels' array can be constructed. As for the pyramidal architecture, the pixels' array has to be constructed from the inner ring towards the outer ring by translating the pixel from one axis to another, then by mirroring the constructed ring segment around all the axis of symmetry until the whole ring is done before jumping to the next ring and so on. During this process, the ring's shared signals, namely the PSEL and PRST (pixel select and reset signals respectively) will be connected, consequently, at the end of pixel ring's formation. Through this point of view, the pyramidal CMOS imager architecture and its pixel along with the involving symmetries looks very similar to the crystallographic structures (nets) formation with their building blocks (basis) found in nature [79]. The fact that in order to construct the

pyramidal imager rings, one has to start from inner rings towards the outer rings makes the layout of the pyramidal CMOS imager similar to how crystals grow in nature. Fig 3.6 shows the first two inner rings.

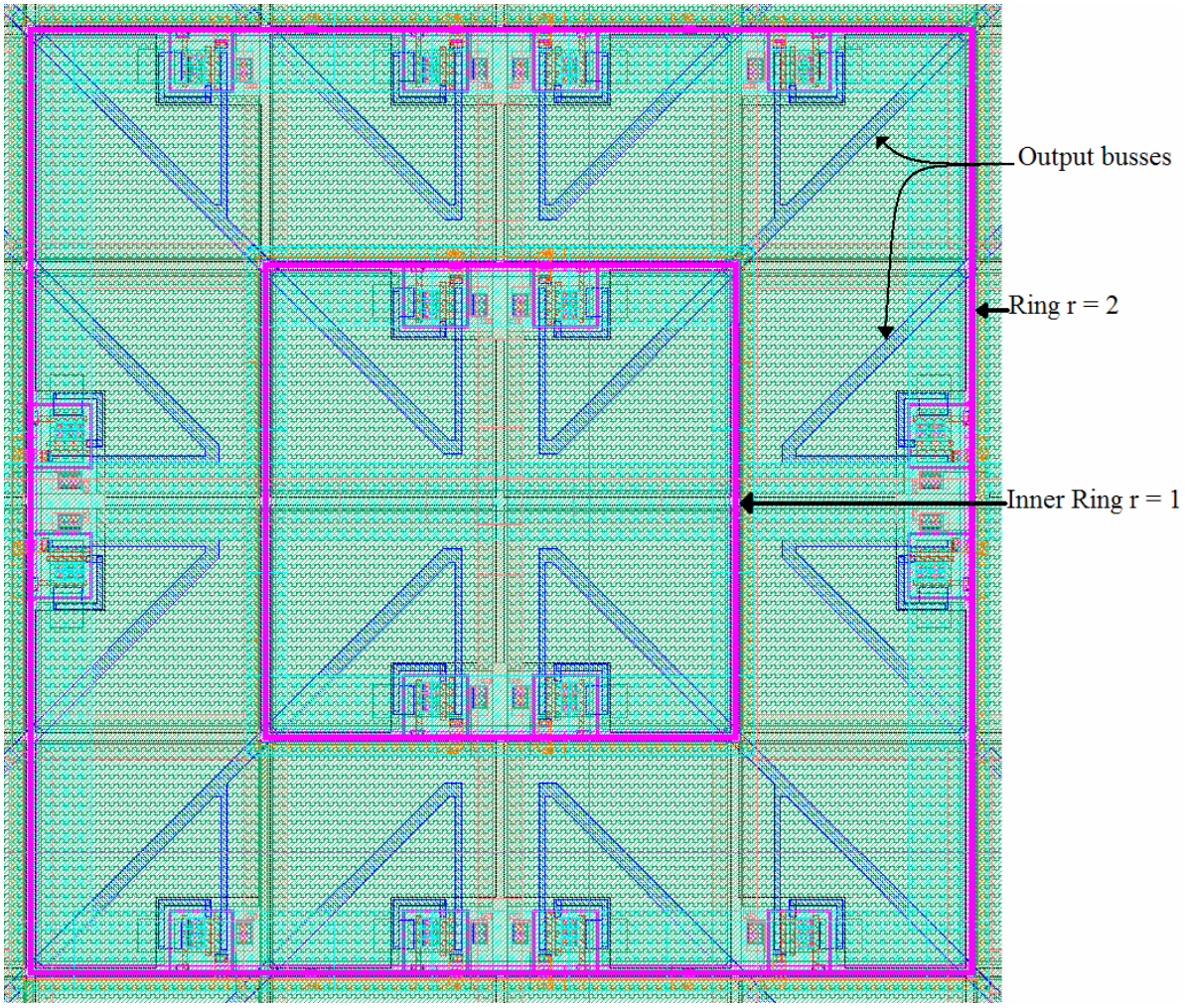


Fig 3.6 Layout of the two first inner rings of pyramidal CMOS imager

One may note that the diagonal pixels (that include the inner ring pixels) are shared between adjacent segments and hence are sampled by S&H capacitors twice as large as any other

pixels. This will lead to two diagonal artefacts in the sampled image as shown in Fig 4.4. The remedy to this problem is either by halving the sampling capacitors of these diagonal pixels or choosing one segment S&H for these diagonals which creates and irregularity of the readout process[13].

Before closing this section it is worth mentioning that the MOSFETs used in building the pyramidal pixel were of thick oxide that supports gate voltages of 3.3V that is also the power supply V_{dd} of the pixel. Two reasons are behind this choice. The first is to have higher electrical dynamic range (or voltage swing) at the output of the pixel. The second reason is to minimize the leakage current through the gates of the RESET and the source-follower NMOS transistors (see Fig 3.4). For deep submicron CMOS technology (such as CMOS 0.18μm), this gate leakage current potentially becomes higher than dark current and even comparable to photocurrents under normal lighting conditions [80]. This gate leakage current is mainly due to the decrease of gate oxide thickness into the nanometer range, thus allowing direct tunnelling through the gate oxide [81]. The gate leakage current is also known as the Fowler-Nordheim tunnelling current. It is modeled by I_{FN} in the following formulation [75]:

$$I_{FN} = C_1 W L E_{ox}^2 e^{-\frac{E_0}{E_{ox}}} \ldots (1.3)$$

$E_{ox} \approx \frac{V_{gs}}{t_{ox}}$, t_{ox} is the oxide thickness, W and L are the width and length respectively of the MOSFET and C_1 and E_0 are constants.

[13] Because one segment will be larger than the other by one diagonal pixel

Consequently, not only the source follower and RESET NMOS transistors are made of the thick oxide MOSFETS but also the select transistor as well. Thus, the controlling signals namely RESET and SELECT signals take values of 0V and 3.3V corresponding to the 0 and 1 logic states.

Table 3-1 shows a summary of the characteristics of the pyramidal pixel.

Table 3-1 Physical characteristics of the pyramidal pixel

PIXEL CHARACTERISTICS	PHYSICAL VALUE
Pixel pitch	**16 μm**
Pixel area	**256 μm^2**
Active area	**199.66 μm^2**
Fill factor	**78 %**
Active area perimeter	**60 μm**
Power supply	**3.3 V**

3.5.3 Voltage Conversion block

From the previous section, control signals RESET and (pixel) SELECT are chosen to have 3.3 V as high logic voltage and therefore the controlling logic must be built using thick oxide MOSFETs. To minimize the cost of power and area in designing the logic blocks responsible of generating RESET and SELECT signals, a voltage converting unit is built. The voltage conversion unit is an interface that can convert the high voltage of 1.8 V to 3.3 V while passing the low voltage of 0 V unchanged. Therefore, the 1.8V-to-3.3V voltage conversion unit will allow us to design the control logic, i.e. decoders, generating RESET and SELECT signals using thin-oxide transistors. This way the imager design will benefit from the low power and area saving by using thin-oxide transistors in designing the logic blocks as they are insensitive (due to their digital nature) to the gate leakage current.

A circuit diagram of the voltage conversion unit is shown in Fig 3.7 and its layout is shown in Fig 3.8.

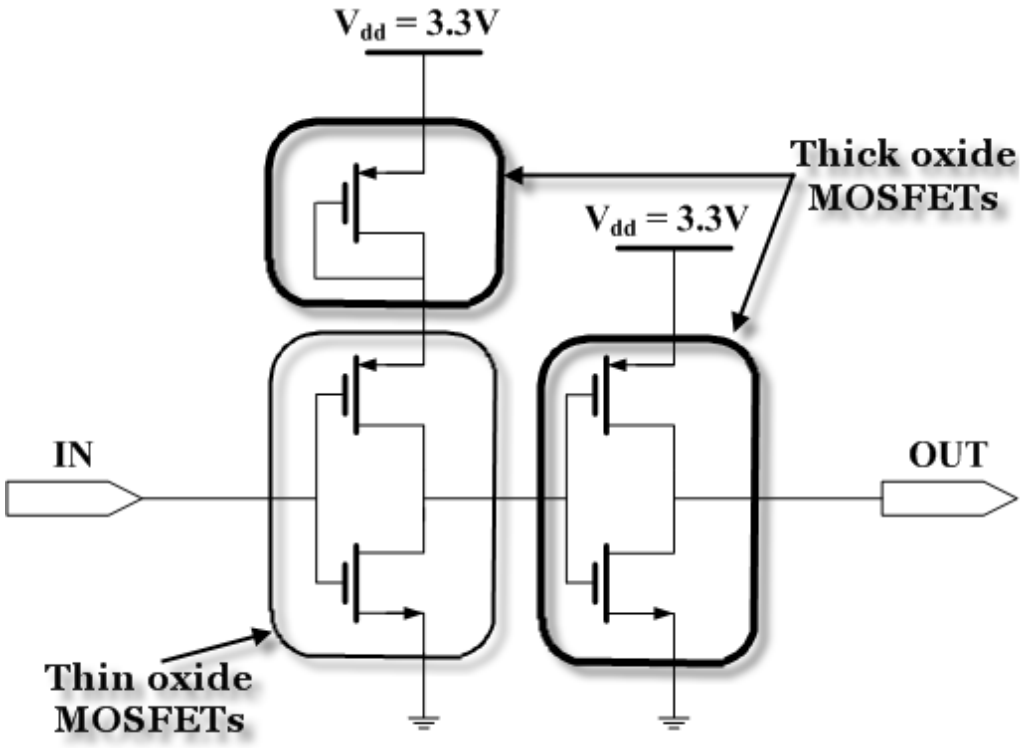


Fig 3.7 Schematic of the 1.8V-to-3.3V voltage conversion block

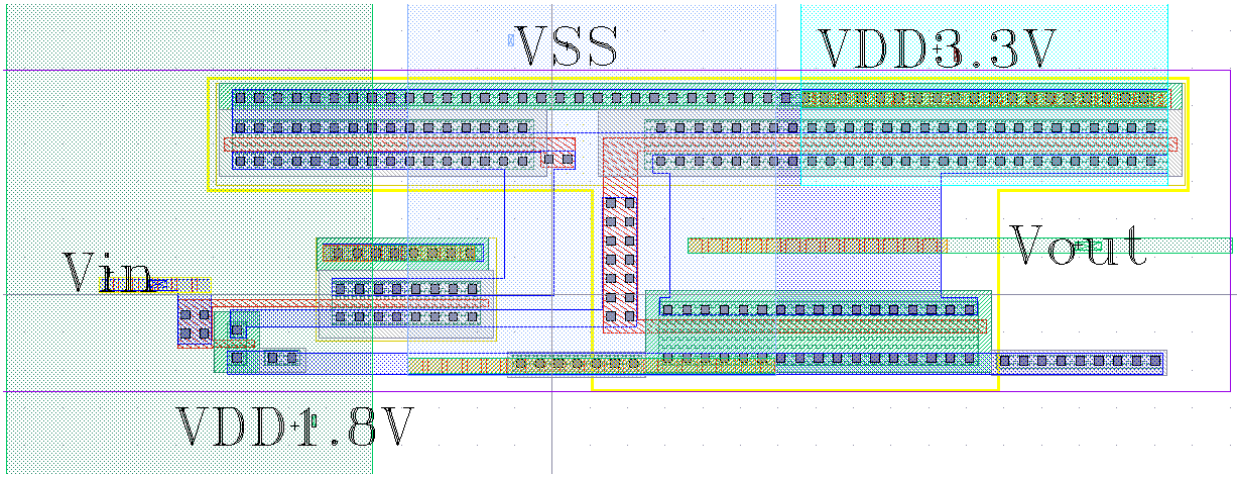


Fig 3.8 Layout of the 1.8V-to-3.3V voltage conversion block

In addition to the voltage conversion functionality of the present block, it has another more important role that is buffering. It is clear that the control signals (RESET and SELECT) are shared among rings that have different lengths. The inner rings are smallest and length progressively increases towards the outer rings. These control ring-buses have different resistances *R* and capacitances *C* that are related to the buses lengths and thus monotonically increasing from inner rings towards the outer rings. Therefore, the *RC* constants, which characterize the charging time of the different buses, are increasing with the square of the perimeter. This issue is solved by the voltage conversion block because it acts like a buffer interface between the logic circuitry, responsible for generating the control signals RESET and SELECT, and the rings' control buses. Thus the *RC* differences among the control ring buses become insignificant.

3.5.4 Sample and Hold block

Around the outermost ring, sample and hold (S&H) blocks are laid out and connected to the output busses that are shared between the pixels on the same diagonal, as shown in Fig 3.6. Each sample and hold block is composed of two metal-insulator-metal (MIM) capacitors sharing an NMOS load transistor to bias the active pixel's source follower. The two capacitors, Cs and Cr, are made for sampling and storing Vs (the pixel photo-voltage) and Vr (the pixel reset voltage) respectively, as shown in the schematic of Fig 3.9.

It is worth mentioning that the power supply used in this circuit block is 3.3V and its transistors are therefore thick MOSFETs due to its analog nature, like the pyramidal pixel, thus minimizing the gate leakage and preserving the voltage swing of the pixel.

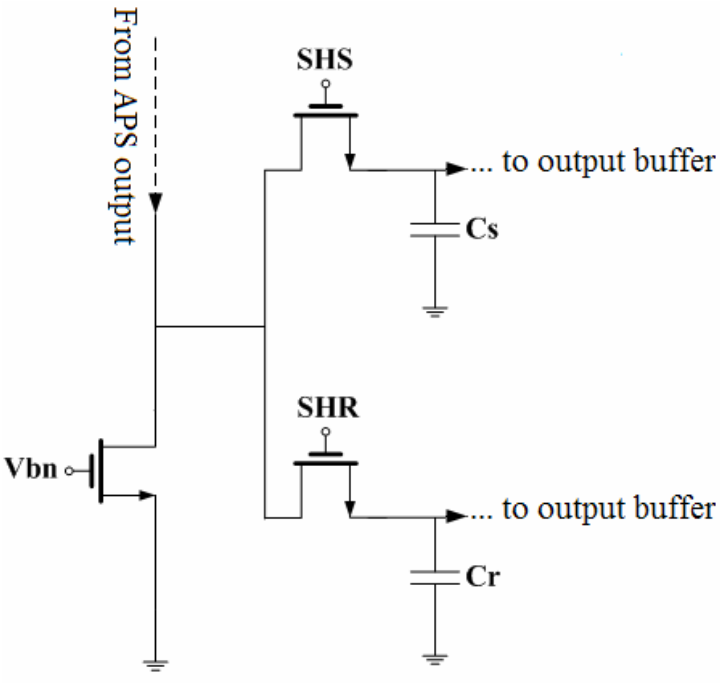


Fig 3.9 Sample and hold circuit schematic

It is worth mentioning that the sampled voltages Vs and Vr which are held in the capacitors Cs and Cr, respectively, are to be buffered out through each PMOS source follower with a gain close to unity.

Cs and Cr are both of 2pF capacitance, occupying each about 2000 μm^2 (about 160μm by 12.5μm) in an elongated shape. The sample and hold circuit layout is shown in Fig 3.10 below where VOUTS and VOUTR replace Vs and Vr, respectively, and VBIASN replaces Vbn. The lengths shown in this figure are in microns.

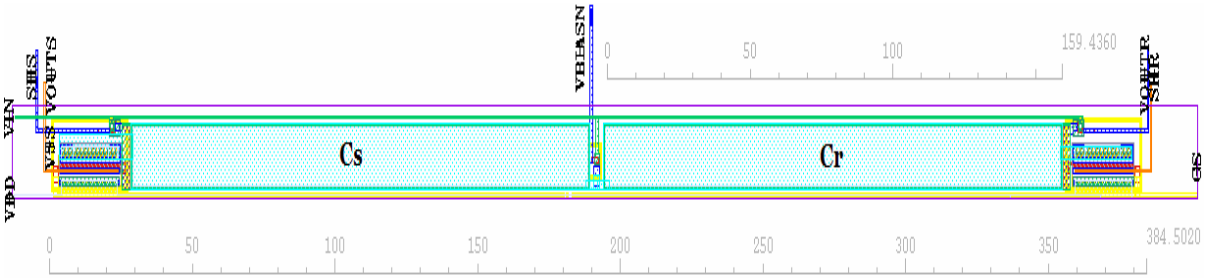


Fig 3.10 Layout of the sample and hold block

The sample and hold capacitors are made as large as possible in order to minimize the *charge injection* noise or what is commonly called *clock feedthrough* noise arising from the switching activity during the sampling process [82]. As the photo-transduced current flows from the source follower through the output diagonal bus, it passes the NMOS active resistor (controlled by Vbn) to be converted into a voltage value. Whether it is Vs or Vr, the converted voltage value get sampled into Cs or Cr through SHS or SHR NMOS switches respectively. The overlapping capacitances between sampling gates (SHS and SHR) and their drains (at Cs and Cr respectively) cause the feedthrough noise which is minimized by using large sampling capacitors as mentioned previously. After the sampling is done, decoders are used to buffer out Vs and Vr, stored in Cs and Cr respectively, by activating the PMOS switches through CS. The buffered voltages Vs and Vr will pass though the output buffer before being sampled off chip by the data acquisition system. Similar to the S&H structures used in conventional CMOS imagers, the column circuits can contribute significantly to the

fixed pattern noise owing to column-to-column mismatch between devices; this will be further examined in Chapter 6.

3.5.5 Output Buffer

In reality, the output buffer has been split between the sample and hold circuit, described in the previous section, containing the PMOS source follower (just before the CS switch) and a PMOS biasing transistor shown schematically in Fig 3.11 (surrounded by polygons for Vs and Vr buffering) and in the layout in Fig 3.12.

This output buffer in the current design of the pyramidal CMOS imager (as for many classical CMOS imagers) is used to meet two objectives. First, it provides current driving capability to buffer out the sampled voltages; second it compensates (with the S&H source-follower) for the V_{th} voltage drop lost in the NMOS source follower in the active pixel sensor. The first V_{th} occurs after the reset transistor that charges the photodiode up to (V_{dd}-V_{th}) instead of V_{dd} and the second V_{th} drop occurs after the select transistor in the APS pixel.

Finally, following the pixel and sample and hold blocks, the output buffer uses thick MOSFET therefore its power supply voltage is 3.3V. The main reason is to maintain the power supply voltage rail (V_{ss} to V_{dd}) from pixel to the off chip output.

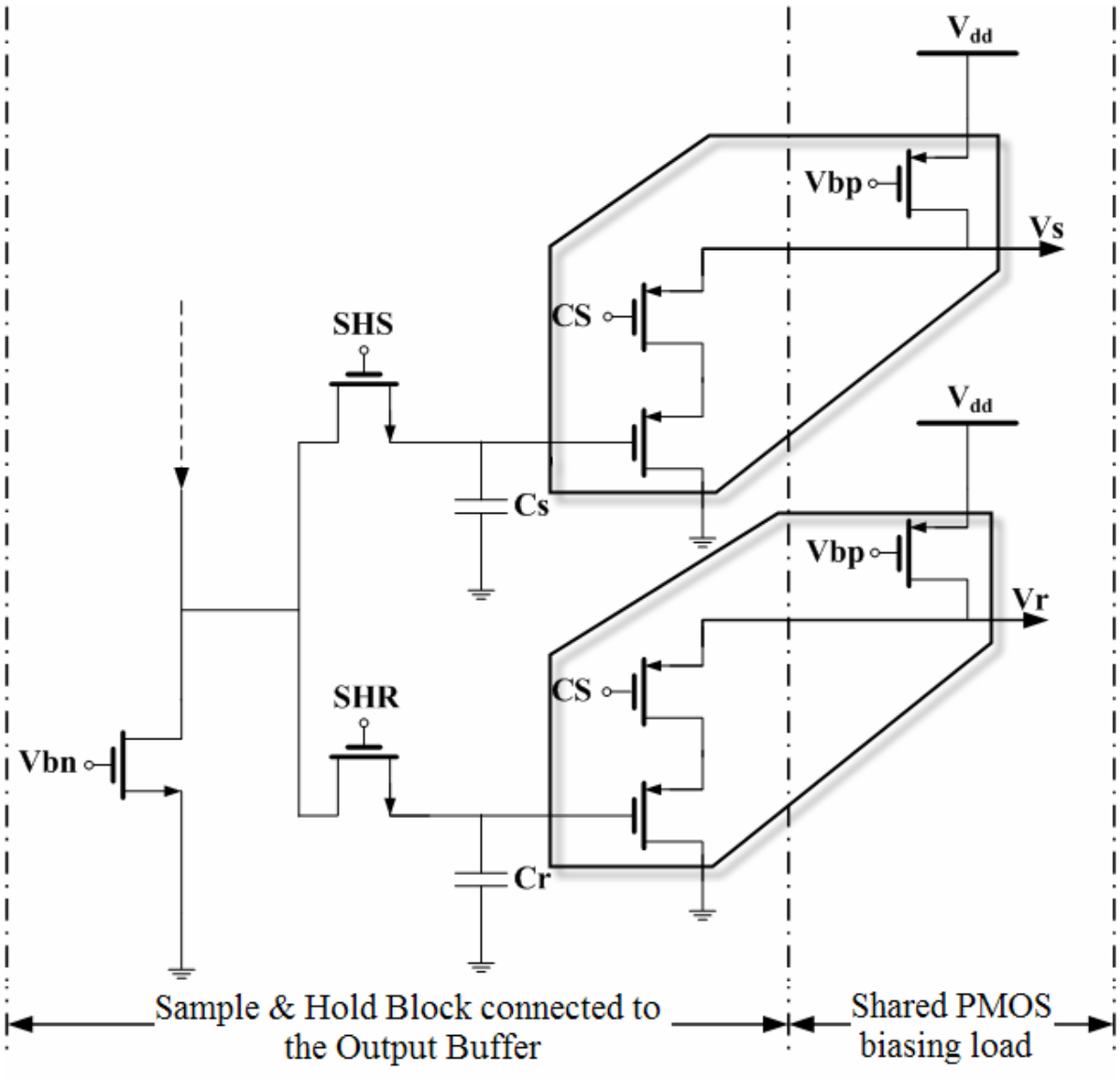


Fig 3.11 The output buffer split between the S&H block and the shared PMOS bias

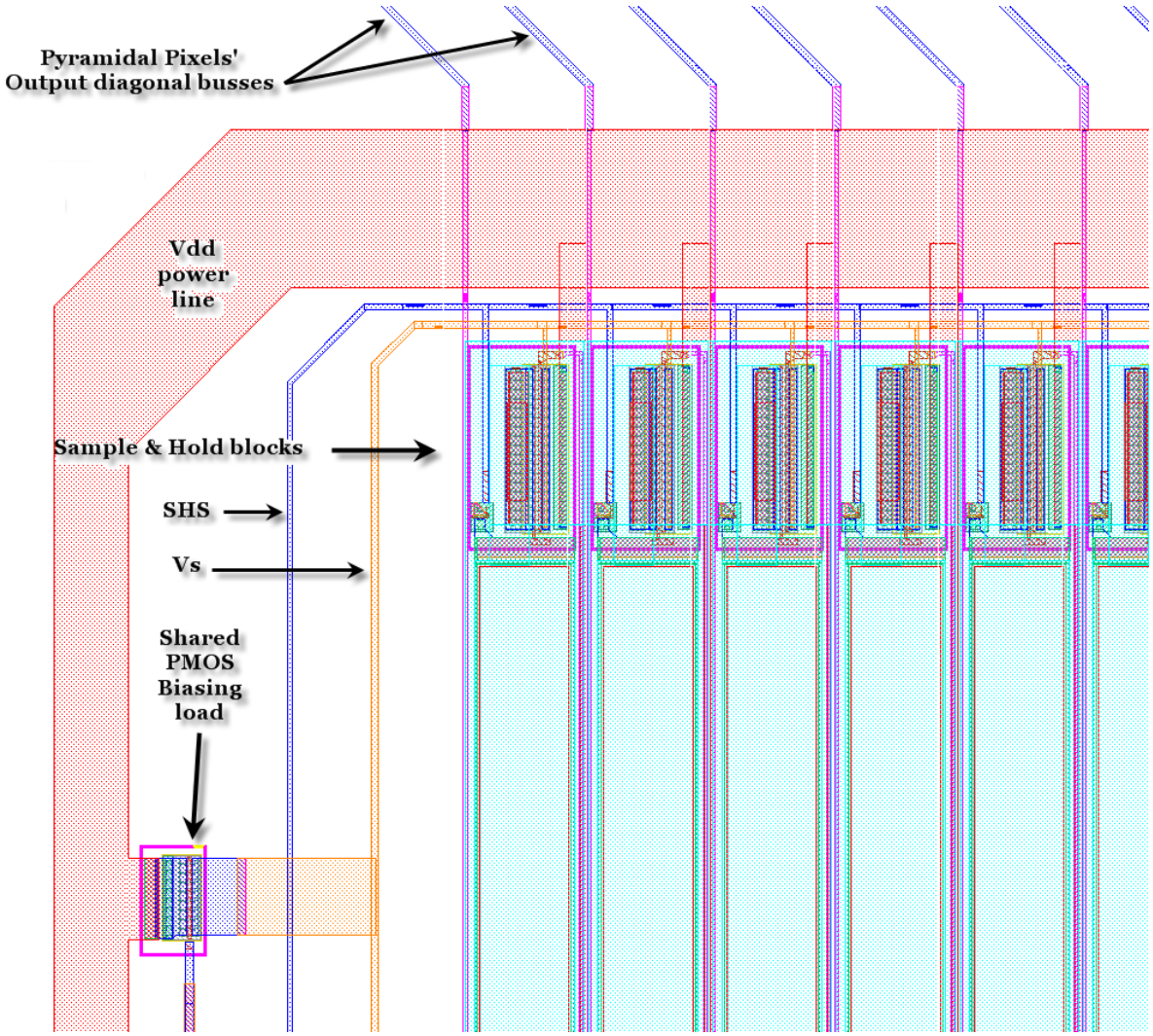


Fig 3.12 Layout of the output buffer within the S&H blocks

3.5.6 Decoders for Column (Diagonal) Select

At this stage all the components of the pyramidal CMOS imager have been described except the decoders, used in ring select or ring reset, which are standard circuits [83] mainly used in memories that are thought to be the closest standard CMOS systems to CMOS imagers[14].

Fig 3.13 shows a simple gate level block diagram of a typical 3 to 8 decoder that can be expanded to any m-to-2^m decoder. RESET and ENABLE signals are used to control the decoder that will have all its output low when RESET=0 (logic), and will have all its output high when ENABLE=1, otherwise the output ports will decode the input ports logically [83].

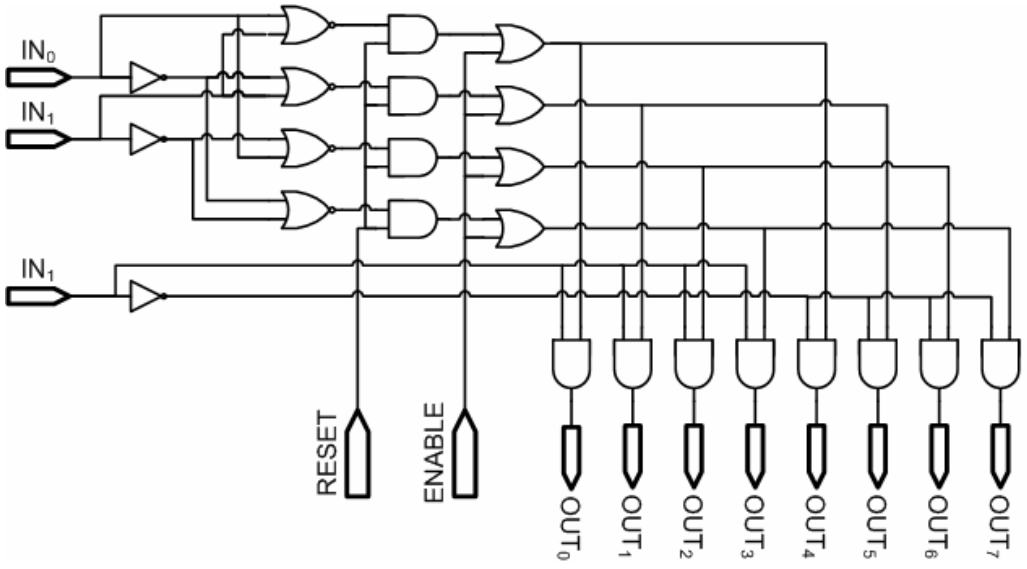


Fig 3.13. Gate level block diagram of 3 to 8 decoder

[14] This is the main reason why a memory manufacturing giant like Micron Corp. established its imaging division in Boise, IDAHO and Pasadena, California in the USA.

As mentioned in previous sections, the MOSFETs used in the digital logic blocks (decoders) are of thin oxide type (due to their gate leakage current immunity) for lower power and area consumptions.

3.5.7 Pyramidal Imager Chip design

Fig 3.14 shows the structural layout of the designed 64x64 pixel (32 ring) pyramidal CMOS image sensor occupying an area of about 4mm x 4mm while Fig 3.15 shows the photo-micrograph picture of the fabricated imager.

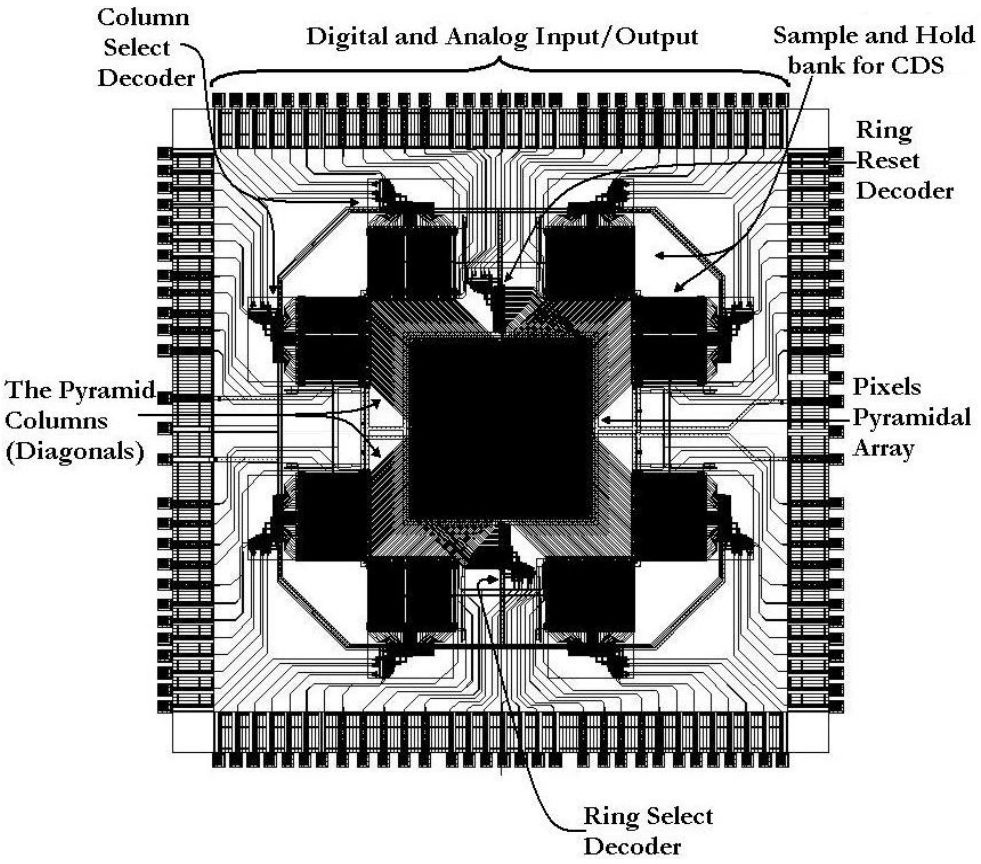


Fig 3.14 Pyramidal CMOS image sensor structural layout

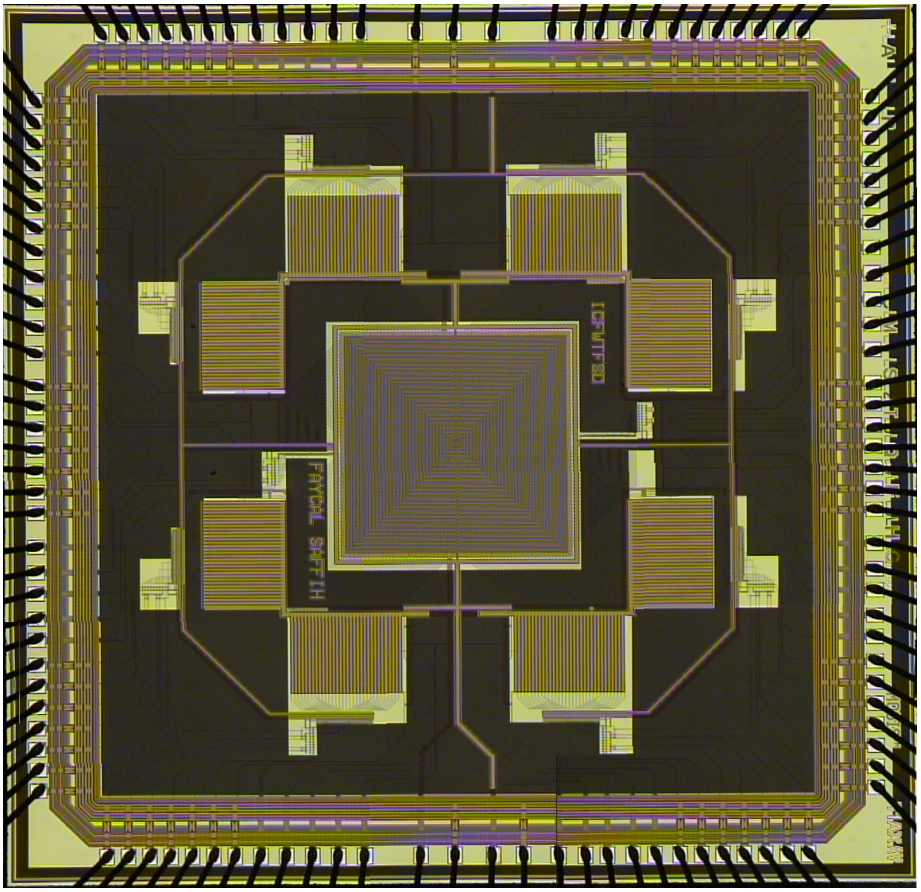


Fig 3.15 Optical micrograph image of the pyramidal CMOS image sensor

Finally, the general characteristics of the designed pyramidal CMOS image sensor are shown in Table 3-2.

Table 3-2 General characteristics of CMOS pyramidal imager

PYRAMIDAL CMOS IMAGER CHARACTERISTICS	PHYSICAL VALUE
Imager area	**3.9898x3.9898 mm^2**
Imager resolution (in pixels)	**64 x 64**
Design technology	**Standard CMOS 0.18µm (1P6M)**
Power supply voltage	**3.3V and 1.8V**
Voltage swing	**400 mV***
Dynamic range	**56.6 dB***
Sensitivity	$0.091\frac{V cm^2}{\mu W s}$ *

3.5.8 Bouncing Scanning

In the previous sections, the pyramidal CMOS imager architecture has been presented along with its implementation in standard CMOS 0.18µm technology. In this section, a new control mechanism is suggested in sampling images in addition to the implementation of the classical raster (rolling) scanning scheme in the pyramidal architecture. The suggested sampling

* This value will be extracted in Chapter 5

scheme, called bouncing, is proposed as an alternative to raster scanning through the statements developed in section 1.2.2.

Fig 3.16 depicts the difference between the raster scanning used in classical CMOS image sensor and the bouncing scanning used in the pyramidal CMOS imager.

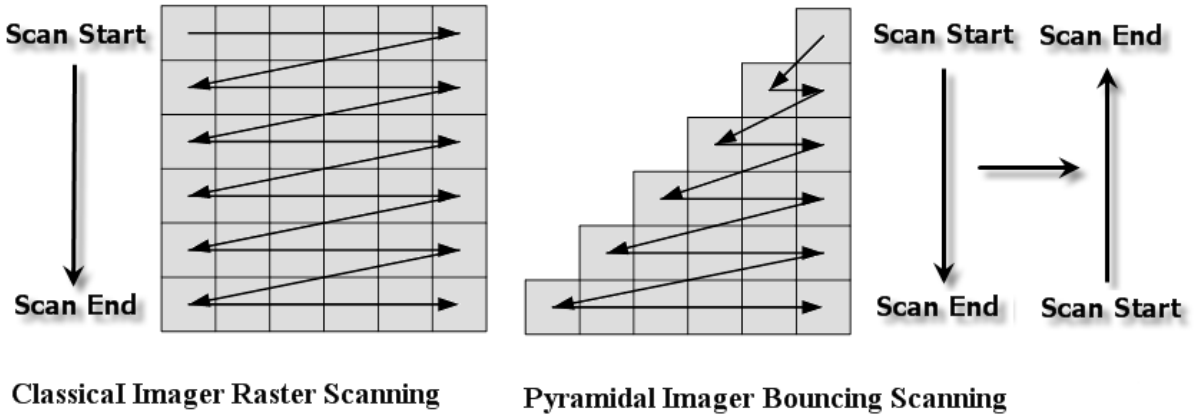


Fig 3.16 Classical imager raster scan and bouncing scan in a pyramidal imager segment

Raster scanning has been introduced earlier in section 2.1.4, and is also called the rolling scan because after sampling and resetting the last row, the first row is next in the scan chain and so on. In the pyramidal sensor, raster scanning is implemented by starting in the centre of the sensor and incrementing the rings; after the outer ring, the scan starts again in the centre. The scanning is therefore in the form of concentric circles. In contrast to raster readout of a conventional sensor, where the pixel update rate is different in the horizontal and vertical directions, raster scanning of the pyramidal sensor is symmetric. This property is very useful to minimize motion distortion as demonstrated in section 3.5.9.

In bouncing scan, however, the scan is bounced backward after reaching the last row rolling the scanning back towards the first row. After reaching the first row the scan is bounced back towards the last row and so on as shown in Fig 3.16. In this scanning strategy the integration time is not uniform and every row will have two integration times, one integration time for the inward scanning (from outer ring towards the inner ring) and another integration time for the outward scanning. More analysis of the bouncing scanning is developed in section 4.1.1. Although the non uniformity of the integration time might be an issue, fusing the resulting two frames (one from inward scanning and the other from the outward scanning) will result in a uniform photo response of the imager. This result is shown in section 4.1.2 namely through equation (4.6) and Fig 4.4.

3.5.9 Pyramidal Ring Sampling and Blur Symmetrization

Motion blur is an artifact affecting images acquired when the object image at the imager focal plan moves during image integration[15]. This image smear is more pronounced when spatial sampling speed of the image is relatively close to the motion speed of image at imager's focal plan. The motion of the image over the imager's integrating pixels will share the light intensity over a number of pixels resulting in an apparent smear in the image. The most common case of this artifact happens when using rolling shutter in CMOS imagers when acquiring fast moving objects or when used with pulsed-illumination [84]. With higher motion speed of the sampled objects in the focal plane, another type of motion artifact known

[15] The motion blur could be caused either by imaging moving objects or when the imager itself is moving.

as motion distortion takes place. This is mainly due to the missed areas from the sampled object in the resulting image. The cause of this artifact is explained by the violation of (or approaching) the Nyquist temporal sampling limit [85]. This sampling limit in its temporal domain says that the imager sampling rate should be at least twice as fast as the sampled moving object otherwise signal distortion (or temporal aliasing) occurs. An example of motion distortion is shown in Fig 3.17 using a Kodak CMOS image sensor in rolling shutter mode [86] acquiring a moving bus. Rolling shutter is typically similar to the dual sampling technique in the way how the readout signal follows the reset signal with a predetermined time delay determining the integration time of the rolling segment. Once reaching the bottom of the imager, the rolling shutter continues with the top rows of the imager as if the imager was a cylinder. The main reason to use the rolling shutter is to separate the integration time from the frame time during high speed imaging. Note that plain raster scanning is just one extreme of the rolling shutter, where the shutter is fully 'open'.

Fig 3.17 Impact of rolling-shutter on CMOS sensor causing motion blur

The common solution for the motion distortion when imaging moving objects is by using a global shutter [84] and thus the imager (or a ROI) is reset synchronously and after an integration time period sampled synchronously as well. The sampled image is separated from the photosensitive area (ex. Photodiode) and stored into capacitor to be readout afterward sequentially. Fig 3.18 shows how the global shutter solve the motion distortion (a typical global shutter pixel circuit diagram is introduced in previous chapter in Fig 2.11). Despite the motion distortion is not far less in global shutter sampling (Fig 3.18.B) than in raster scan sampling (Fig 3.18.A), the motion blur is more noticeable in the former than in the latter as shown in the fan blades edges. This is because the pixels (rows) exposure to incident light is longer in the global shutter regime than in the raster scan.

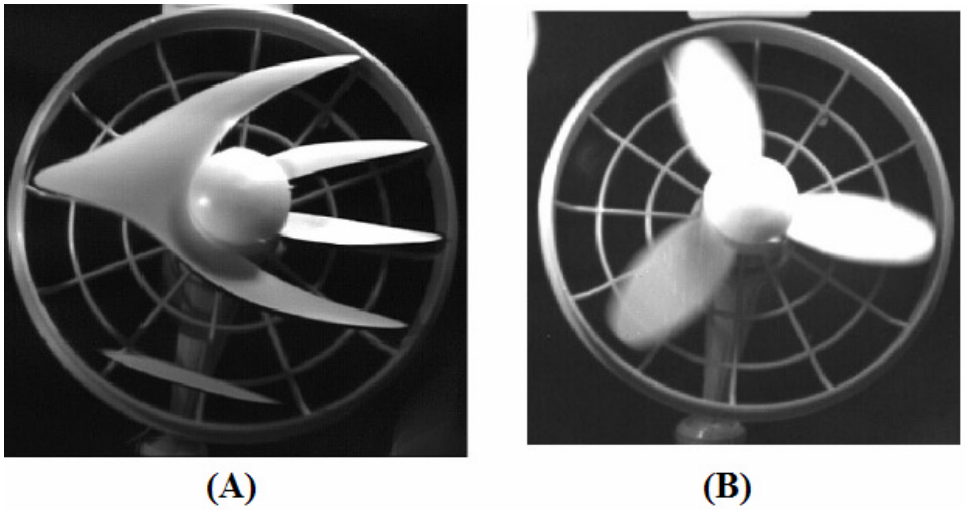


Fig 3.18 Global shutter (B) versus rolling shutter (A) and motion blur distortion [84]

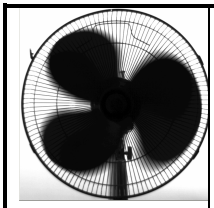
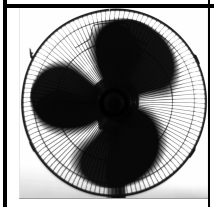
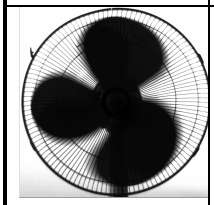
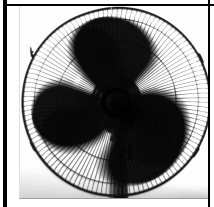
The main disadvantage with using global shutter in fast imaging (beside the shutter efficiency affected by the shutter leakage and the limited in-pixel sampling capacitance) is its limited application to still imaging because of its bottleneck in transferring a large amount of 2D data at high frame rates. This problem will be more pronounced with increasing resolution. Despite good improvement of the shutter efficiency reaching close to 100% which has been reported recently ([87], [88]), these solutions come at the expense of fill factor.

It is clearly visible from Fig 3.17, that the cause of the motion blur was the inequality in the spatial sampling rate between rows and columns which is faster in the latter than in the former. In raster scanning, the whole columns of a row are sampled simultaneously then readout serially before sampling the next row, which makes column sampling faster that row sampling in classical CMOS imager. Consequently, row motion blur is higher than that of the columns which result in a horizontal dislocation of the bus image in Fig 3.17.

The higher spatial sampling symmetry inherent in the pyramidal sensor serves to minimize the motion blur artifact or, at least, to distribute it evenly to minimize the artifact. It is also advantageous for architecture to have multiple output channels to increase the acquired image transfer speed throughput thereby reducing the need for local storage of the sampled image and eliminating the global shutter. The high speed property due to the 2D ring sampling property and the 8 output channels of the pyramidal architecture will be extensively studied in section 4.3 and 4.4. Here, we will investigate the pyramid scanning and its impact on minimizing the motion blur on the acquired image.

Shown in Table 3-3, are images acquired by a PixeLINK™ 1.3 Mega Pixel monochrome PL-A653 CMOS camera. The moving object is a fan moving at a peed of 1 rotation per second and lit from behind. The imager samples a window of 784x784 pixels with a programmable exposure (or integration) time of 43.2ms and with a resulting frame rate of 6 frames per second (fps) or 1 frame every 166.67ms.

Table 3-3 Standard CMOS image sensor raster scanning motion blur demonstration.

	**Fan leaf passing by the right side clockwise. (a)**
	**Fan leaf passing by the bottom side clockwise. (b)**
	**Fan leaf passing by the left side clockwise. (c)**
	**Fan leaf passing by the top side clockwise. (d)**

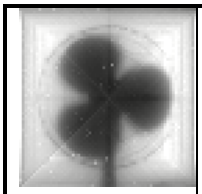
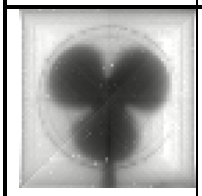
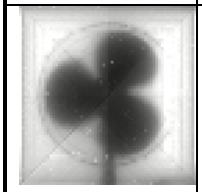
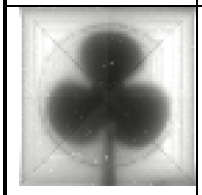
From the images in Table 3-3, the motion blur is more prominent on the fan leaf passing by the right and left side of the fan as seen by the deformation of the fan leaf form. This blur is less noticeable when the fan leaf passes by the top and bottom of the fan. Moreover, the motion blur is larger when the leaf passes by the right side and shorter when it passes by the left side, thus deforming the leaf appearance. This is explained by the fact that, at the focal plan when the moving object image is moving on the same direction of vertical sampling (row sampling), the relative speed of the object to the imager sampling direction is small thus smearing the integrated image in ongoing direction. This effect fact is noticeable with the image shown in Table 3-3.a. This fact demonstrates that the PixeLINK™ CMOS imager is scanning the rotating fan from top to bottom. In the other case when the object image is moving opposite to the vertical sampling direction, the relative speed of the object looks faster with respect to the sampling scan, and hence some portion of the moving image will not have enough time to be integrated at the pixels (rows) and the resulted acquired image will look smaller. This fact is clearly shown in the figure at Table 3-3.c where the fan leaf is moving clockwise from the bottom to the top. Finally, the images Table 3-3.b and Table 3-3.d suffered less from the motion blur as the image of the moving object is moving orthogonal to the row sampling direction.

Corresponding images acquired using the pyramidal image sensor under the same lighting conditions and fan speed as the previous test are shown in Table 3-4. The integration time the 64x64 pyramidal CMOS imager was 43.24ms, a difference of 0.04ms from the previous imager exposure time.

The images acquired with the pyramidal CMOS image sensor shows almost unnoticeable motion blur artifact. The is the resulting effect of the ring sampling in which the fan leaf image at the imager focal plane is moving orthogonal to ring sampling direction. The moving object is moving circular in our test case, however for any arbitrary movement direction; one can see the limitation of the architecture. In fact, the pyramidal ring sampling performance is still achievable for any arbitrary moving object as long as the horizontal segments of the pyramidal imager are used for sampling vertically moving objects and the vertical segmented are used for the horizontally moving objects acquisition. However, this scenario could not always be verified as no control over moving imaged objects is always possible.

Despite the above limiting scenarios, the pyramidal imager can still acquire images with minimal motion blur and moving objects distortion. At the inner rings area (fovea) the horizontal and vertical sampling speeds get closer and closer making the sampling at the fovea region of the imager similar to a global shutter regime. Therefore, the pyramidal imager fovea is a motion-blur and motion-distortion free sampling zone adding a new foveated attribute to the pyramidal imager architecture beside the foveated dynamic range.

Table 3-4 Pyramidal CMOS image sensor ring scanning motion blur demonstration.

	**Fan leaf passing by the east side. (a)**
	**Fan leaf passing by the south side. (b)**
	**Fan leaf passing by the west side. (c)**
	**Fan leaf passing by the north side. (d)**

3.5.10 Hardware Cost Scaling of the Pyramidal Imager Design

One of the main differences between the pyramidal and classical CMOS imagers is the hardware and especially regarding the sample and hold blocks. These needed blocks are used to sample the 2D rings in the former and to sample the 1D row sampling in the latter. Beside these extra blocks, their control logic and output buffers also come as an extra hardware cost. For the sake of simplicity the logic blocks and output buffers are neglected compared to the relatively large sampling blocks.

From the first glance, it is obvious that the pyramidal CMOS imager sampling blocks are three times larger than that of the classical CMOS imager. With increasing resolution and larger pixel array sizes the hardware cost ratio till remains. However, this cost ratio of the sampling circuitry decreases quickly with increasing array size with respect to the imager pixel array. To clarify this fact, a simple analysis will be demonstrated for clarification.

Let us assume a classical CMOS imager of $2R$ x $2R$ pixel resolution (pixel array size), thus a pyramidal imager of the same resolution would be of R ring size. Consequently, the classical imager needs a 2 R sample and hold blocks whereas the pyramidal imager needs 8 R sample and hold blocks. This hardware cost estimation is based on the description of both architectures in Chapter 3. The total area occupied by the S&H blocks in the pyramidal imager is $\text{Pyramid\_S\&H}_{area} = 8\ R\ \text{S\&H}_{area}$ and the that occupied in classical imager is $\text{Classic\_S\&H}_{area} = 2\ R\ \text{S\&H}_{area}$, were $\text{S\&H}_{area}$ is the silicon area used to design a single S&H block. For both architectures the pixel array occupies $\text{Pixel\_array}_{area} = 4\ R^2\ \text{Pixel}_{area}$ of the silicon area where Pixel_{area} is the size of a single pixel. Finally, the ratio between the areas of the sampling blocks and the sensing array will be $Pyramid\_Area\_Ratio = \frac{2}{R}\frac{S\&H_{area}}{Pixel_{area}}$ and $Classic\_Area\_Ratio = \frac{1}{2R}\frac{S\&H_{area}}{Pixel_{area}}$ for the pyramidal and classical imager respectively. Fig 3.19 shows the profile of both ratios versus parameter R.

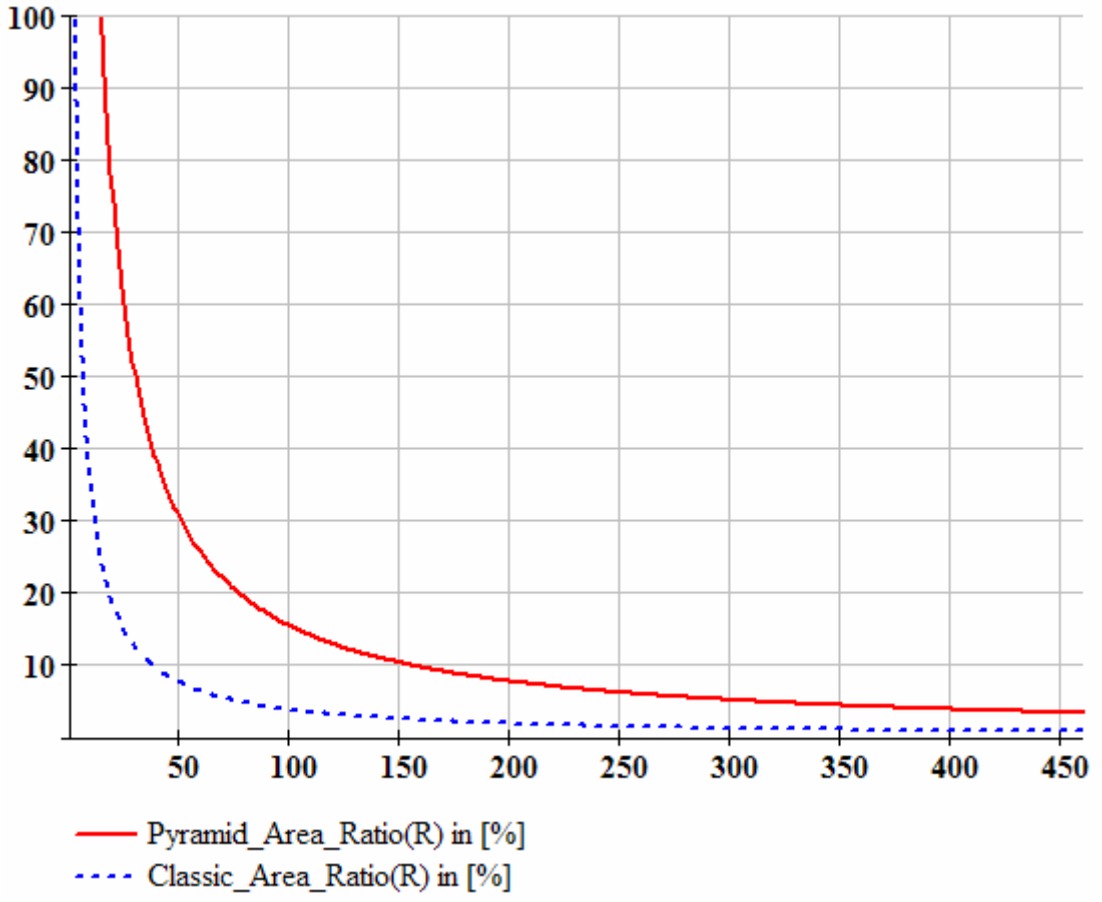


Fig 3.19 S&H hardware cost ratios of the classical and pyramidal imagers

It is visible that the difference of the S&H block area ratio between the pyramidal and the classical imager for smaller pixel array sizes (or low resolutions). However, this difference decreases with increasing resolutions. In fact, at 1 Mega pixel resolution (1024x1024) the S&H block area ratios with respect to pixel array is 1.5% and 0.4% for the pyramidal and classical imagers respectively.

In conclusion, with regular resolutions (of Mega pixels ranges) the hardware cost of the pyramidal CMOS imager with respect to the classical CMOS imager becomes insignificant.

3.6 Multiresolution Imager Design

The Multiresolution CMOS imager design is based on a bottom-up approach. The new architecture is proposed to implement the multiresolution mechanism at the pixel level in order to achieve the expandability and programmability properties that current multiresolution CMOS imagers are lacking ([89], [90]). In the suggested multiresolution architecture, the classical CMOS imager architecture is used and instead of using the classical active pixel sensor (APS) architecture a new pixel design is suggested. The multiresolution pixel (MP) has the ability, beside the photo-electric conversion and sampling to share its stored photo-charge with its next in-row and its next in-column neighbour pixels through a simple MOSFET switches network. Further details are developed in Chapter 7.

3.7 Summary

In this chapter two different implementation of foveated CMOS image sensors were discusses. The time domain foveated CMOS imager was designed in a top down approach as it deals with the sampling architecture at the focal plane level. The spatial domain foveated CMOS imager that is the multiresolution imager (detailed in Chapter 7), on the other side, was designed in bottom up approach. This is because it deals with the resolution management (programming resolution down-scale) at the pixel level. Both imagers benefited from the dual-voltage feature of CMOS 0.18µm technology. 3.3V supply voltage has been used in the

APS pixels and their analog chain (until the output buffer) while 1.8V supply voltage has been used in the controlling digital blocks. This is mainly due to the high gate leakage current of the thin oxide transistors used only in digital (gate-leakage immune) blocks.

Finally, it has been shown that the extra hardware cost ratio (with respect to the pixels array) due to the extra S&H blocks in the pyramidal imager compared to that of the classical imager decrease with increasing imager resolution.

Chapter 4

Mathematical Basis of Foveated Dynamic Range & High Speed Imaging

4.1 Mathematical Analysis of Foveated Dynamic Range

In this section we will present the mathematical foundations and aspects of the proposed Foveated Dynamic Range (FDR) enhancement. It is important to note that this chapter is not including, unless explicitly mentioned, any non-ideal physical parameters of CMOS image sensors such as limited electrical dynamic range, dark current, temporal and spatial noise. In the following sections an ideal linear, noiseless CMOS image sensor will be assumed. We will include some physical limitations towards the end of the chapter in order to evaluate the physical reality of our mathematical claims.

4.1.1 Timing examination of bouncing scanning:

Firstly, the parameters for assessing the integration time of sampled rings in bouncing scanning schemes are discussed. The readout process can be divided into two steps; the row sampling step that takes a period of time T_{spl} (in the range of 1μs to 10μs), and the scanning step to readout serially the sampled pixel values T_s (is in the range of 0.1μs to 1μs) [68]. The sampling period T_{spl} is the time necessary for three operations: the sampling of the ring's photo-signal output, *Vs,* to the corresponding capacitors in the sample and hold bank, the

resetting of the ring to V_{dd} and finally the sampling of ring's reset voltage V_r to the corresponding capacitors in the sample and hold bank.

In the pyramidal architecture, the rings contain different numbers of pixels. We take the inner ring to be the reference ring, or r =1, the subsequent ring is r = 2, and so on. Secondly, we assume the output image is built from the output channels of the sensor's eight symmetrical segments. Therefore, the analysis of the rings' integration time can be reduced to the analysis of any one of the pyramid segments instead of the entire sensor. In other words, since the pyramidal imager has eight output channels read out in parallel, frame scanning is based on the segment timing. In the following development we will focus on one segment for which (from Fig 4.1), there is one pixel in ring r = 1, two pixels in ring r = 2…etc. Hence for the ring of order r there are r pixels and the last ring (outer ring) which is of order N/2 there are N/2 pixels, where N is the square root of the resolution of the imager, assumed here to be square.

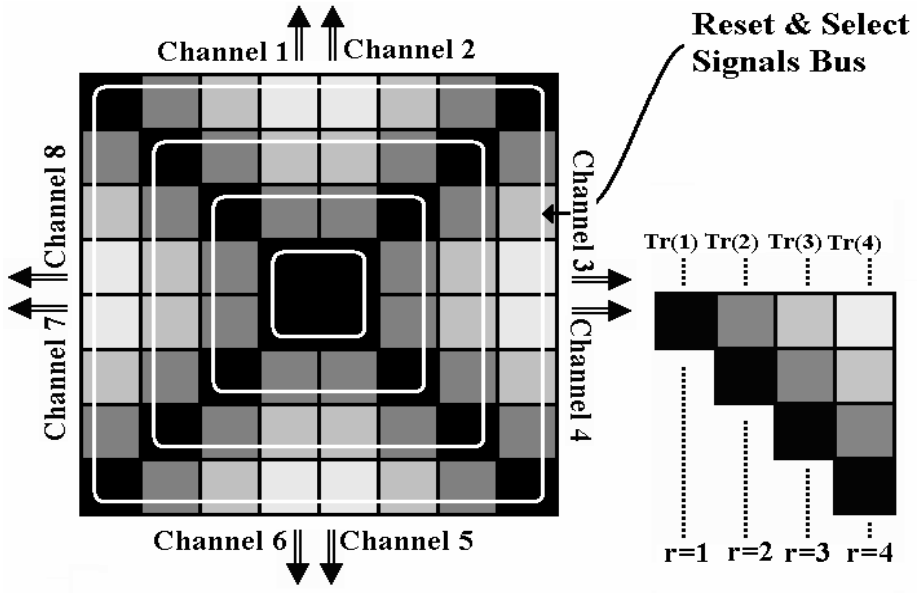


Fig 4.1 Pyramidal imager readout scheme

Consequently, the time required to read a pyramid ring is:

$$Tr(r) = rT_s + T_{spl} \ldots (4.1)$$

For the sake of illustration, the timing diagram of the inner ring (r = 1) and outer ring (r = 4) of a 4x4 pyramidal imager using bouncing scanning scheme for image sampling is shown in Fig 4.2. The inner ring and outer rings are the bouncing scanning edges where scanning direction changes from inward (imager center) scanning to outward (image center) scanning or vice versa. The timing diagram in Fig 4.2 will be useful in calculating the integration time for each ring in the pyramidal imager when bouncing scanning is used.

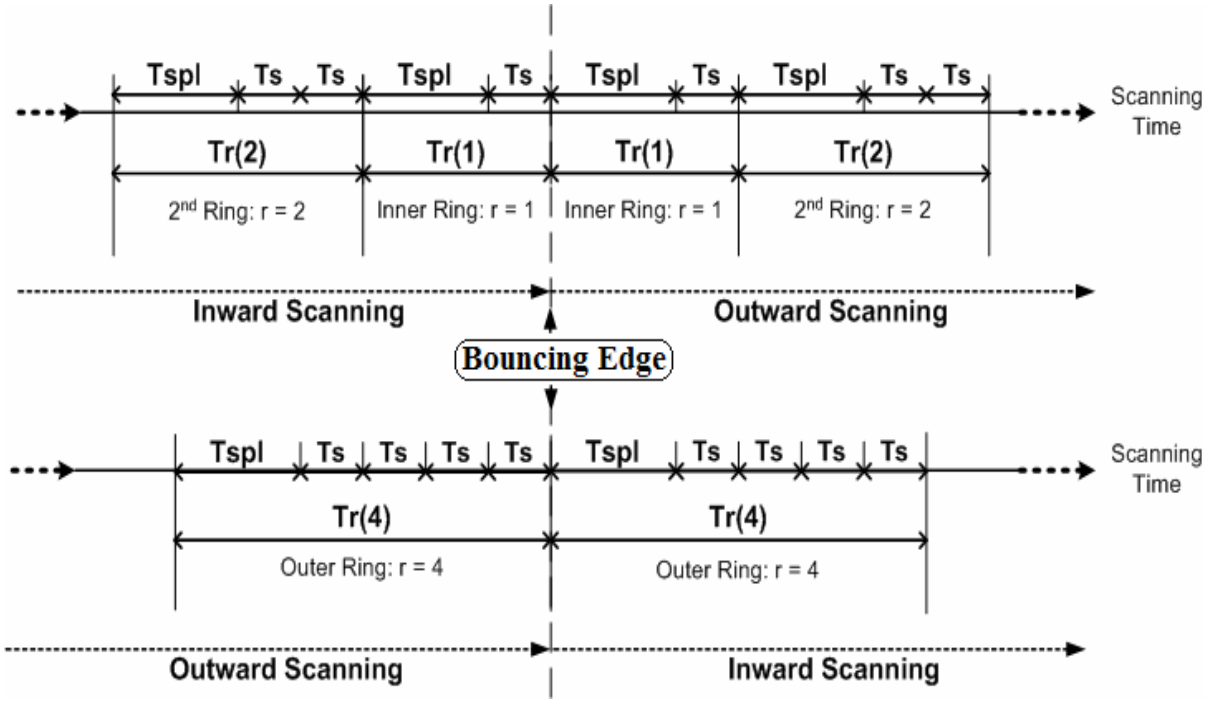


Fig 4.2 Timing diagram of inward and outward scanning

4.1.2 Mathematical Analysis

The integration time for a given ring is the period of time between two consecutive samplings of that ring. At each sampling, the ring photo-signal is sampled into one of the two capacitors of the CDS circuit, then the ring is reset and sampled again into the other capacitor. All these three operations are made during T_{spl}. Using the timing diagram shown in Fig 4.2 a mathematical formulation for integration time for both inward scanning $Tin(r)$ and outward scanning $Tout(r)$ will be derived. Because each ring is sampled one time by inward scan and the next time by outward scan (or vice versa), the ring integration time will be equal to the summation of the time required to read the read the current ring pixels namely ($r\ T_s$)

plus twice the time needed to scan from current ring r until the outer ring $r = N/2$. This result is shown in equation 4.2.

$$Tin(r, R, T_s, T_{spl}) = 2\left[\sum_{\substack{i=R \\ i \to i-1}}^{r+1} iT_s + (R-r)T_{spl}\right] + rT_s \ldots (4.2)$$

Similarly, the integration time $Tout(r)$ will be equal to the summation of the time required to read the current ring pixels namely ($r\ T_s$) plus twice the time needed to scan from current ring r until the inner ring $r = 1$. The final result is shown in equation 4.3.

$$Tout(r, T_s, T_{spl}) = 2\left[\sum_{\substack{i=1 \\ i \to i+1}}^{r-1} iT_s + (r-1)T_{spl}\right] + rT_s \ldots (4.3)$$

The expansion of the equations (4.2) and (4.3) result in the following equations (4.4) and (4.5) respectively.

$$Tin(r, R, T_s, T_{spl}) = -T_s r^2 - 2T_{spl} r + \left(R^2 T_s + RT_s + 2RT_{spl}\right) \ldots (4.4)$$

$$Tout(r, T_s, T_{spl}) = T_s r^2 + 2T_{spl} r - 2T_{spl} \ldots (4.5)$$

Fig 4.3 plots both integration time profiles for 32 rings (a 64 x 64 pixel array) and timing parameters $T_{spl} = 3T_s = 3\mu s$ based on the ranges mentioned in the previous section.

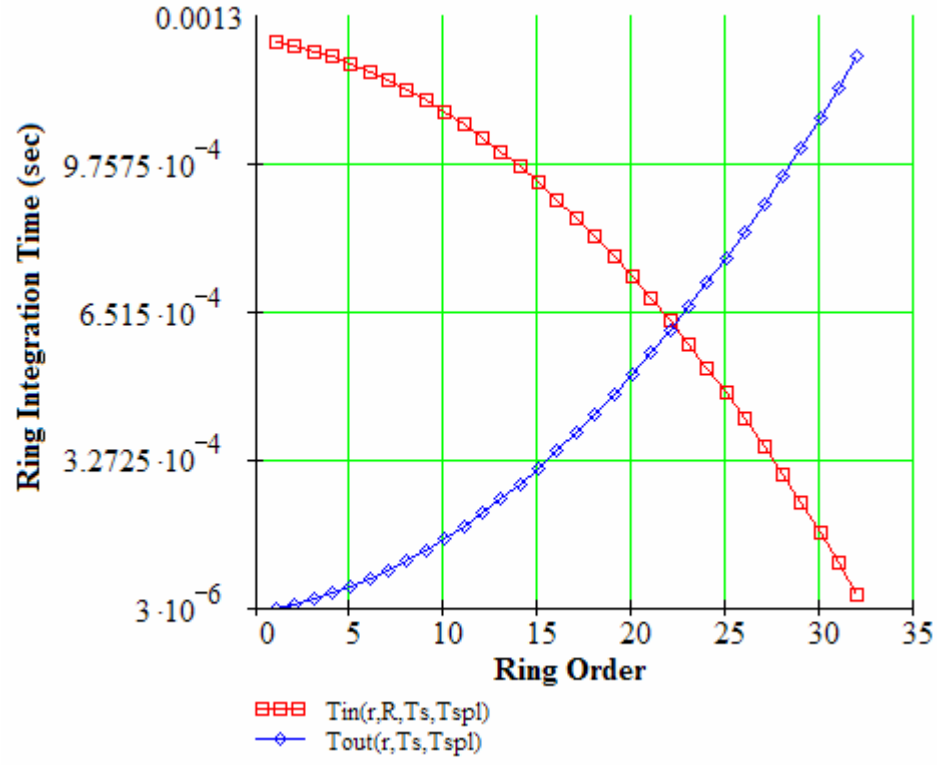


Fig 4.3 Plot of the rings inward and outward scanning integration time

From the above figure and for each ring we have different integration times *Tin*(*r*) and *Tout*(*r*) for the same ring scanned in two directions. An inward scan followed by an outward scan will result in two frames. The first frame would have, for an intermediate light intensity, the brighter areas in the inner rings which result from the inward scanning or *Tin*(*r*) profile. The second frame will have the brighter areas in the outer rings resulting from outward scanning or *Tout*(*r*) profile. This is apparent from the above figure since the maximum integration time for *Tin*(*r*) is at the inner rings whereas Tout(r) has its highest values at the outer ring. In the linear response regime of the pyramidal CMOS imager, the fusion of

inward and outward scan frames will lead to a uniform response. In fact, in the linear response regime of the imager, the sampled photo signals are linearly proportional to the integration time by a constant light-independent, technology dependant parameter called the sensitivity. Hence, fusing the generated inward and outward frames will result an image that is linearly proportional to the sum $Tout(r) + Tin(r)$ by sensitivity parameter. Using equations 4.4 and 4.5 we get:

$$Tin\left(r, R, T_s, T_{spl}\right) + Tout\left(r, T_s, T_{spl}\right) = R^2 T_s + R\left(T_s + 2T_{spl}\right) - 2T_{spl} \ldots (4.6)$$

The summation in equation 4.6 represent the equivalent integration time of the fused image when sampled without bouncing scanning. Note that the fused image integration time is independent of r, in the linear regime, which proves the uniformity of the fused image as shown in Fig 4.4.c. The two diagonal artifact lines in Fig 4.4 are clarified in section 3.5.2.

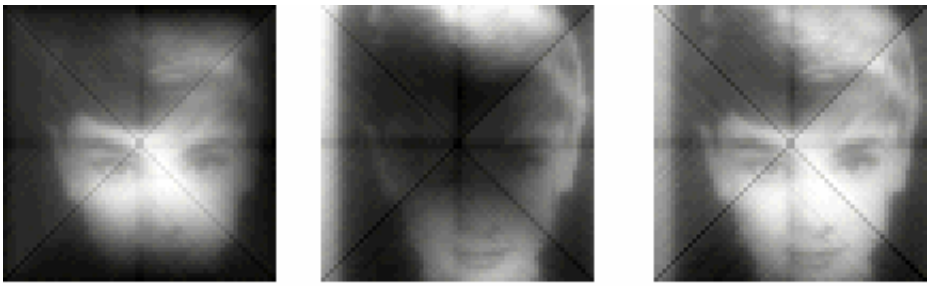


Fig 4.4 Bouncing scanning in linear regime: (a) Inward, (b) Outward, (c) Fused images

Fig 4.3 is reproduced in 3D in Fig 4.5.

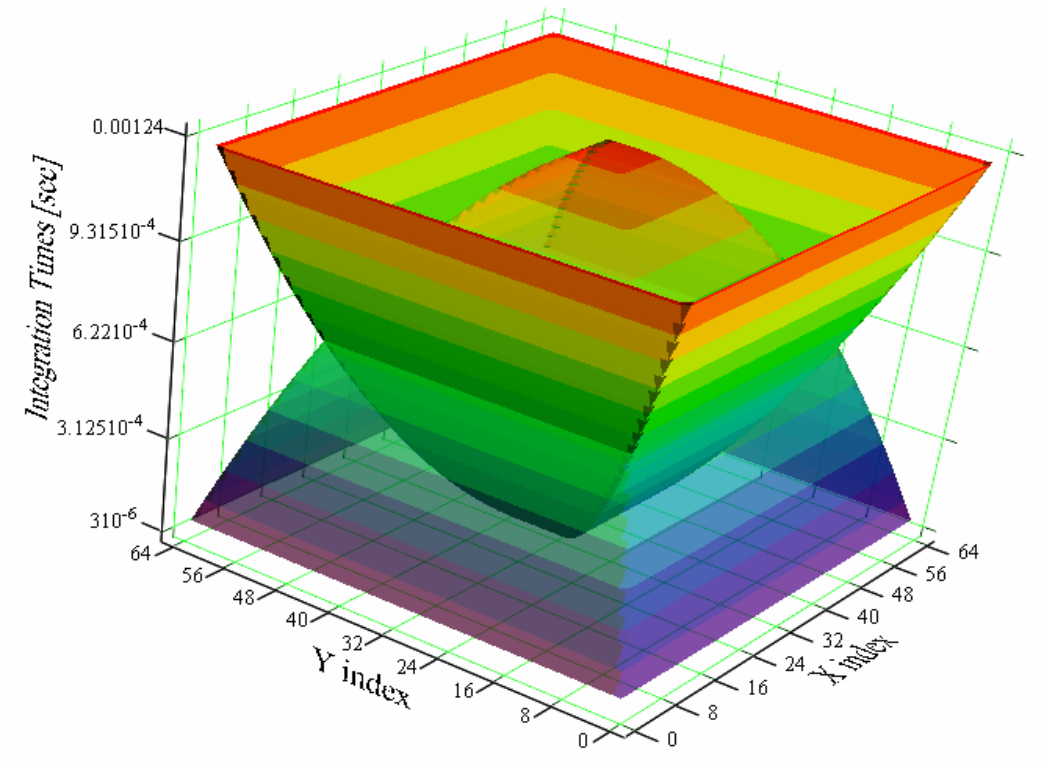


Fig 4.5 3D view of the rings inward and outward scanning integration time profiles

Fusion (by summation) of the resulting images of the two integration time profiles *Tin*(*r*) and *Tout*(*r*) of inward and outward scanning respectively we get a dynamic range enhancement DR_{enh}(r) (in decibels) for each ring based on the following formula [68]:

$$DR_{enh}(r) = 20\log\left[\frac{\max\left(Tin\left(r,R,T_s,T_{spl}\right),Tout\left(r,T_s,T_{spl}\right)\right)}{\min\left(Tin\left(r,R,T_s,T_{spl}\right),Tout\left(r,T_s,T_{spl}\right)\right)}\right] \ldots (4.7)$$

Because the dynamic range enhancement was a result of fusion between two frames, or scenes, the enhancement is called intrascene dynamic range enhancement. The DR profile is show below for R=32, $3T_s = T_{spl} = 3\mu s$.

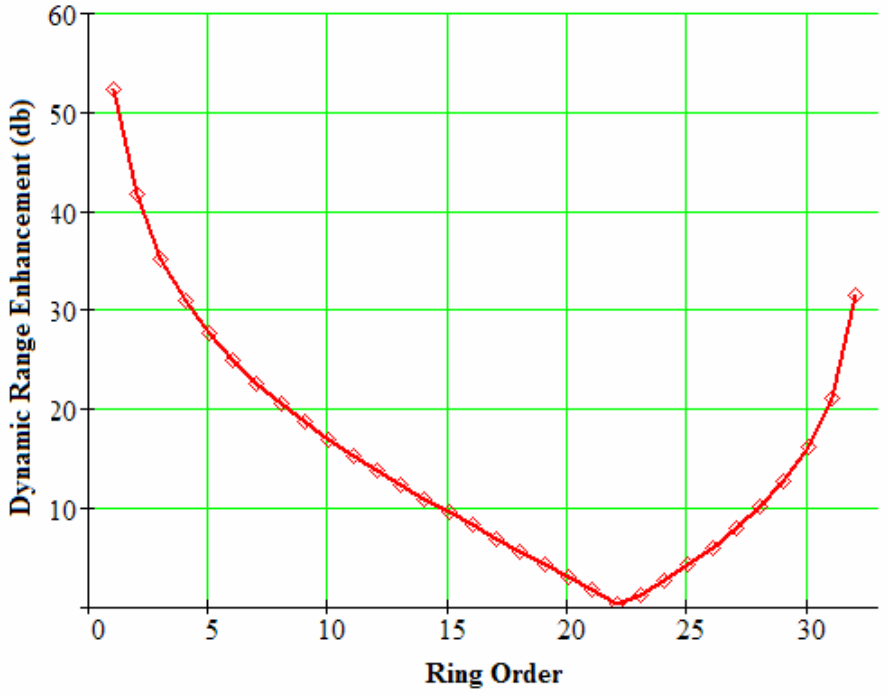


Fig 4.6 Intrascene foveated dynamic range enhancement

The above figure is shown in 2D in the Fig 4.6. The dynamic range enhancement is higher in the inner rings than in the outer ring especially from ring 1 to ring 5 and hence the naming of foveated dynamic range enhancement (FDR).

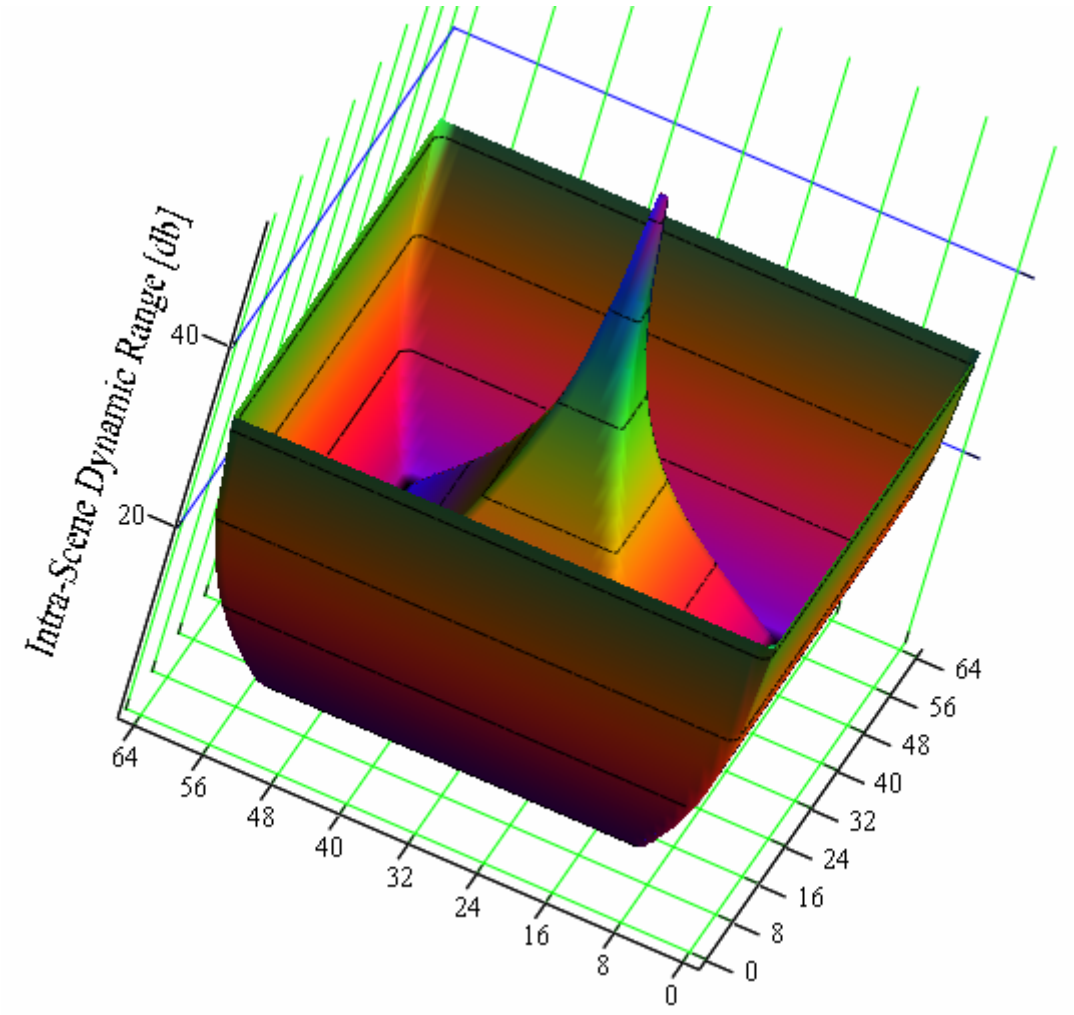


Fig 4.7 Intrascene foveated dynamic range enhancement in 3D view

The minimum of the FDR enhancement coincides with the ring order at which the ring's integration time is equal for both the inward scanning and the outward scanning and thus the dynamic range enhancement is nil. This ring in question is deduced from Fig 4.6 to be ring number 23. From ring 23 until the last ring the foveated dynamic range enhancement increases. Two questions now arise. Can we predict *a priori* the ring at which the FDR

enhancement is zero and after which FDR increases? The second question is how to control the imager in order to make the ring of zero-dynamic range enhancement the last ring in the pyramidal image sensor. In other words how can we achieve monotonic FDR enhancement for the entire sensor, i.e. how to pin down the minimum of the FDR enhancement at the outer ring of the pyramidal imager instead of having this DR enhancement minimum at ring 23? The second question will be answered in the next section. To answer the first question, we equate equations (4.4) and (4.5) in order to find the minimum that we will call the fovea border. Equating the previously mentioned equations and extracting the parameter *r* the fovea border ring order *FOVborder*, we get the following equation.

$$FOVborder(\alpha, R) = \frac{1}{\alpha}\left[\sqrt{1+\frac{1}{2}\alpha^2 R(R+1)+\alpha(R+1)}-1\right], \text{ where } \alpha = T_s/T_{spl} \quad \text{... (4.8)}$$

The *FOVborder* is plotted in Fig 4.8 below as a function of $\alpha = T_s/T_{spl}$.

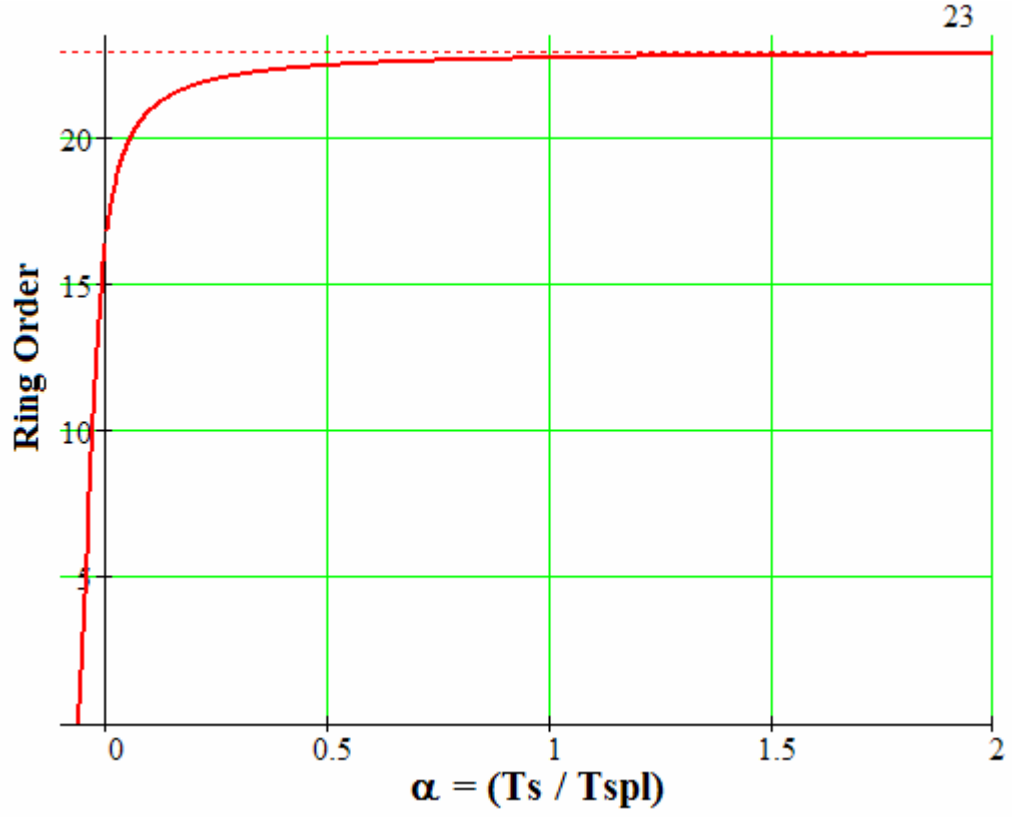


Fig 4.8 Foveated dynamic range border limit for a 32 ring pyramidal imager

It is clearly noticeable from the above graph that the fovea border approach asymptotically ring $r = 23$ when α approaches infinity. In fact, the ring order of the fovea border approaches 23 even at as low values of α as 0.5. It is clear that we are interested only on the positive values of α as T_s and T_{spl} are positive values.

To prove the above limit, let us examine the limit of equation (4.8) when $\alpha \to \infty$ which is shown in the below equation:

$$\lim_{\alpha\to\infty} FOVborder(\alpha, R) = \lim_{\alpha\to\infty} \frac{1}{\alpha}\left[\sqrt{1+\frac{1}{2}\alpha^2 R(R+1)+\alpha(R+1)}-1\right] = \frac{R}{\sqrt{2}}\sqrt{1+\frac{1}{R}} \quad \ldots (4.9)$$

As $R >> 1$ the limit in equation (4.9) tends to:

$$\lim_{\alpha \to \infty} FOVborder(\alpha, R) = \frac{R}{\sqrt{2}} \ldots (4.10)$$

Hence, the fovea border does not depend on α but depend linearly on the imager ring size with a slope of $1/\sqrt{2} \cong 0.71$. In other words, the fovea border of our system is at about 71% of the ring size of the pyramidal imager. This explains why the fovea border for a 32 ring pyramidal imager tends to 23 as 71% of 32 is equal approximately 22.66 or ring number 23.

An interesting result of the above study of the fovea border would be to make the fovea border coincide with the outer ring of the pyramidal imager. This will clean the extra tail of the dynamic range enhancement after the minima as stated at the beginning of this chapter.

4.1.3 The Control of the Fovea Border

The dynamic range fovea border can be pinned down on the edge of the pyramidal imager by scanning the imager (using the bouncing scanning scheme) and assuming that that imager is virtually of a larger dimension. This is realized by adding an extra readout period at the end of the outward scanning. The size of the virtual pyramidal imager would be of ring size equal to $\hat{R} = R\sqrt{2}$ because fovea border limit will be, as demonstrated above, at $\hat{R}/\sqrt{2} = R$. This virtual ring extension will affect only the inward scanning ring integration time, which is a function of *R* that is the ring size of the pyramidal imager. The outward scanning integration time is not affected because it is not dependent on the number of imager rings. This is

obvious because the inward scanning is limited to the total number of rings as it is bounded by the inner ring of pyramidal imager, whereas the outward scanning is not so limited.

The method described above is equivalent to adding an amount of time we call *Txd* to the inward scanning to make it equal to the outward scanning at the outer ring of the pyramidal imager. After reaching the outer ring of the pyramidal imager at the end of outward scanning, a period of time is spent before starting scanning inward to the imager inner ring. This is the practical side of the previous method, but how much is this period of time that we called *Txd*?

In order to extract mathematically the value of *Txd* and examine its dependencies, the expression of rings' integration time for the inward scanning will be written assuming the extended virtual pyramidal imager (similar to *Tin*(*r*) in equation 4.4) and we call it *TinVPyr* by replacing R by $R + R_{\Delta}$ in equation (4.4). We will get as a result equation (4.10) shown below.

$$TinVPyr\left(r, R, T_s, T_{spl}, R_{\Delta}\right) = -T_s r^2 - 2T_{spl} r + \left(\left(R + R_{\Delta}\right)^2 T_s + \left(R + R_{\Delta}\right) T_s + 2\left(R + R_{\Delta}\right) T_{spl}\right) \quad (4.11)$$

In equation (4.11), R_{Δ} is the extra rings needed to reach the virtual pyramidal imager' ring size as shown below.

$$R_{\Delta} = R\sqrt{2} - R = R\left(\sqrt{2} - 1\right) \ldots \quad (4.12)$$

Next, we will use the mathematical formulation of the inward scanning integration time *Tin*(*r*) as introduced in equation (4.4), this time without assuming the virtual pyramidal imager but instead increasing *Tin* by *Txd,* as initially proposed. We get the following equation

$$TinTxd\left(r, R, T_s, T_{spl}, Txd\right) = Tin\left(r, R, T_s, T_{spl}\right) + Txd \quad \ldots (4.13)$$

Equating equations (4.11) and (4.13), we get the expression of *Txd* and call it *Txd_approx*.

$$Txd\_approx\left(R, T_s, T_{spl}\right) = R^2 T_s + R\left(\sqrt{2} - 1\right)\left(T_s + 2T_{spl}\right) \quad \ldots (4.14)$$

It is apparent that *Txd_approx* is a function of the imager ring size and sampling parameters T_s and T_{spl}. The above equation of *Txd_approx* is not the exact value of *Txd* because it has used the approximation of the limit in equation (4.11). The exact value of *Txd* is determined by equating equation (4.13) and (4.5) at the outer ring (*r* = R) and extracting the exact value of *Txd* that we call *Txd_exact*. The extracted value of *Txd_exact* is shown in the following equation (4.15).

$$Txd\_exact\left(R, T_s, T_{spl}\right) = R^2 T_s + \left(2T_{spl} - T_s\right)R - 2T_{spl} \quad \ldots (4.15)$$

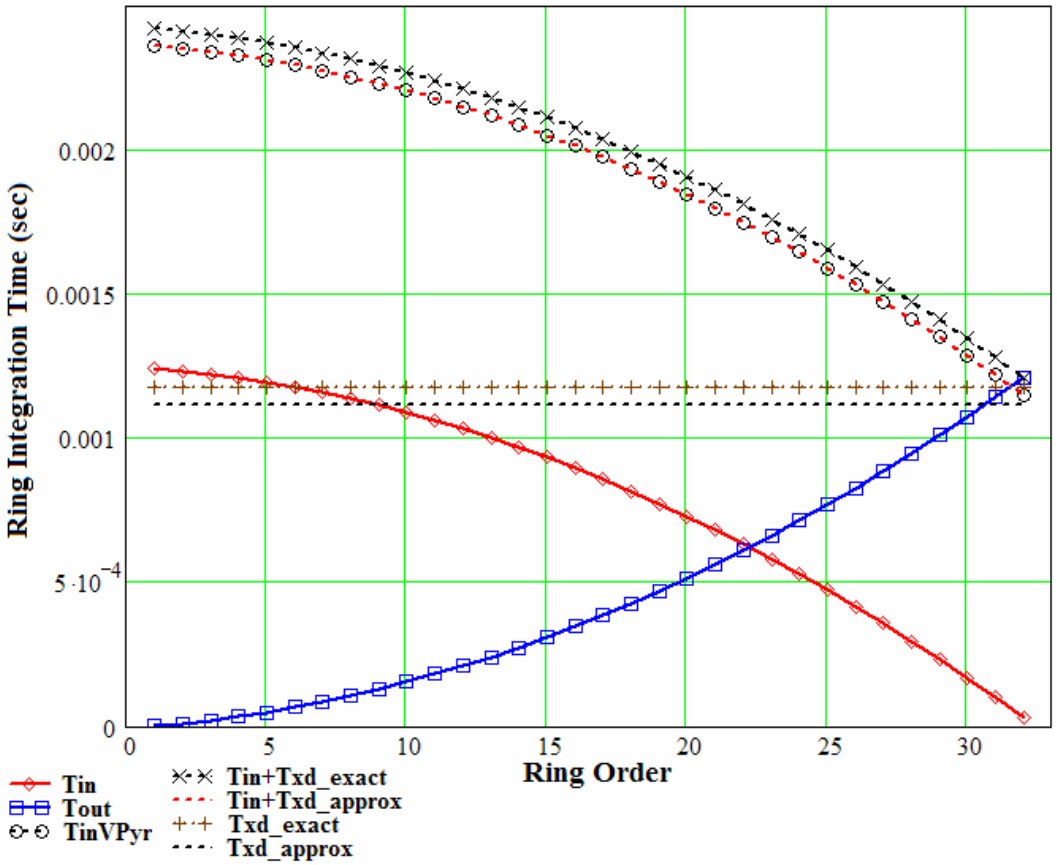


Fig 4.9 Foveated dynamic range border control technique

From the Fig 4.9, we can see that the introduction of *Txd* has just shifted up the inward scanning ring integration profile *Tin*(*r*, *R*, *Ts*, *Tspl*) so that the inward and outward integration time profiles cross each other at the outer ring where the FDR border will be pinned as shown in Fig 4.10 below. In Fig 4.9 also we can see that *TinVPyr* (see equation 4.11) coincides with *TinTxd* (equation 4.13) when *Txd* is replace with its approximate value (equation 4.14). When *Txd* is replaced with its exact value (equation 4.15) in the expression of *TinTxd* the later coincides exactly with *Tout*(*r*) at the last ring $r = 32$ in our pyramidal

imager whereas *Tout*(*r*) coincides with *TinVPyr* at a ring earlier $r = 31$ due to the latter approximate approach. Although, the approximate approach has been used to introduce the concept of the virtual pyramidal imager, however, the exact value *Txd_exact* (equation 4.15) will be used in the subsequent sections. It is visible from Fig 4.9 that the introduced time delay to pin FDR border at the outer ring is approximately one frame time and thus halving the frame rate of the pyramidal imager. This impact will be further estimated in section 4.5.

In Fig 4.10 we show the foveated dynamic range enhancement, we call *FDRenh*, of the pyramidal imager as plotted in Fig 4.6. Also shown in the same figure is dynamic range enhancement calculated based on the same parameters used in Fig 4.10 but with the inward scanning integration time *Tin* is replaced by *TinTxd* mentioned in equation 4.15. We will call this DR enhancement *FDRenhTxd*. The figure shows clearly that the fovea border of the *FDRenhTxd* is at the edge of the pyramidal imager. Also shown in the same graph, the difference *FDRenhTxd-FDRenh* which shows that *FDRenhTxd* is over most of the rings is greater than *FDRenh*. A numerical integration of *FDRenhTxd* and *FDRenh* results 588 dB and 498 dB respectively which shows that *FDRenhTxd* is higher that *FDRenh* over all the imager.

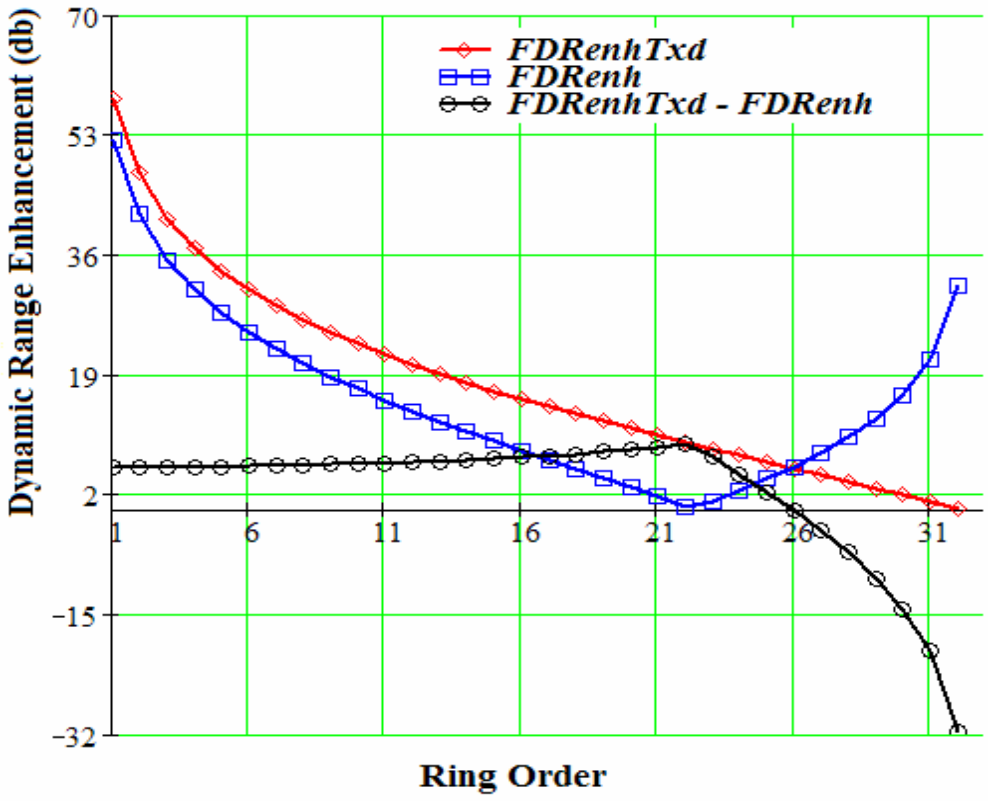


Fig 4.10 Foveated dynamic range enhancement after border pinning

Fig 4.11 represents a 2Ddisplay of the ring integration times with inward scanning correction *TinTxd* along with the outward scanning *Tout*(*r*). The corresponding intrascene foveated dynamic range enhancement *FDRenhTxd* is sketched in 2D as shown in Fig 4.12. The cost of achieving the dynamic range in a foveated form is the extra time delay *Txd* injected on *Tin*(*r*) which will impact the frame rate of the pyramidal imager. This timing cost will be revisited and estimated at the end of the present chapter.

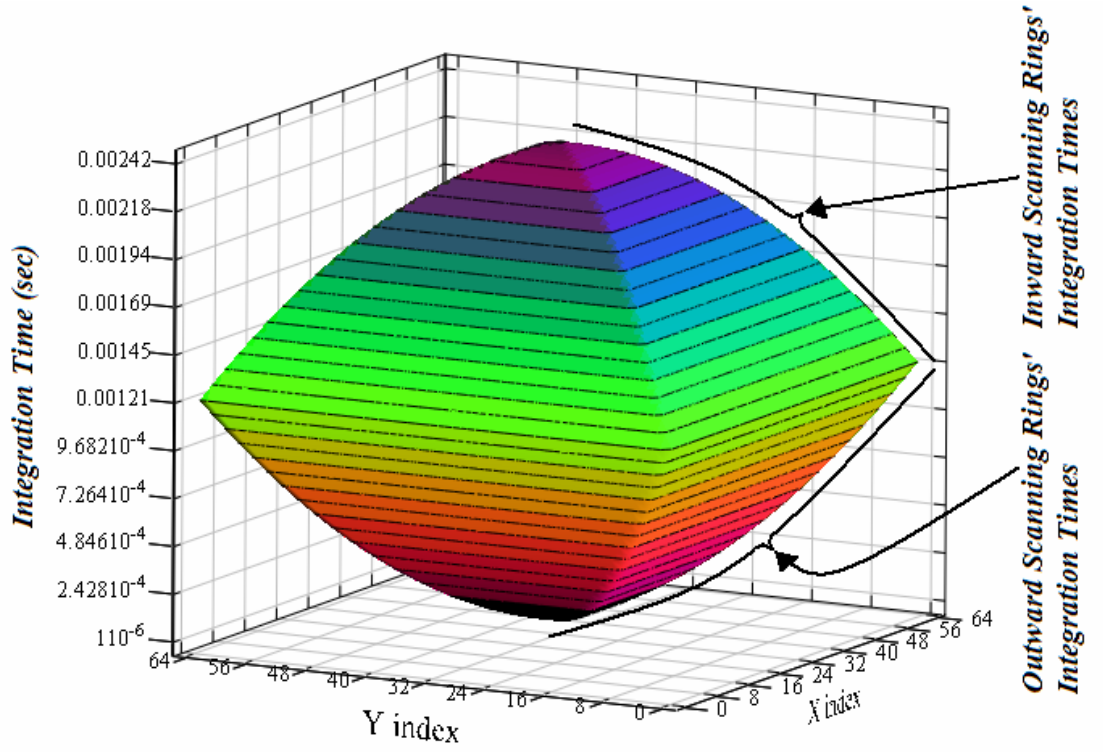


Fig 4.11 Bouncing scanning integration times after FDR border pinning in 3D view

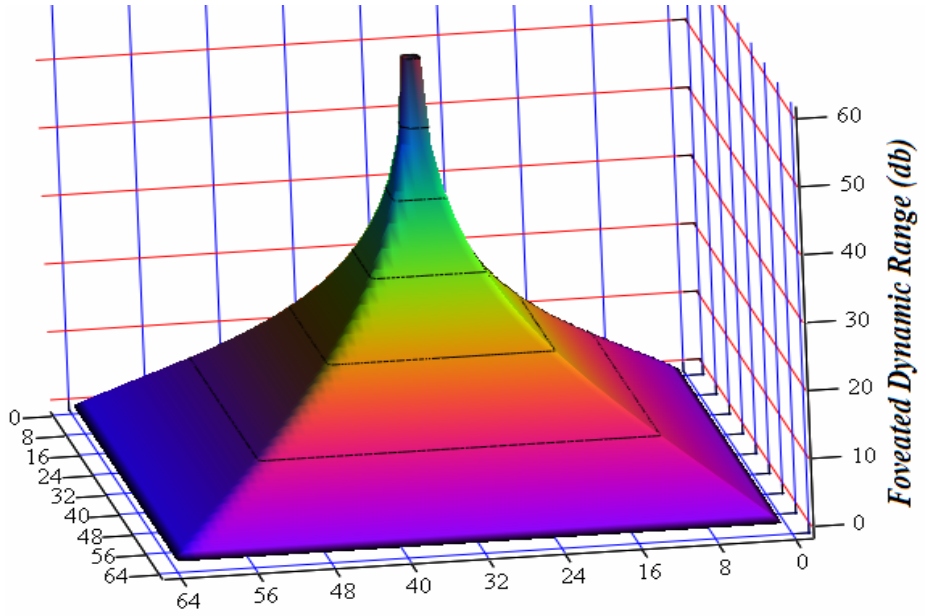


Fig 4.12 Foveated dynamic range enhancement after border pinning in 3D view

4.1.4 Inverse-Foveated Dynamic Range

In the previous section, we have manipulated scanning timing for the sake of pinning the foveated dynamic range border at the outer ring of the pyramidal imager by extending (or up-shifting) the rings inward scanning integration times. This manipulation is equivalent to breaking the bouncing scan (outward scanning) and the inward scanning at the outer ring (see Fig 4.2) and adding a delay time equal to *Txd between* them before closing it up again (i.e. repeat the scan for next frame). Effectively, the manipulation was based on an "assumed" bouncing at a virtual pyramidal outer rings and scanning back to reach the real outer ring of the existing imager.

Now, what will happen when the outward scanning integration time is manipulated the same way the inward scanning was manipulated? To answer this question let us examine the plot shown in Fig 4.9. The alternative manipulation is realised by up-shifting the outward scanning integration time so that the inward and outward scanning coincide at the inner ring. The amount of time *ToutTxd* needed to be added to *Tout(r)* to achieve this up-shifting is extracted in the following:

$$ToutTxd\left(R,T_s,T_{spl}\right)=Tin\left(1,R,T_s,T_{spl}\right)-Tout\left(1,T_s,T_{spl}\right),$$

Hence, we get:

$$ToutTxd\left(R,T_s,T_{spl}\right)=T_sR^2+\left(T_s+2T_{spl}\right)R-2T_s-2T_{spl} \quad \ldots (4.16)$$

The result of the outward scanning integration time up-shifting is shown in the plot of Fig 4.13. The new outward scanning integration time coincide with the inward scanning integration time at the inner ring $r = 1$.

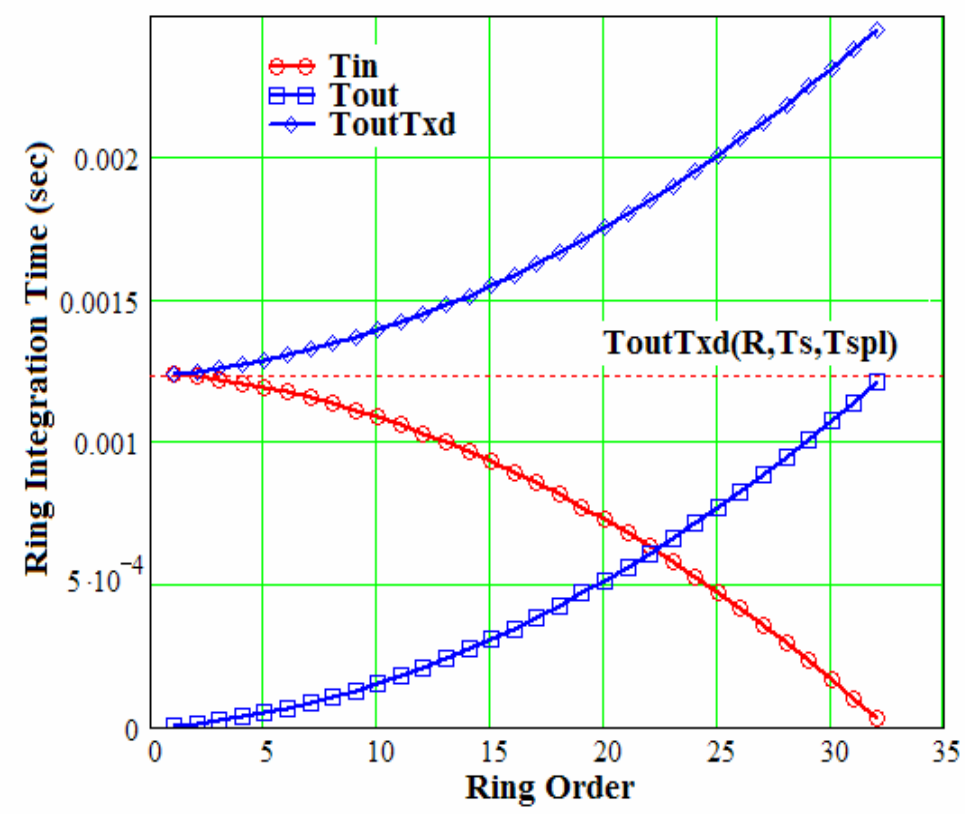


Fig 4.13 Inverse foveated dynamic range enhancement technique

The corresponding dynamic range enhancement we call *DRenhRevFov* that is due to the new outward scanning integration time is calculated based on equation (4.5) and is sketched in Fig 4.14.

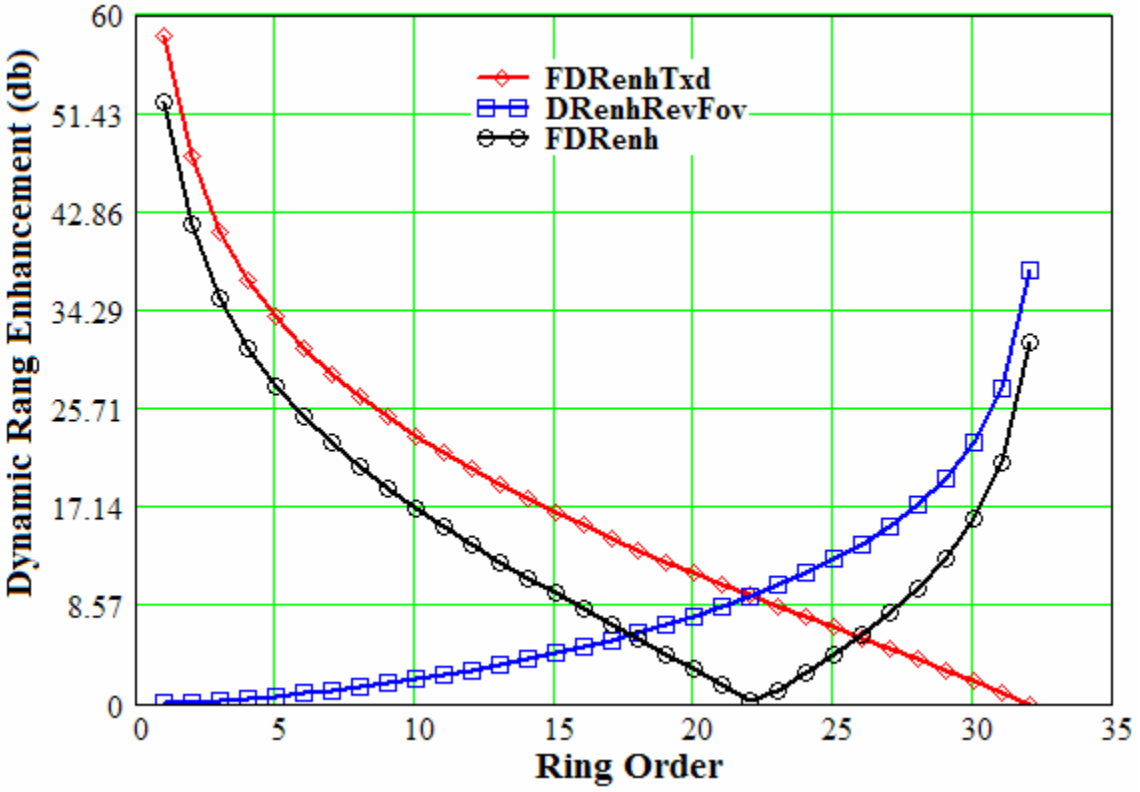


Fig 4.14 The different foveated dynamic range enhancement profiles

Although, the resulting dynamic range enhancement *DRenhRevFov* is increasing from the inner rings to outer rings, that we may call it an inverse-Fovea Dynamic Range Enhancement (*DRenhRevFov*), the new dynamic range profile has not gained similar enhancement like in the previous foveated dynamic range when pinned to the pyramid imager borders *FDRenhTxd*. The reason of this fact is quite apparent and is mainly due to the difference in pixels between the outer and the inner rings of the pyramidal imager which has led to inward scanning integration time being smaller at inner rings than outward at the outer rings

4.1.5 Pyramid-bordered FDR Profile Control

In this last mathematical manipulation of the FDR, we are concerned about the control of the FDR that has its border pinned at the outer ring of the pyramidal imager. This manipulation is based on a time shift of the ring integration time in both inward and outward scanning while preserving the condition of the outer ring integration time being the same for both inward and outward scanning schemes, as shown in the Fig 4.15 below.

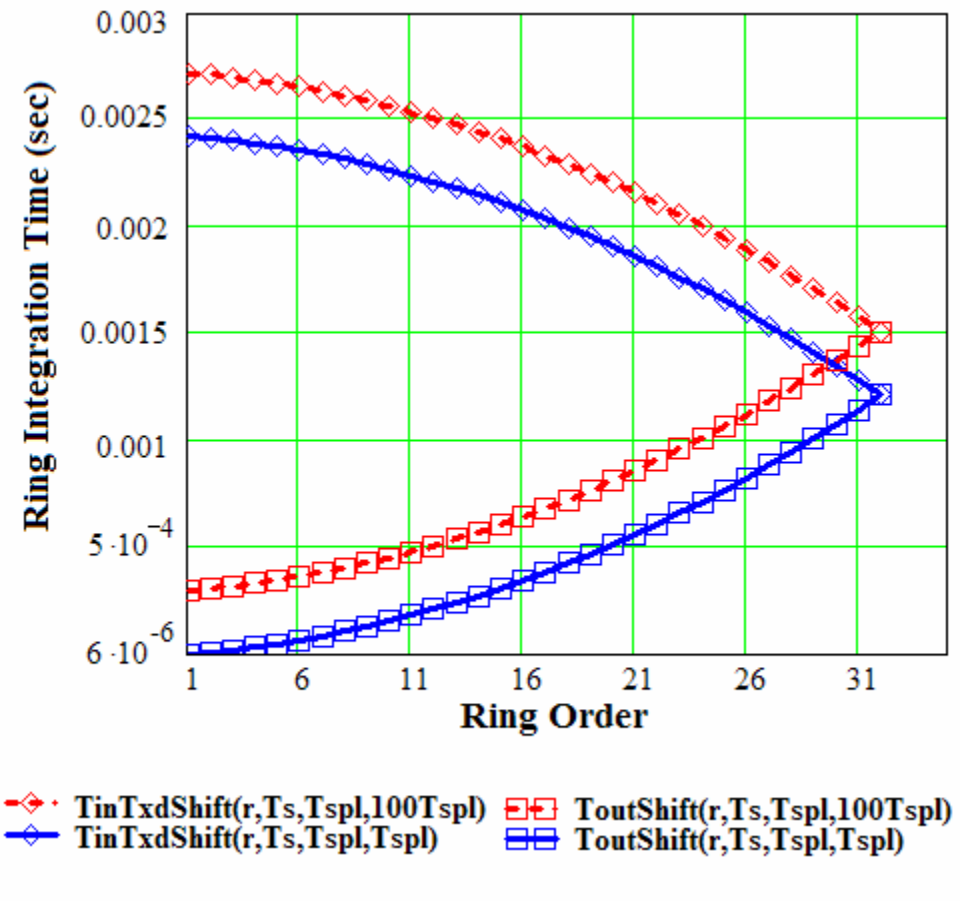


Fig 4.15 Control technique for pinned foveated dynamic range enhancement

Fig 4.15 shows the shifting of the inward and outward scanning in a pinned foveated dynamic range regime by a time period equal to *Tspl* (in solid lines) and 100*Tspl* (in dashed lines). In practice the above realization is made by delaying the start of the inward scanning, after each outward scanning, with the previously mentioned *Txd* period of time (see equation 4.15). The next step is to wait an extra equal delay to the start of the inward and the outward scanning and this is what meant by "shift" above (Fig 4.15). Thus for example, *TinTxdShift*(*r,R,Ts,Tspl,100Tspl*) is the inward scanning ring integration time shifted by shift=100*Tspl* and similarly is *ToutShift*(*r,Ts,Tspl,100Tspl).* These definitions are mathematically developed.

$$TinTxdShift\left(r,R,T_s,T_{spl},Txd\right)=TinTxd\left(r,R,T_s,T_{spl}\right)+Shift \ \ldots \ (4.17)$$

$$ToutShift\left(r,R,T_s,T_{spl},Txd\right)=Tout\left(r,T_s,T_{spl}\right)+Shift \ \ldots \ (4.18)$$

But what is the impact of this time shift? To answer this question let us plot the resulting dynamic range using equation 4.5 for shifts of 0s, 50*Tspl* and 100*Tspl*. The resulting intrascene dynamic range enhancement for the above three case is shown in Fig 4.16.

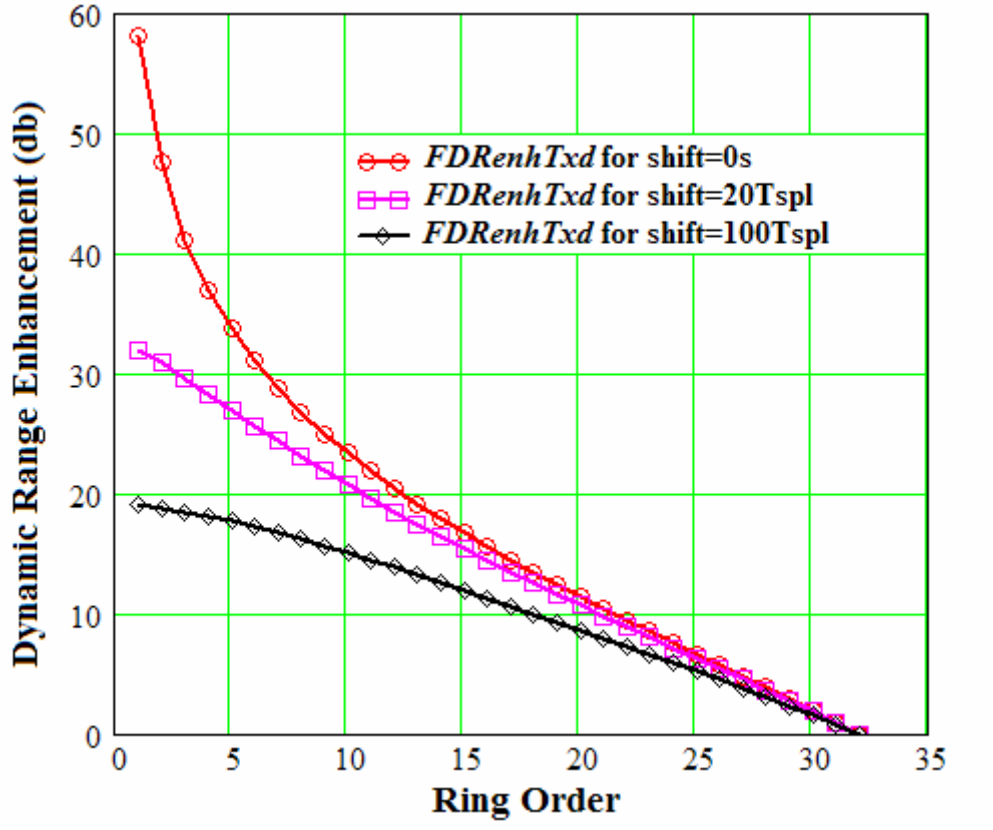


Fig 4.16 Pinned foveated dynamic range enhancement versus shifting time

It is apparent that increasing the shift of the pinned inward and outward scanning integration time results in a degradation of the FDR especially at the inner rings where its shows a clear change of the FDR shape. To have a clear view of the FDR degradation, the profile of the resulting FDR variation is plotted in 3D with respect to the values of the shift. This profile is plotted in Fig 4.17 below where shift periods are taken as multiple of *Tspl*.

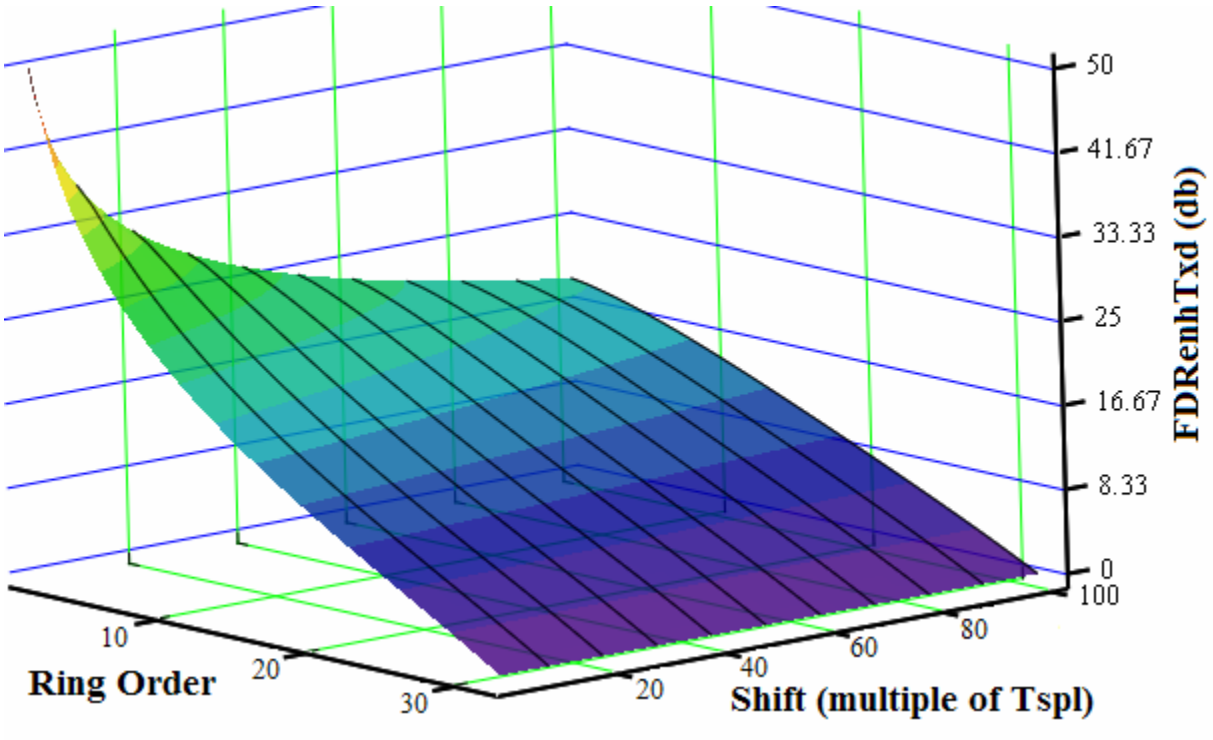


Fig 4.17 Pinned foveated dynamic range enhancement versus shifting time in 2D

With the change of shape of the FDR near the inner rings as mentioned earlier, the *foveation* property of the FDR fades away. However, an interesting finding can be drawn in this point when an intermediary shift of 18*Tspl* (18x3μs=54us) result an interesting shape of the FDR as shown in Fig 4.18.

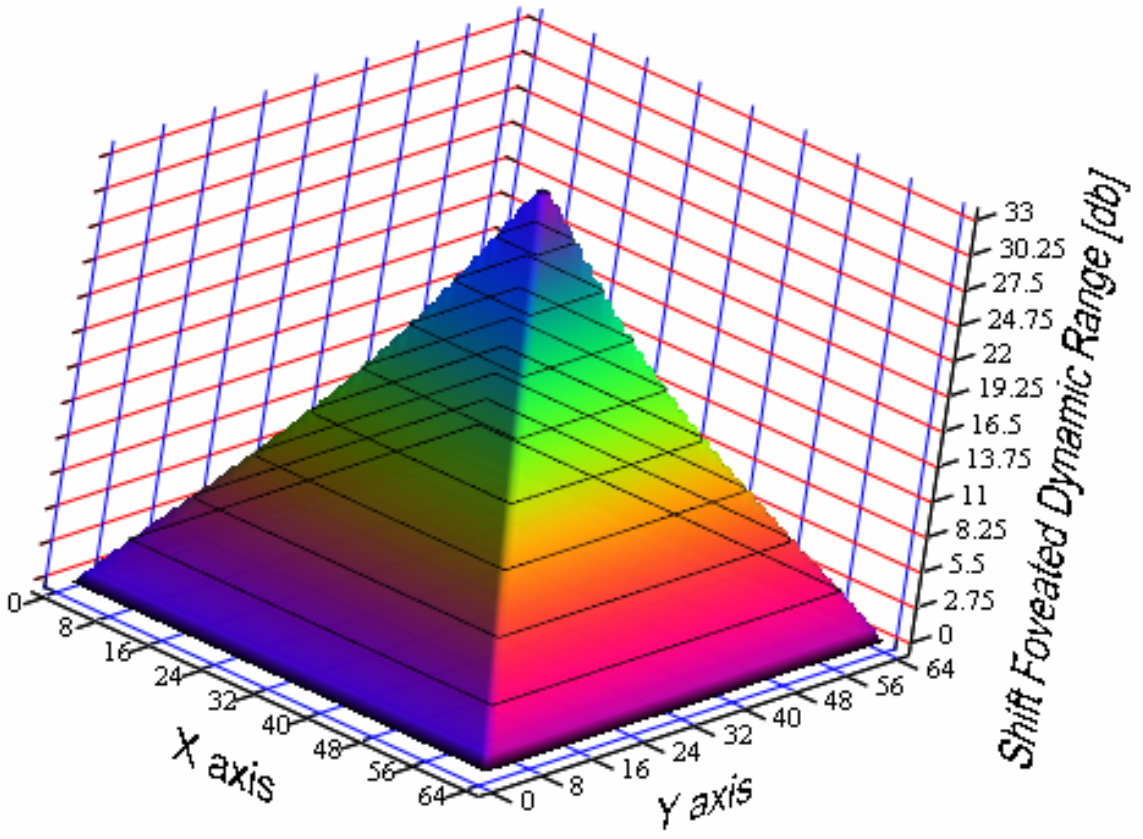


Fig 4.18 Pyramidal dynamic range enhancement profile

An interesting observation from Fig 4.18 is that the pyramidal imager hardware architecture supported with bouncing scanning scheme and with some scanning timing manipulation has generated a pyramidal dynamic range.

4.2 Restrictions of the FDR Mathematical foundations

The above mentioned methods do not take into account the constraints of the real imager, especially the saturation of the photodiode voltage and the impact of various noise sources. In the following chapters the physical applicability of foveated dynamic range is examined, including the consequence of adding the physical constraints to this model.

4.3 High Speed Imaging of Pyramidal Imager

In this section, the sampling speed of the pyramidal CMOS architecture is compared to the equivalent (in timing and array size) classical CMOS architecture. First, the frame integration time is calculated for both architectures. Then, the ratio of the two integration times is then deduced and plotted for different sizes of CMOS imager. Note that we are dealing so far with square CMOS imagers and that we are using the parameters defined in the previous chapter section 4.1.1.

We will examine the frame integration time of a rolling scanning of the pyramidal CMOS imager in which all rings will have the same integration time. In the rolling scanning, like in the raster scanning, the ring sampling will start from either the inner ring or the outer ring and will progress successively until the last ring after which it will start over from the first ring as shown in Fig 4.19, and hence the name of rolling scanning. This will lead to a uniform integration time for all imager pixels. Using the same parameters as in section 4.1.1 (namely R, T_s and T_{spl}) and using simple mathematical calculation steps the estimation of pyramidal imager frame integration time $FrameTint_{Pyr}$ for a rolling scanning will be developed.

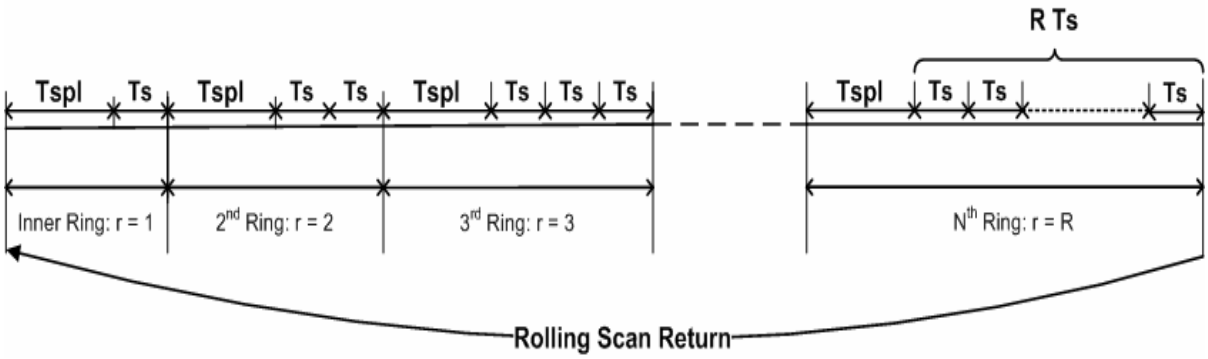


Fig 4.19 Rolling scanning timing diagram

From the timing diagram shown in Fig 4.19 frame integration time $FrameTint_{Pyr}$ is estimated through the following equation.

$$FrameTint_{Pyr} = R\,T_{spl} + T_s \sum_{s=1}^{R} s \text{, that is } FrameTint_{Pyr} = R\,T_{spl} + T_s \frac{R}{2}(R+1) \quad \ldots (4.19)$$

After simple mathematical manipulations $FrameTint_{Pyr}$ formulation is simplified to:

$$FrameTint_{Pyr} = \frac{T_s}{2} R^2 + \left(T_{spl} + \frac{T_s}{2} \right) R \ldots (4.20)$$

It is worth mentioning that $FrameTint_{Pyr}$ was calculated for a single segment out of the 8 segments of the pyramidal imager because all of the segments are sampling the integrated image simultaneously. Therefore, frame integration time will the same as the single segment integration time.

Using the same array of pixels with the same scanning timing parameters but with classical architecture of CMOS imager, the frame integration time $FrameTint_{Class}$ is extracted. First, recall the number of rings in the pyramidal imager is equal to half of the number of pixels in

one single row (or column) namely $R=N/2$ or $N=2R$. Besides, every row in classical CMOS imager will take the same amount of time to sample its data which is equal to $T_{spl} + NT_s$. Multiplying the row sampling time by the number of rows $2R$, the total time needed to integrate the whole image ($FrameTint_{Class}$) will be equal to:

$$FrameTint_{Class} = 4T_s R^2 + 2T_{spl} R \ldots (4.21)$$

In order to simplify the comparative frame acquisition speed analysis we replace in the above two equations the value of T_{spl} with $3T_s$. This replacement if due to the fact that *Tspl* includes 3 switching cycles, while T_s include just one switching cycle, all mentioned in previous section. After replacement one gets the following two equations.

$$FrameTint_{Pyr} = \frac{1}{2}T_s R^2 + \frac{7}{2}T_s R \ldots (4.22)$$

$$FrameTint_{Class} = 4T_s R^2 + 6T_s R \ldots (4.23)$$

The frame rate is the inverse of the frame integration time, and from this definition we can evaluate the ratio of the pyramidal imager frame rate FR_{Pyr} over the classical imager frame rate FR_{Class} by the following equation:

$$Ratio(R) = \frac{FR_{Pyr}}{FR_{Class}} = 2\frac{4R+6}{R+7} \ldots (4.24)$$

It is clear now from equation (4.24) that that the pyramidal imager is faster than the classical CMOS imager in acquiring images using similar timing parameters T_s and T_{spl} and for a similar square size of pixel array imagers. In fact, as the size of the imager increases the ratio between the pyramidal and classical imagers frame rates increase to a limit of 8 as shown by equation 4.24 and Fig 4.20. It is interesting to note that the ratio between the two frame rates

is also independent of the *Ts*, if we keep the ordinary approach of *Tspl*=3*Ts*, as shown in equation (4.25).

$$\lim_{R\to\infty}\frac{FR_{Pyr}}{FR_{Class}}=8 \quad \ldots (4.25)$$

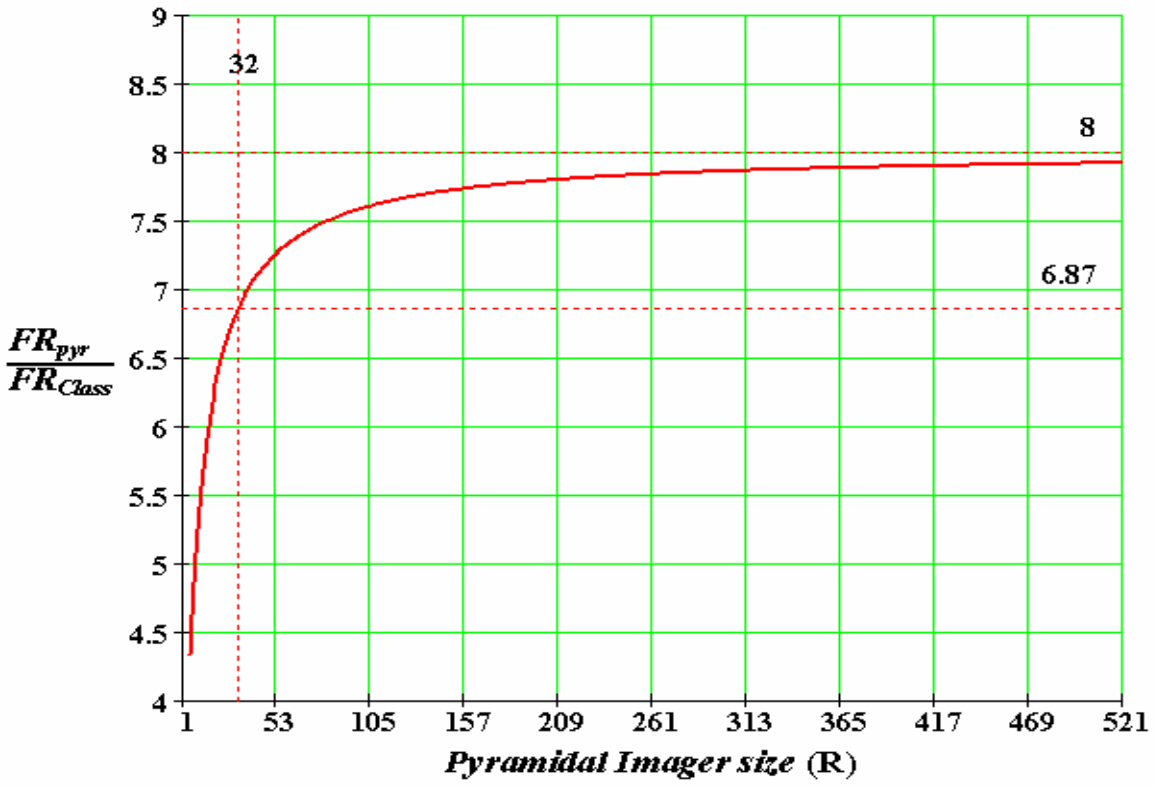


Fig 4.20 Fame rate ratio between pyramidal and classical imager of size 2*R* by 2*R*

From previous analysis, it can be deduced that the high speed acquisition of the pyramidal CMOS imager compared to the classical CMOS imager is mainly due to the segmentation of the pyramidal imager into eight parallel sampling segments. This is why the ratio of frame rate between the pyramidal CMOS imager and the equivalent (in timing and size) CMOS

imager approaches the limit of eight. To generalize this fact, we conclude that any imager having N parallel sampling segments will outperform the sampling speed of the classical imager N times. This technique of dividing the imager into parallel sampling segments has been used previously [91] [92] to achieve higher frame rates, however, for the case of pyramidal imager the technique was rather a direct result of the 2D sampling (ring sampling) than an arrangement to segment the imager acquisition system. In fact, the only way to read out the sampled 2D ring is through the diagonal output busses that ultimately will need 8 sample and hold banks aground the pyramidal imager.

4.4 Pyramidal Acquisition Speed with Serial readout.

In this section, the pyramidal frame rate speed is analysed and compared with that of the classical CMOS imager, assuming serial readout instead of the 8 segment parallel readout studied in the previous section. The serial readout of the pyramidal CMOS imager is made by sampling whole frame, ring by ring, into the sample and hold capacitors banks, and then reading the sampled values serially pixel by pixel through a single output buffer.

First, the frame integration time (then frame rate) $FrameTint_{PyrSerial}$ ($FR_{PyrSerial}$) formula of pyramidal imager using the serial readout as described above and the timing parameters described in section 4.1.1, is developed. Then, a frame rate ratio between $FRate_{PyrSerial}$ and FR_{Class} is calculated and analysed.

Frame integration is calculated by calculating the time needed to read a given ring (after sampling it to sample and hold capacitor bank) serially in circular shape. Summing these times will lead us to the whole frame integration time.

In order to sample and readout a given ring *r,* one needs to spend T_{spl} to sample the ring to the sample and hold capacitors to get both *Vs* and *Vr*, then to shift serially the (8*r* - 4) pixels of the ring for a period of time of (8*r* - 4) T_s. Therefore, $FrameTint_{PyrSerial}$ is calculated from equation 4.26 below.

$$FrameTint_{PyrSerial} = \sum_{r=1}^{R} \left(T_{spl} + (8r - 4) T_s \right) \text{ or } FrameTint_{PyrSerial} = 4 T_s R^2 + T_{spl} R \ldots (4.26)$$

Knowing that the frame rate is the inverse of frame integration time, and using equations 4.26 and 4.23, we deduce the ratio $SerialRatio_{PyrSer\_Class}$ of $FRate_{PyrSerial}$ (inverse of $FrameTint_{PyrSerial}$) over FR_{Class} as shown in equation 4.27 below.

$$SerialRatio_{PyrSer\_Class}(T_s, T_{spl}, R) = \frac{4 T_s R^2 + 2 T_{spl} R}{4 T_s R^2 + T_{spl} R} \ldots (4.27)$$

It is clear from equation 4.27 above that for high values of *R* (imager size) the ratio of $FRate_{PyrSerial}$ over FR_{Class} approaches 1. However, for smaller value of *R* the ratio of $FRate_{PyrSerial}$ over FR_{Class} is higher than 1 and hence the pyramidal imager has faster frame rate than its equivalent classical imager.

Recall that T_{spl} is the time spent for sampling a ring (or a row) into sample and hold capacitor banks for both *Vs* and *Vr* (after resetting the ring or the row), while T_s is the time needed to buffer out the sampled data from the sampling capacitor to off chip. Therefore, it is clear that the minimal value of T_{spl} is *2* T_s and the maximal value is relatively unbounded. Thus let us define a variable $\beta = (T_{spl} / 2 T_s)$ in order to determine the impact of the relative sampling speed, between the ring (or row) sampling speed and pixel readout speed, on the frame ratio.

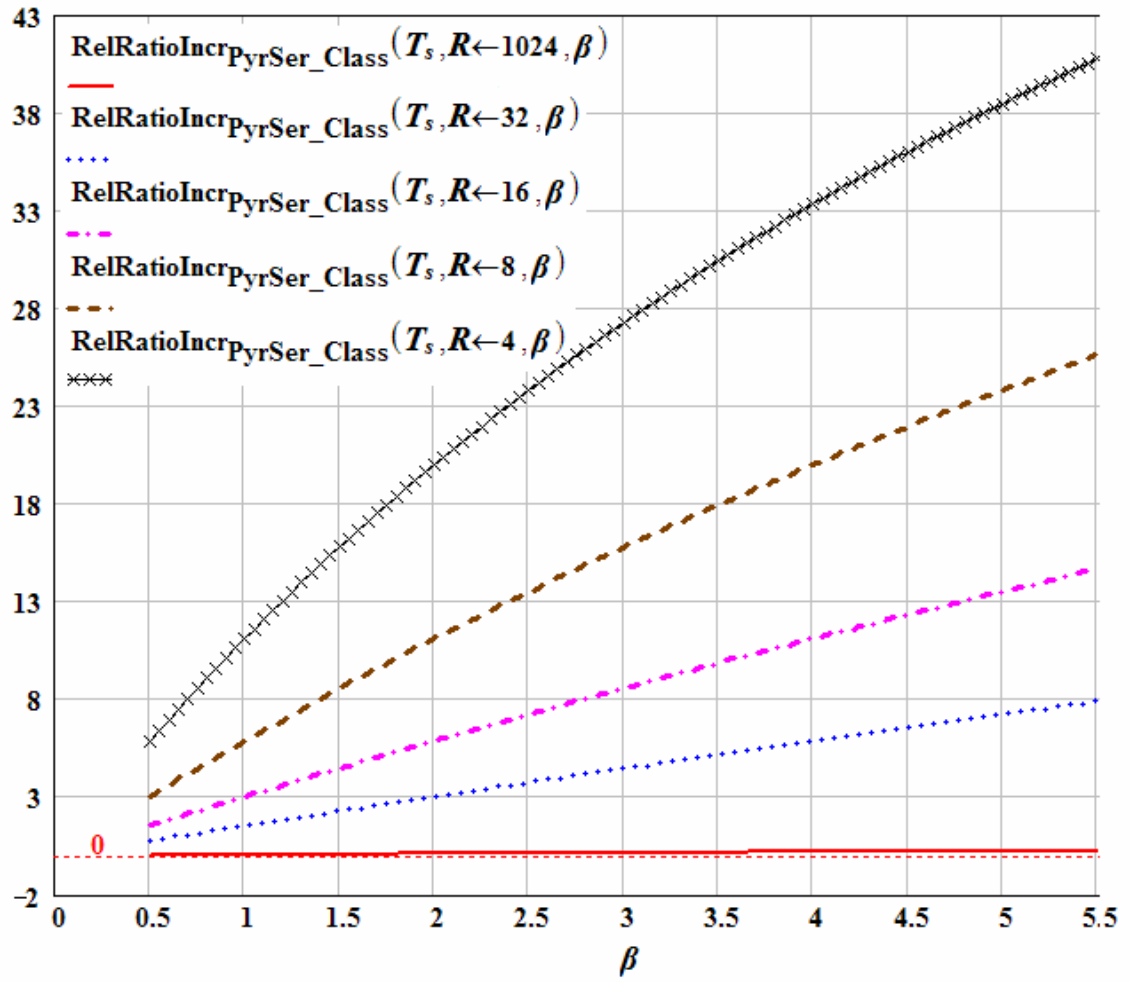


Fig 4.21 Relative ratio of pyramidal imager rate (serial readout) over classical imager's

The graph in Fig 4.21 shows plots of relative increase $RelRatioIncr_{PyrSer\_Class}$ (in percent) of frame rate of pyramidal CMOS imager serially readout over the classical CMOS imager for different size of imager resolutions R, versus β. This increase is calculated based on the following formula.

$$\mathrm{RelRatioIncr}_{\mathrm{PyrSer\_Class}}(T_s, R, \alpha) = \left(SerialRatio_{PyrSer\_Class}(T_s, R, \alpha) - 1\right)100\% \ldots (4.28)$$

Developing equation 4.28 leads to the following simple equation 4.29 showing the independence of $RelRatioIncr_{PyrSer\_Class}$ from T_s.

$$RelRatioIncr_{PyrSer\_Class}\left(T_s, R, \beta\right) = \frac{\beta}{\beta + 2\,R} \ldots (4.29)$$

It is apparent that for any values of β and R, the frame rate of the pyramidal CMOS imager is higher than that of the classical CMOS imager because of the positive value of $RelRatioIncr_{PyrSer\_Class}$. This relative increase of frame rates gets lower and lower for high imager resolutions as R increases until the pyramidal frame rate using serial readout will be exactly similar to the classical imager frame rate. Finally, it can be seen that the influence of the parameter β becomes apparent only at low resolutions, where it increases the ratio of pyramidal imager frame rate over that of the classical imager. This is obvious because as β increases, the scanning time will be more dependent on ring (row) sampling than on pixel out-buffering. Adding to this the fact the pyramidal imager has a number of rings that is half the number of rows in its classical equivalent imager, the influence of β becomes clearer. In fact as β increases, the ratio $RelRatioIncr_{PyrSer\_Class}$ approaches 100% meaning that the pyramidal imager frame rate is two times higher than that of the classical imager.

Practically, the values of β is in the range of 1 to 2 leading to a maximum frame rate out performance of the pyramidal CMOS imager (using serial readout) that is 20% better than the classical CMOS imager for a resolution of 8x8 pixels. While this is clearly too small compared to common imager sizes, this result shows that inner rings (up to the 3rd ring for an 8x8 array for example) are sampled faster using the serial scanning than their equivalent classical arrays (by about 20% for up to the 3rd ring for example). Due to the fact that inner

rings could be selected (and readout) independently of the remaining rings (for a specific imaging needs or applications) shows that even with the serial readout the pyramidal imager readout its fovea faster than an equivalent classical imager. This is an important aspect of the fovea region (which is the region of interest) of the pyramidal imager as it will possible to track (in time) fast moving objects compare to the classical imager. Therefore, this fact is another foveated attribute of the pyramidal imager beside the FDR enhancement.

The present analysis shows that pyramidal CMOS image sensor is always faster than classical CMOS imager of equivalent size and scanning timing parameters. However, this high speed feature is more prominent in the case of acquisition segmentation into parallel readout channels, which is appropriate and original to pyramidal architecture, than in the case of serial readout. To conclude this section, its is worth mentioning that parallel segment readout scheme of the pyramidal imager is the readout of choice not only because of the fast frame rates its achieved, furthermore, because it is the most natural scheme to this architecture due to the fact that every segment has its own independent sample and hold banks of capacitors

4.5 FDR Pinning Cost on Frame Rate

In order to study the impact of the extra time *Txd_exact* paid for the pinning of foveated dynamic range, let us recall its value from equation 4.15 and divide it by the frame integration time of the pyramidal image sensor when sampling using bounced scanning here called $FrameTint_{Pyr\_Bsc}$. In order to estimate the value of $FrameTint_{Pyr\_Bsc}$ one can use the diagram shown in Fig 4.19. Instead of the step back (rolling scan return) from the last (outer)

ring to first (inner) ring, the scan has to continue sampling from last ring to the first ring. Thus, it is straightforward to estimate $FrameTint_{Pyr\_Bsc}$ to be;

$$\text{FrameTint}_{\text{Pyr\_Bsc}} = 2\ \text{FrameTint}_{\text{Pyr}} \ldots (4.30)$$

Therefore, using equations 4.15, 4.20 and 4.30 the ratio of *Txd_exact* over $FrameTint_{Pyr\_Bsc}$ will be;

$$\frac{Txd\_exact}{FrameT\,\text{int}_{Pyr\_Bsc}} = \frac{R^2 T_s + (2T_{spl} - T_s)R - 2T_{spl}}{T_s R^2 + (2T_{spl} + T_s)R} \ldots (4.31)$$

From equation 4.31, the limit of the ratio of *Txd_exact* over $FrameTint_{Pyr\_Bsc}$ will be approaching 1 independently from the scanning parameters T_s and T_{spl} as shown in equation 4.32 below.

$$\lim_{R\to\infty} \frac{Txd\_exact}{FrameT\,\text{int}_{Pyr\_Bsc}} = 1 \ldots (4.32)$$

Equation 4.32 implies that in order to pin the foveated dynamic range the imager has to delay the bouncing scan frame integration time by 100% for relatively large pyramidal imager array. Causing the bouncing scan frame integration time to double to pin the foveated dynamic range, the frame rate of the pinned FDR bouncing scan will be half of the regular FDR frame rate.

Recalling that that bouncing scan frame rate is half of that of the non-bouncing scan (rolling scan) as deduced from equation 4.30, and recalling the ratio of the rolling scan of pyramidal imager of that of the classical imager frame rate from equation 4.25, one can reach the following equations;

$$\lim_{R\to\infty}\frac{FR_{Pyr\_Bsc}}{FR_{Class}} = 4 \ldots (4.33.a)$$

$$\lim_{R\to\infty}\frac{FR_{Pyr\_pFDR}}{FR_{Class}} = 2 \ldots (4.33.b)$$

FR_{Pyr_Bsc} and FR_{Pyr_pFDR} are frame rate of regular bouncing scan FDR and pinned FDR respectively. In conclusion, the extra delay spent in bouncing scan in order to realize the pinned FDR has caused the bouncing scan frame rate to be reduced to half of its regular value. Nevertheless, the pinned FDR frame rate is two times the frame rate of the classical image sensor frame rate using similar imager size and scanning time constants T_s and T_{spl}.

4.6 Benefit of Foveated Dynamic range

In this section, we will show one particular benefit of FDR in image acquisition if the pyramid sensor were to be constructed with analog-to-digital conversion (ADC).

First, let us introduce the formula that converts dynamic range from decibels into significant bits needed in ADC. The formula is shown in equation (4.34).

$$DR_{bits} = \log_2\left(10^{\frac{DR_{dB}}{20}}\right) \ldots (4.34)$$

Now, all formulae regarding DR in decibels can be converted into binary bits. In particular, let us convert the pinned FDR shown in Fig 4.12 and plot the result in the following Fig 4.22. Fig 4.22 shows the extra bits achieved due to intrascene dynamic range enhancement resulting from image fusion, to be added to the original bits generated from electro-optical ADC conversion bits. In uniform DR enhancement schemes such that mentioned in [68] the

extra binary bits generated by DR are identical all over the enhanced image. However, in the FDR developed in our work, extra bit generated by DR enhancement is higher (in number) in the central region (fovea) and decreases towards the perimeter of the imager. This means, FDR actually minimizes the size of the generated images by allocating more resources (data) to the fovea region (which is assumed to be centred on the region of interest similar to the case of human fovea). That is exactly, the case for the human fovea, which has more interconnectivity between its photocells (cones) and the processing neurons (ganglions) than that of the peripheral retinal photocells (rods) [4]. In fact, the former have a many-to-one interconnectivity (between photocells and processing neurons) configuration in addition to their relatively higher density, which explains their higher light sensitivity dedicating them for initiating vision at low light intensities. On the other hand, the cones have a one-to-many interconnection configuration with ganglion cells inferring dense representation of cones which explains the high contrast of human vision (and many mammalians' vision) in the fovea (central vision) compared with the peripheral vision. In conclusion, as the low light vision (*scotopic*) is initiated at the retina periphery (despite the low contrast) and the high light vision (*photopic*) is well sampled in the fovea, this may implicitly mean the human higher DR is made by fusion between the two photocells making the dynamic range appear higher at the central region of retina and decreasing at the periphery, analogous to the pyramid sensor.

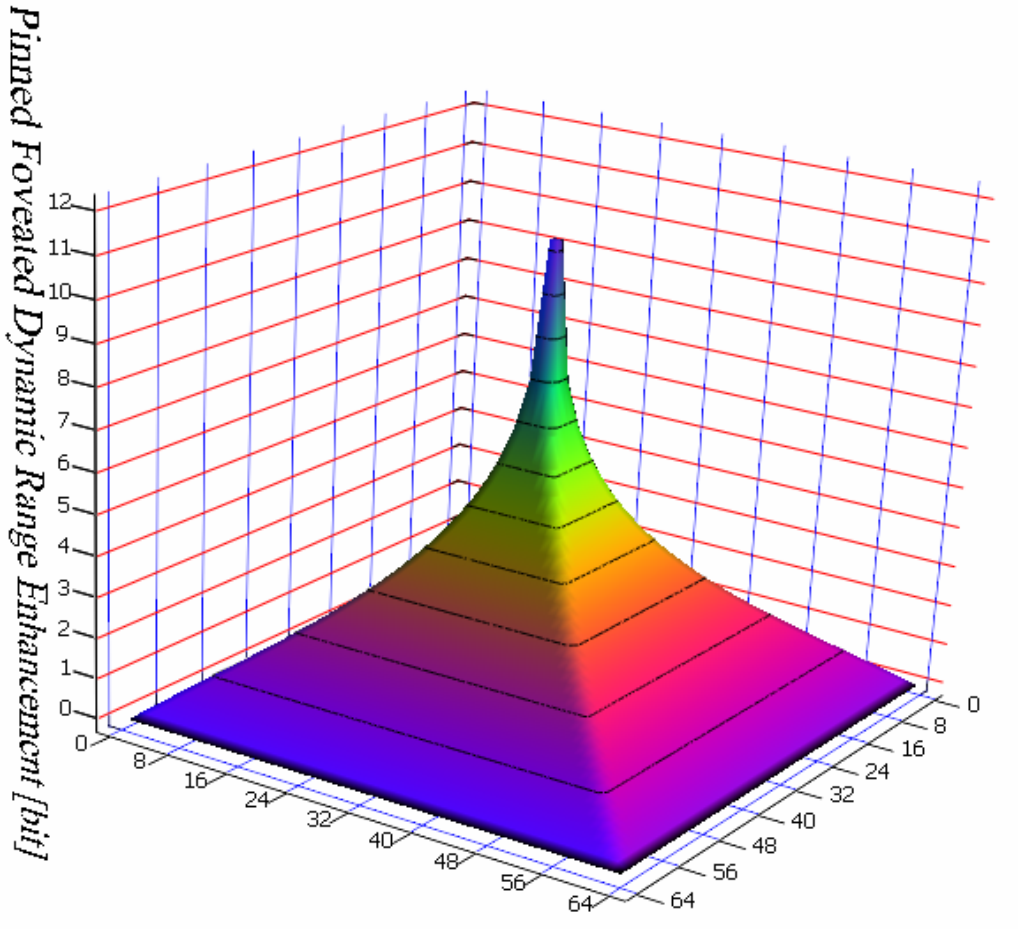


Fig 4.22 3D view of the pinned FDR expressed in binary bits

4.7 Summary

In conclusion, the pyramidal architecture, through its 8 parallel output channels and sample and hold banks and thanks to its 2D ring sampling, exhibits high speed acquisition capability compared to similar classical image sensors by a factor of up to eight. In fact, the minimum value of this ratio of frame rates of pyramidal architecture sampling over the classical imager

frame rate is equal to 2.5 in the case of 4 pixels (2 rows and 2 columns in classical imager or one ring in pyramidal imager). The value of the previous ratio increases very rapidly as shown in Fig 4.20 to reach the limit of 8. Our designed imager of 64x64 pixels is in fact about 7 times faster than its equivalent classical version of CMOS image sensor when using the same scanning parameters namely *Ts* and *Tspl*. Even using serial readout of the pyramidal imager, its frame rate performance outperform that of the classical CMOS imager again due to the 2D nature of rings sampling compared to the 1D sampling of the classical CMOS imager.

We have also shown that using bouncing scanning in acquiring images by the pyramidal image sensor leads to an interesting foveated shape of intrascene dynamic range. Initially the border of this fovea was independent of the scanning parameters at about 71% of the ring size from the origin inner ring as demonstrated in section 4.1.2. We have suggested a practical method of how to pin the foveated dynamic range enhancement outer ring using some timing manipulation, as discussed in the section 4.1.3. A similar technique was tried out to get an inverse foveated dynamic range in section 4.1.4.

The cost paid to get pinned FDR is a reduction of the frame rate of the pyramidal architecture. Finally, FDR enhancement cost was estimated in terms of memory consumption through needed binary bits for DR enhancement information storage. This has unveiled some similarity between the designed pyramidal architecture with FDR enhancement (memory) cost and the neural interconnectivity configurations in human fovea with its outstanding image sampling quality.

Chapter 5

Testing and Characterization

In the present Chapter we discuss the characterization process of the pyramidal imager starting from the experimental setup and continuing to the data analysis and the extraction of imager characteristics. The main objective is also to verify the mathematical analysis of the foveated dynamic range (FDR) developed in the previous Chapter. To do this, we develop a simple empirical model of the sensor photo-response which enables calculation of the output of any pixel for arbitrary illumination and integration time. From the model, the dynamic range enhancement of the pyramidal image sensor under bouncing scanning can be calculated and compared favourably with the ideal value found in the previous chapter.

5.1 Testing Setup

The setup of the testing apparatus is shown in Fig 5.1.

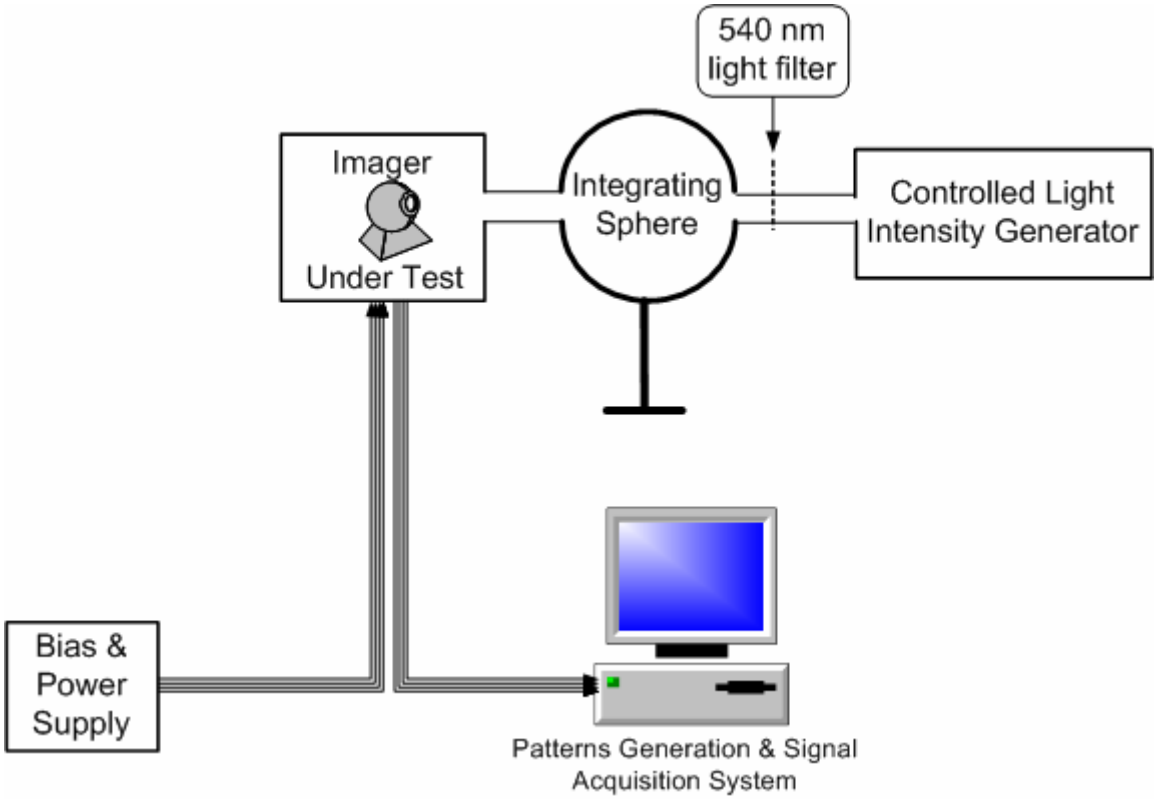


Fig 5.1 Experimental setup for the characterization of Pyramidal CMOS imager

The testing equipment is built around a 500MHz INTEL™ Pentium III computer in a Windows2000 operating system environment. The computer system is used to synthesize and control the testing pattern of the different controlling signals to be sent to the imager under test. The imager will respond according to the electrical signals sent by the PC and the light intensity generated by a light intensity generator through a 540 nm (green) optical filter. To achieve an approximately uniform spatial intensity, the illumination is passed through an integrating sphere. Over the entire output aperture of the integrating sphere (approx. 20 mm diameter), the uniformity is expected from manufacturer's specifications to be 98%. Hence,

the illumination is assumed to be perfectly uniform for the relatively 4mm x 4mm sensor die. We used two Industry Standard Architecture (ISA) Master-Slave 32-bit output cards of the CompuGen3250 family of cards to generate electrical signals to control functionality of the imager. Analog signals were acquired from the imager, by a 2 channel, 16 bit resolution, 2.5 MS/s acquisition Peripheral Component Interconnect (PCI) CompuScope1602 card. Both cards were made by Gage-Applied technologies corp.

The signals patterns were generated using Matlab™ and Gagebit software that comes with the signal generation card. The latter generates the patterns in ASCII format and the former converts them into a binary format before loading them into the generation card. This method was later modified to automate the testing by using the Application Program Interface (API) that Gage Corp supplies with its card in order to interface them with other programming tools. To this end and to automate our acquisition we have chosen the graphical programming tool known as Labview™ along with the API's of CompuGen and CompuScope to automate the generation and acquisition of images sampled at the image sensor under test, as Fig 5.2 depicts.

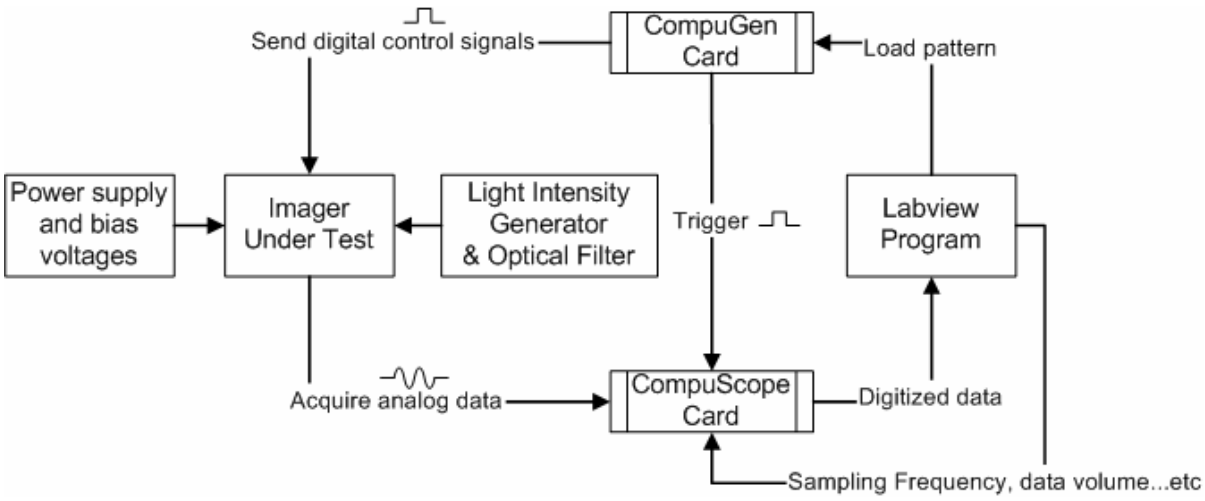


Fig 5.2 Software and hardware acquisition system

The Labview program loads the digital pattern of the controlling signals into the CompuGen card. Among those controlling signals is a pulse (trigger) signal that will be used to externally trigger the CompuScope card to start sampling the data coming from the imager. Labview also controls the CompuScope card by specifying its sampling rate and number of samples to be acquired from the imager. Finally, the Labview program is used to automate the sweeping of the sampling frequency and save the acquired data for each sampling frequency.

The pyramidal imager that is under test is a 64x64 pixel imager laid out on a 120 pin chip soldered on board and interfaced by 21 digital input signals and 16 analog output channels. These 16 analog output channels correspond to the VS and VR signals of the 8 pyramidal imager segments (as discussed in chapter 3). The digital signals are used for ring reset and select decoders as well as for the diagonal select decoders which output the sampled ring

down to the sample and hold capacitor banks around the imager (see chapter 3). Fig 5.3 shows the different segments of the pyramidal imager and their corresponding channels beside the reset and select busses. Because of the limited number of acquisition channels (2 channels of CompuScope) we multiplex the imager 16 channels to the 2 channels of the acquisition card by using three 16-channel/dual 8-channel (16-to-2) differential high performance CMOS analog multiplexers MAX307CPI. This multiplexing scheme is shown in Fig 5.4.

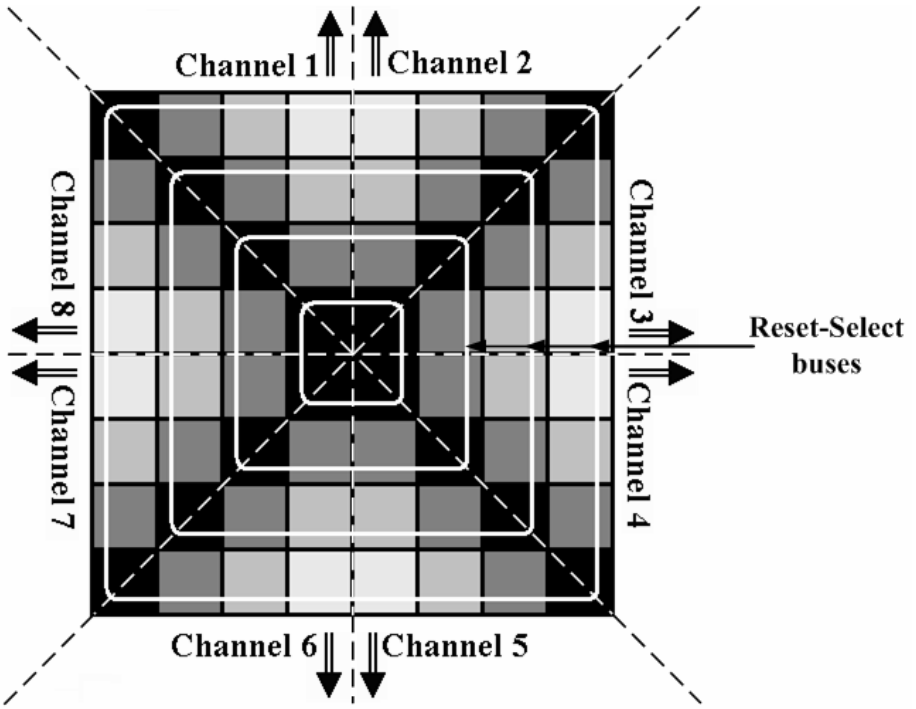


Fig 5.3 Pyramidal imager segments and channels.

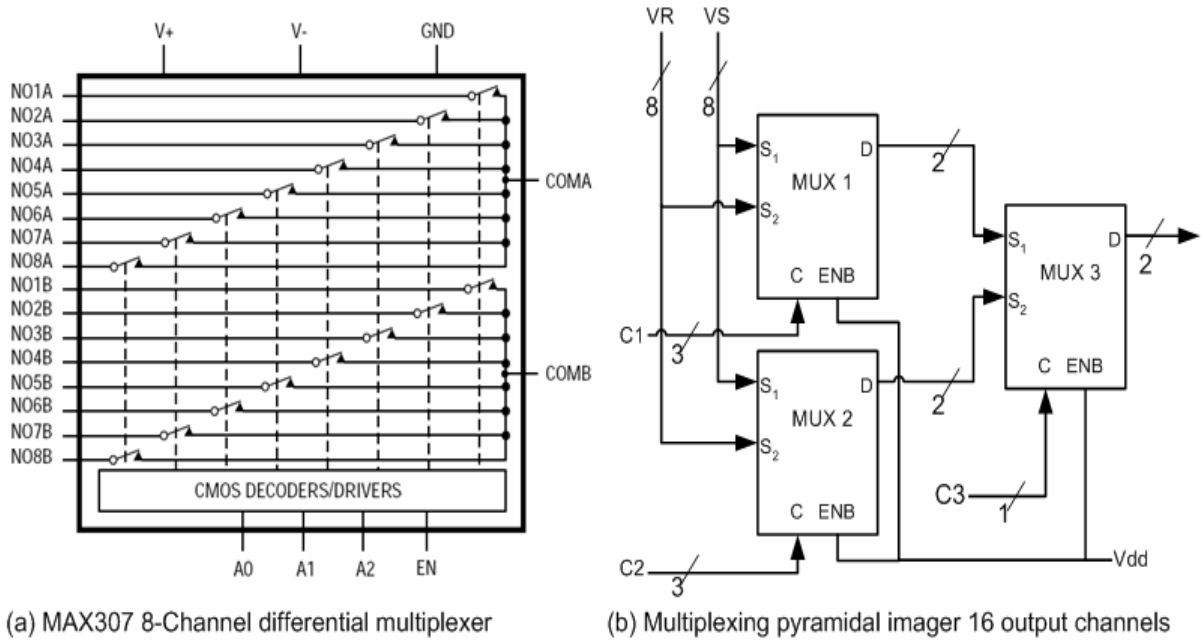


Fig 5.4 Multiplexing pyramidal imager 16 output channels to 2 acquisition channels

Each multiplexer, Mux1 and Mux2, will select a pair of (VS, VR) corresponding to one of the 8 pyramidal imager segments according to the 3-bit binary codes C1 and C2 respectively. The output of Mux1 and Mux2 is fed to the input of Mux3 that will finally select from the two pairs (VS, VR) provided by Mux1 and Mux2. Based on the 1-bit address code, C3, Mux3 multiplexes two possible output pairs (VS1, VS2) or (VR1, VR2). This mechanism of multiplexing between the pyramidal imager clusters allows us to scan the imager segment by segment and hence we called this type of scanning, *segment scanning*. As can be seen, segment scanning is very flexible and can be circular or random, based on the image acquisition need. Furthermore, segment scanning is a kind of imager segmentation (different from image segmentation) that will allow the imager user to localize the acquisition to just a

part of the imager focal plane. For the pyramidal imager this area is of a triangular shape with its tip at the center of the imager. This technique is similar to multiple output tap techniques used in standard high-speed systems.

In the multiplexing scheme described above, Mux3 outputs once (VS1, VS2) or (VR1, VR2) and hence two pulses of C3 are needed in order to perform CDS. Another method of multiplexing can be made by grouping (VS1, VR1) signals from the output of Mux1 and Mux2 at the S1 input of Mux3, respectively, and (VS2, VR2) at the other input port of Mux3. In this case we will have (VS1, VR1) or (VS2, VR2) multiplexed at the output of Mux3 depending on the address code C3. Therefore, we can perform CDS at any given value of C3. Thus, we have developed two multiplexing schemes:

- For one value of C3 VS signals (or VR signals) are available from two segments; to make the CDS we need to apply the other value of C3.
- VS and VR are available simultaneously for one segment at a given value for C3.

We called the first multiplexing scheme MaxINFO multiplexing and the second we called MaxFPN as the first presents two segment values (maximum information from imager) at a given time (or a value of C3), whereas the second present the maximum denoising possibility of VS from FPN noise by providing VS and VR at the same time.

In all the subsequent data acquisition we will use MaxFPN multiplexing scheme and scan the imager's 8 clusters one by one. One note worth mentioning in this step is that the frame rate is exactly equal to segment rate because all segments can be read out independently and in a parallel fashion. The multiplexing is only made for the sake of image reconstruction and

due to the limited number of acquisition channels. Fig 5.5 shows the board on which the imager was fixed with the multiplexers for the two multiplexing scheme.

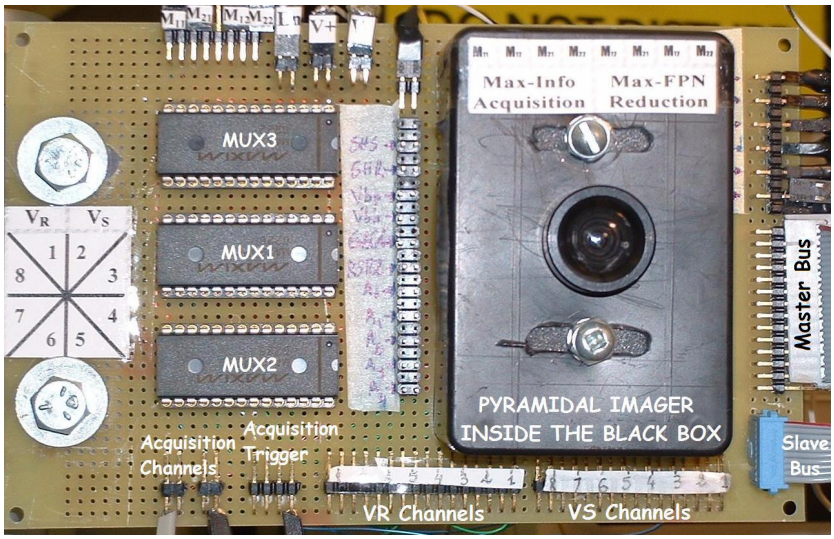


Fig 5.5 Board of the imager under test with the multiplexing implementation.

5.2 Dynamic Range Calculation

Dynamic range (DR) is a characteristic evaluation of any acquisition sensor determining the ability of the imager to sense low light and high light intensities. Usually DR is the ratio between the maximum detectable light intensity and lowest detectable light intensity in decibels. The maximum detectable light intensity is the light intensity that will start saturating the output voltage at a given integration time. The minimum detectable light intensity corresponds to the noise floor inferred light intensity. Corresponding interpretation

of these two limits is the maximum and minimum values of light intensity between which the imager is linear as shown in Fig 5.6.

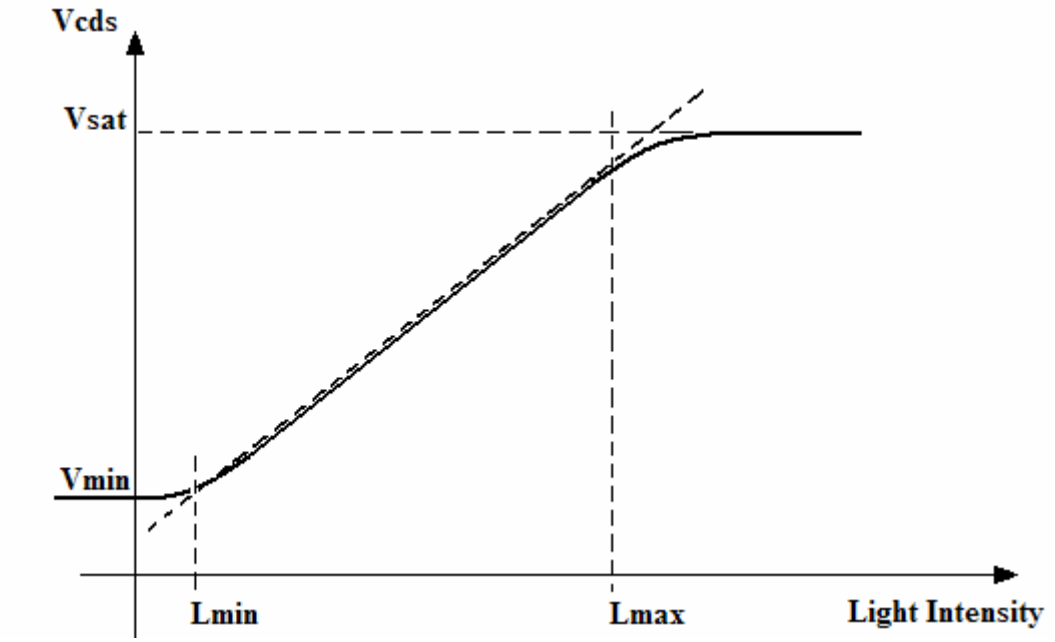


Fig 5.6 Dynamic range extraction from light intensity transfer function

Here, Vcds refers to the output signal after CDS denoising that is (VR-VS). This parameter is used because of the low spatial noise of Vcds (as determined experimentally) and because VS represents the discharged voltage of the pixel photodiode, which is decreasing as light intensity increases.

5.3 Pyramidal Data Structure

Image sampling using the pyramidal image sensor is different from that of the classical image sensor architecture as discussed in chapter 3. Consequently, its image readout is also different. Image reconstruction (or interpolation) is closely related to the image sampling technique. In the classical imager, the sampled rows are fed continuously to the image reconstruction memory locations in a serial fashion. In other words, the data structure indexing the classical imager pixels is based on two loops, a column loop inside the row loop. So, as the data acquired from the imager comes from raster scan, image reconstruction runs the two loops. As soon as column loop finished locating incoming data in the memory the column loop increments, and so on until the imager finishes the whole frame acquisition. For the pyramidal imager, we are dealing with rings and ring sequence, instead of rows and columns, which add more complexity to the pyramidal imager compared with the classical imager. In particular, each ring contains a different number of pixels, increasing from the inner ring towards the outer ring. This is not the case for the classical imager, in which the number of pixels in each row is equal and independent of the order of the sampled row. The data structure indexing the two imagers is shown in Fig 5.7.

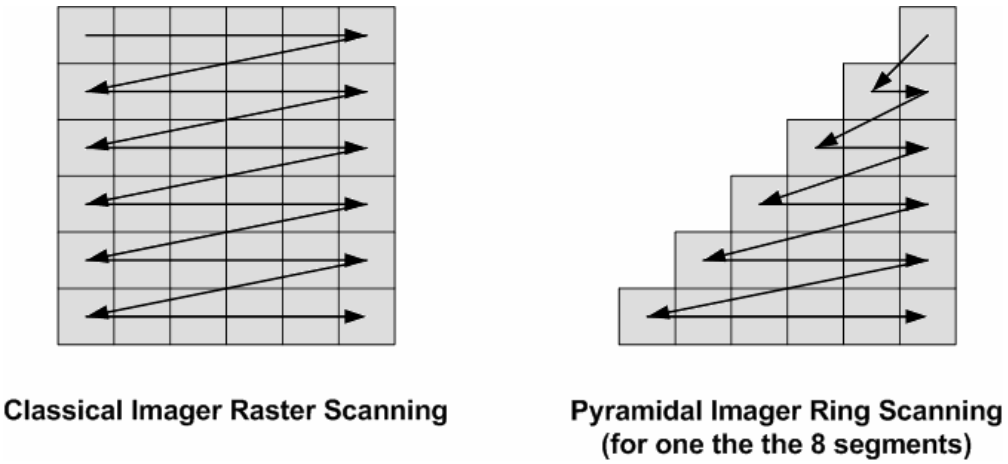


Fig 5.7 Sampled data structure in pyramidal and classical imagers

5.4 Data Analysis

In this section we will first analyze the photo-response of the non-bouncing scanning before repeating the process for the bouncing scanning. We will extract the sensitivity of the imager, which is an intrinsic characteristic of the imager. Subsequently, this parameter will be used to build a model from which we can extract dynamic range of pyramidal imager and the enhancement FDR that bouncing scanning provides.

5.4.1 Non-Bouncing Scanning data analysis

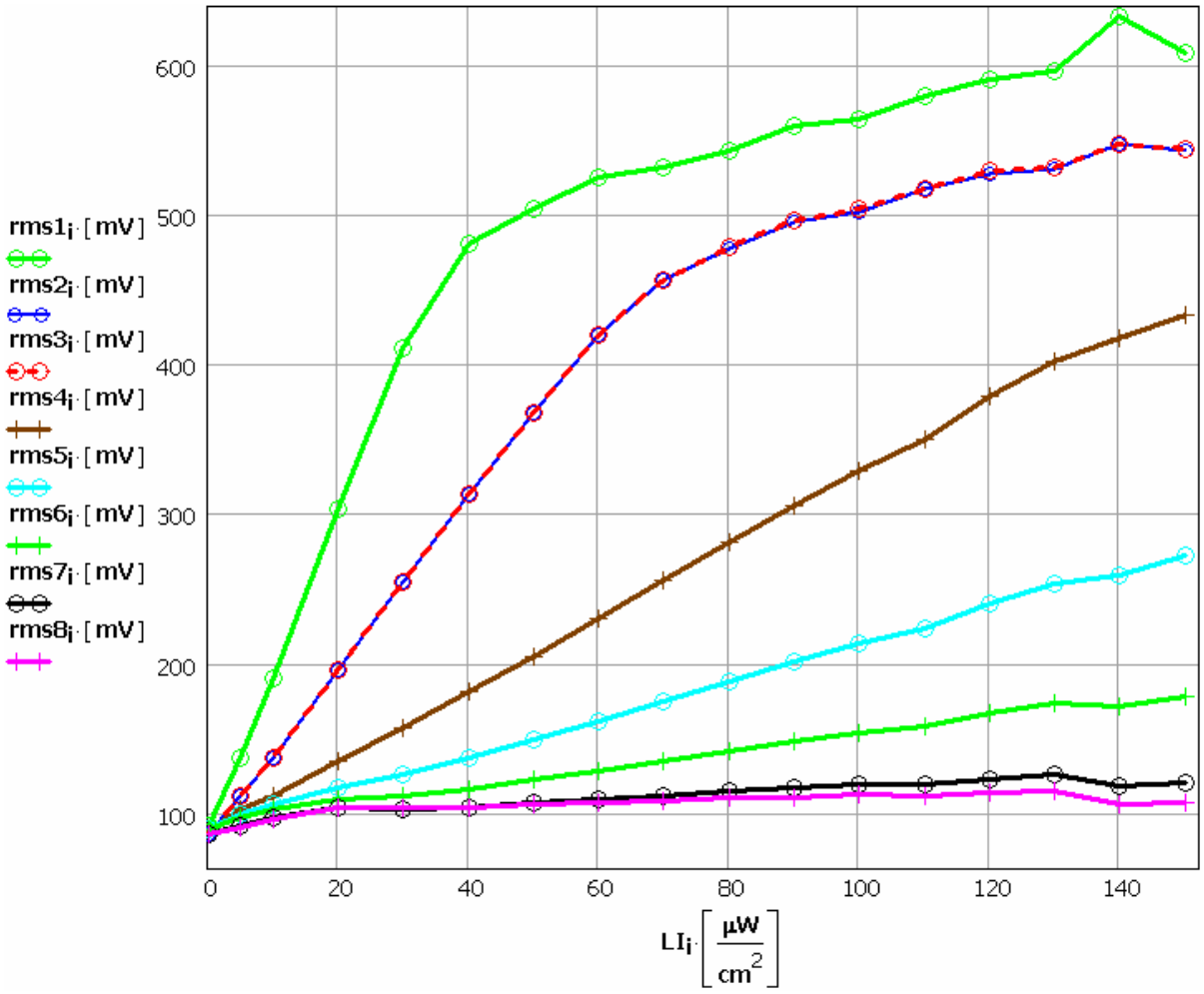


Fig 5.8 Vcds in RMS voltage the whole pyramidal imager for 8 integration times

The plots rms1, rms2,…rms7, rms8 in Fig 5.8 represent the root-mean-squared (RMS) values of Vcds in millivolts (mV) for the 8 sampling frequencies 10KHz, 20KHz, 25KHz, 50KHz, 100KHz, 200KHz, 500KHz and 1MHz, corresponding to integration times of 124.8 ms, 62.4ms, 49.92ms, 24.96ms, 12.48ms, 6.24ms, 2.496ms, and 1.248ms, respectively.

5.4.2 Sensitivity

Sensitivity is an intrinsic imager characteristic that shows the change of the output voltage with illumination. From Fig 5.8 we extract the sensitivity of the pyramidal imager by calculating the slopes of the different linear regions and plotting these slopes versus their corresponding integration times. The sensitivity of the imager (for Vcds) is the slope of the curve in Fig 5.9, namely $0.088\frac{V\,cm^2}{\mu W\,s}$.

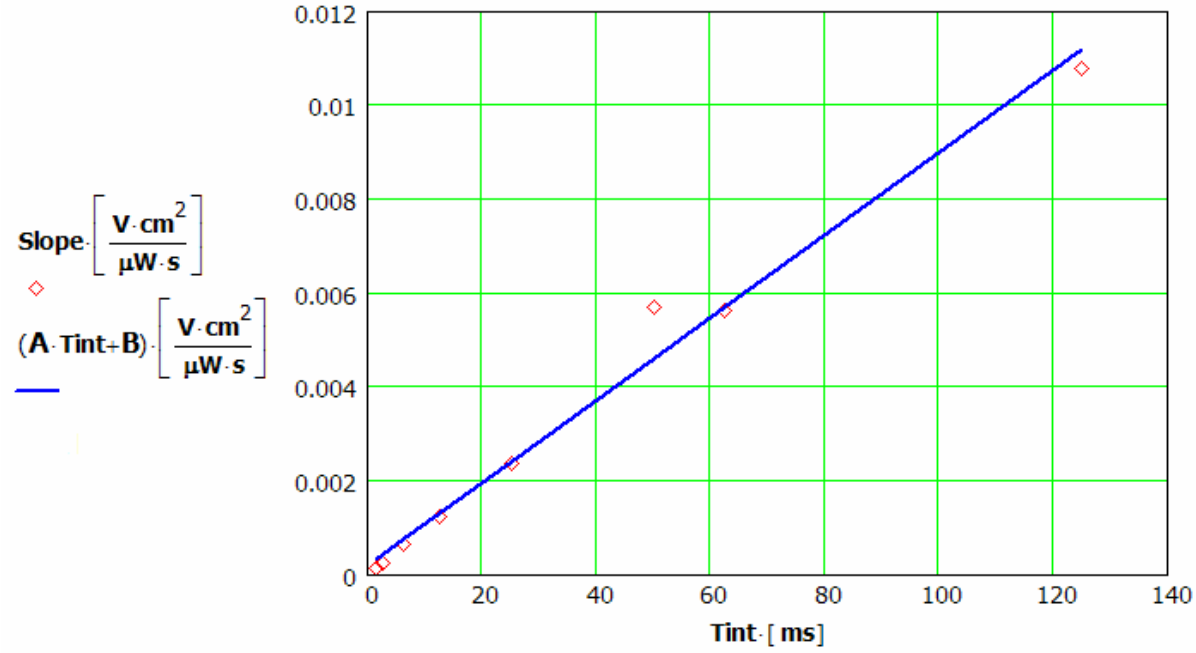


Fig 5.9 Slopes of linear regions of photon transfer curve in Fig 5.8

Applying the same procedure on Vs (instead of Vcds) leads to the similar sensitivity value of $0.091\frac{V\,cm^2}{\mu W\,s}$. This is a good match confirming that the pyramidal imager sensitivity value extracted is correct.

5.4.3 Photon-response Modeling

Using the extracted sensitivity value, one can verify the linearity of the imager with a simplistic model and hence extract the behaviour of the pyramidal imager in a bouncing scanning regime. Furthermore, the effect of the bouncing scanning on the deviation of the imager from the model will be discussed. The modeling of the imager is also useful for extracting the optical dynamic range of the imager and determining the enhancement that has been achieved by using this scanning scheme. These results are compared to the theoretical analysis presented in the previous chapter.

From Fig 5.6, we have developed out our model for the Vcds (called hereafter VCDSout) knowing the value of the sensitivity and the by visual inspection of the graph in Fig 5.8 we can extract Vmin and Vsat (limits of linear region) and we find:

$$
\begin{aligned}
&\text{VCDSout}(LI, Tint) := \\
&\quad Vmax \leftarrow 0.48062V \\
&\quad Vmin \leftarrow 0.1049V \\
&\quad Ss \leftarrow 0.088 \cdot \frac{V}{\mu W \cdot s} \\
&\quad Vout \leftarrow Ss \cdot LI \cdot Tint + Vmin \\
&\quad Vout \ \text{ if } \ Vmin < Vout < Vmin \\
&\quad Vmin \ \text{ if } \ Vout \leq Vmin \\
&\quad Vmax \ \text{ if } \ Vout \geq Vmax
\end{aligned}
\qquad \ldots (5.1)
$$

Similarly we get the following for VS (or VSout)

$$
\begin{aligned}
&\text{VSout}(LI, Tint) := \\
&\quad Vmax \leftarrow 2.6V \\
&\quad Vmin \leftarrow 2.2V \\
&\quad Ss \leftarrow 0.091 \cdot \frac{V}{\mu W \cdot s} \\
&\quad Vout \leftarrow -Ss \cdot LI \cdot Tint + Vmax \\
&\quad Vout \ \text{ if } \ Vmin < Vout < Vmax \\
&\quad Vmin \ \text{ if } \ Vout \leq Vmin \\
&\quad Vmax \ \text{ if } \ Vout \geq Vmax
\end{aligned}
\qquad \ldots (5.2)
$$

The above model is not the exact model that represent the photo-transfer characteristic of CMOS imagers because of the lack of knowledge regarding the exact value of Vmin and correspondingly L_{min} after which the imager start to linearly follow the light intensity [2]. Vmin (and hence also L_{min}) correspond to noise floor of the imager which depends on the integration time and hence is not fixed. This will be clearer when examining the model plot in Fig 5.10 beside the plots shown in Fig 5.8 in the Fig 5.9.

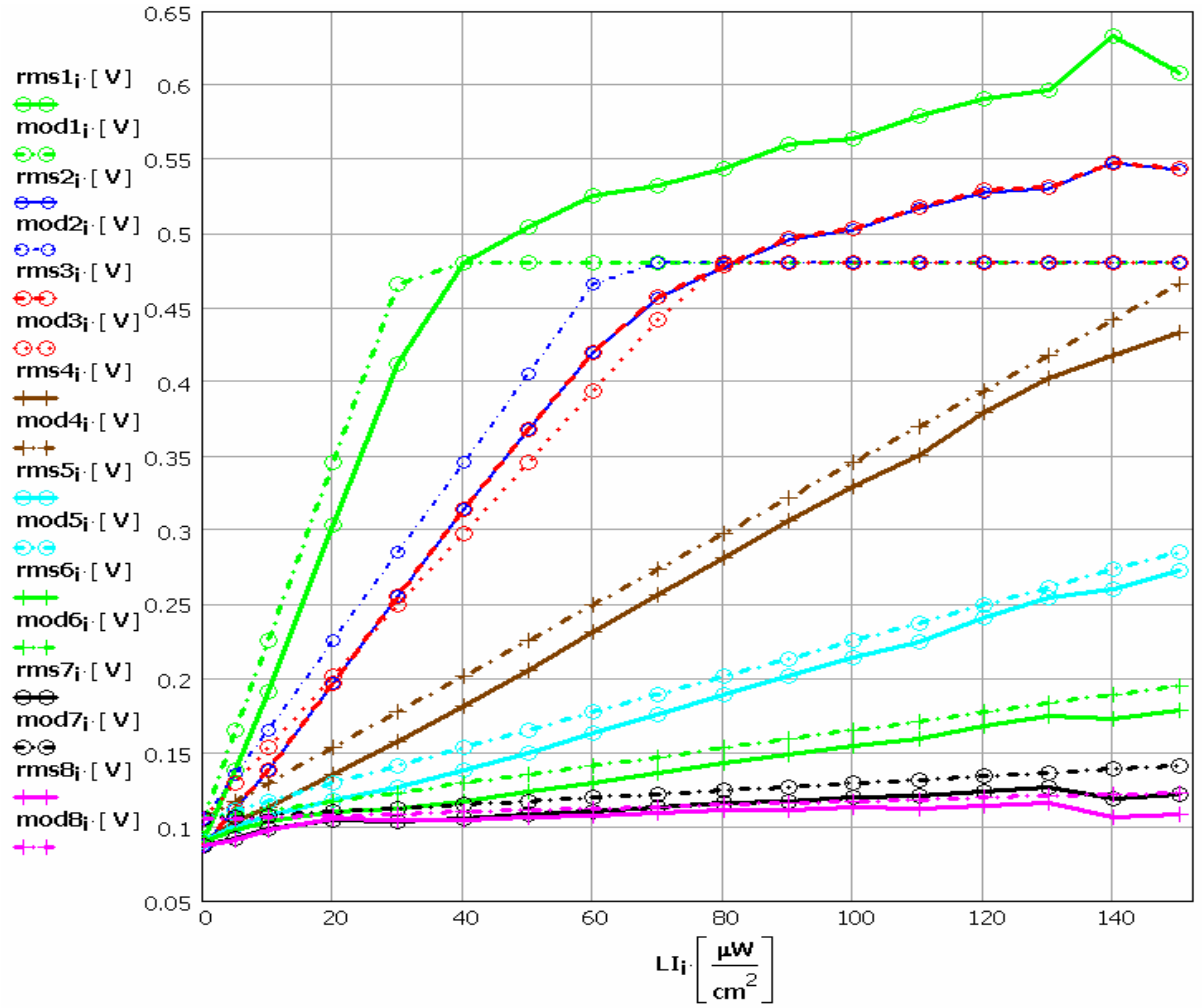


Fig 5.10 Model plots and exact values of Vcds

It is clear the gradients of the modelled curves follow the experimental photo-response curves of Vcds reasonably well, with the exception of some offset. In order to have a better vision about how well the model fits we plot the correlation of each Vcds photo transfer curve and its corresponding model curve and visualize the result over integration time as shown in Fig 5.11.

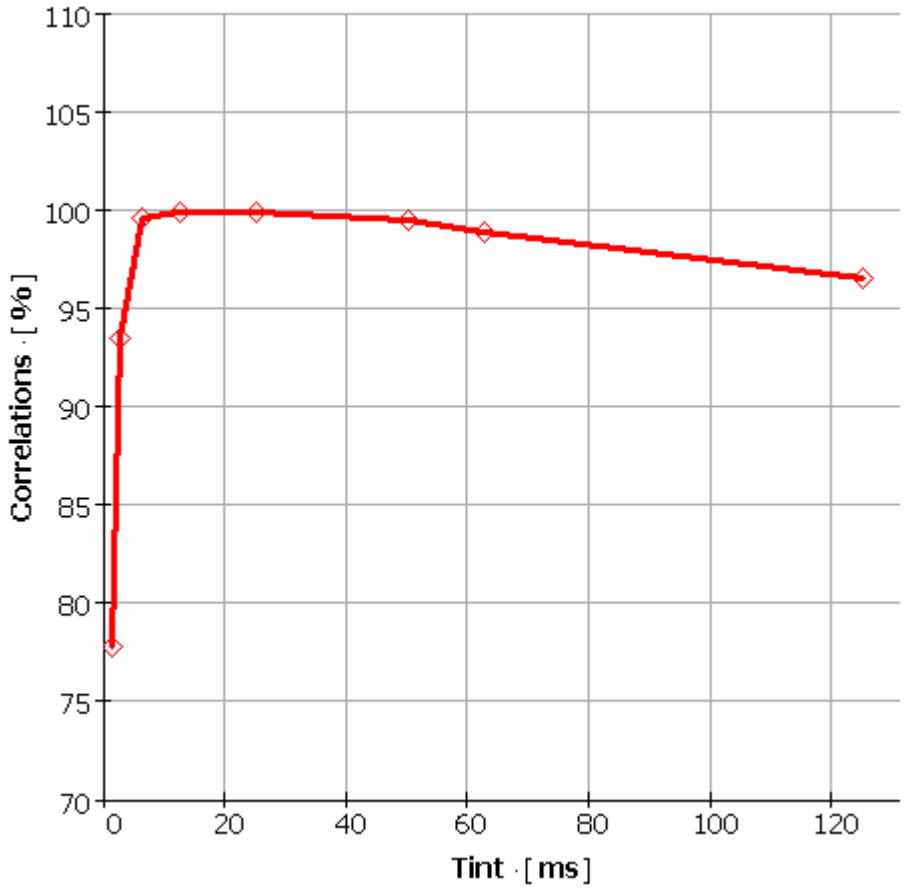


Fig 5.11 Correlation between the model exact curves of Vcds versus integration time.

From above plot in Fig 5.11 we can see good correlation between the model and the Vcds curves but this correlation decrease for longer integration times and decreases further with short integration times. This means that the Vmin used in the in model was precise only for intermediate integration times such as 6.24ms, 12.48ms and 24.96ms. We will estimate later how this assumption affects the dynamic range enhancement calculation.

5.5 Bouncing Scanning and Dynamic Range calculation

In this section we will present our method that enables us not only to get the intrascene dynamic range enhancement through the experimental data but also the dynamic range of the imager itself with the bouncing scanning. Therefore, in the next paragraphs will be dealing with the system dynamic range and the enhancement dynamic range.

In order to calculate the dynamic range, we need to know the maximum light intensity detectable without saturation by the pyramidal imager (L_{max}) as well as the minimum light intensity above which the imager starts its linear conversion of light intensity of electrical signal (L_{min}). Then by applying the equation 5.2 one can get the value of the imager dynamic range that encompasses both the system dynamic range and the enhancement:

$$DR = 20\log\left(L_{\max}/L_{\min}\right)\ldots \quad (5.3)$$

The value of L_{min} corresponding to Vmin is not clearly extractable; by visual inspection we estimate the $L_{min} \approx 5\mu W$. Now, in order to estimate DR we will use the equation 5.1 (or 5.2) to calculate L_{max} at which Vcds will reach its maximum value before saturation. We get for L_{max} for both inward and outward scanning and then we choose the maximum between the two values and use it in equation 5.3 for determining the system dynamic range.

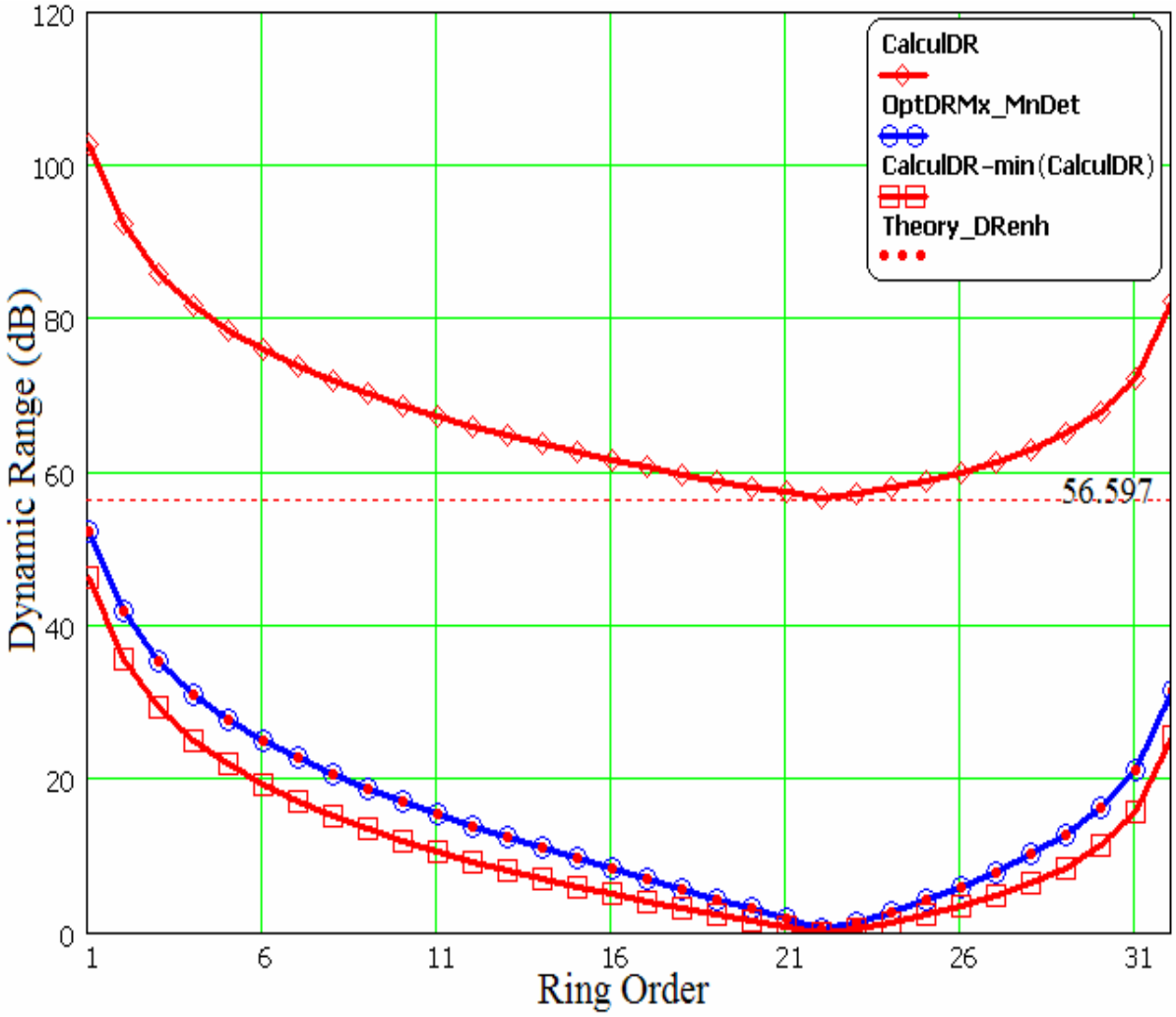


Fig 5.12 The pyramidal imager system and enhancement dynamic ranges

In the graph plotted in Fig 5.12, *CalculDR* is the calculated dynamic range of the pyramidal imager as described in the previous paragraph (equation 5.3), *OptDRMx_MnDet* is dynamic range calculated from equation 5.4 below and finally, *Theory_DRenh* is the dynamic range calculation from equation 4.7.

$$OptDRMx\_MnDet(r) = 20\log\left[\frac{\max(MaxDetLI_{inward}, MaxDetLI_{outward})}{\min(MaxDetLI_{inward}, MaxDetLI_{outward})}\right] \ldots (5.4)$$

First observation of the plots in Fig 5.12, which have been extracted at sampling frequency of 1MHz, we can see the *Theory_DRenh* curve and *OptDRMx_MnDet* curve coincide. This is demonstrates that the intrascene dynamic range enhancement using two [68] or many [93] integration times primarily relies on expanding L_{max} rather than minimizing L_{min}. This comes in accordance with the development of the DR enhancement techniques as discussed in section 2.6.

We have seen in section 4.2 that there is no DR enhancement for ring 23 (in fact in between ring 22 and 23) and at therefore from the plot in Fig 5.12 we can easily deduce that the system dynamic range = 56.6dB which corresponds to the minimum of *CalculDR* at ring 22. Finally, after subtracting this value of the system minimum dynamic range from the imager dynamic range *CalculDR* we get the intrascene dynamic range enhancement plotted with square point in Fig 5.12. It is clear that the measured (*CalculDR*) and calculated (*OptDRMx_MnDet*) foveated dynamic range enhancement is close to the expected foveated dynamic range enhancement (*Theory_DRenh*) demonstrated in Chapter 4.

In the following, some acquired images using bouncing and non bouncing (rolling) scanning are shown to demonstrate the foveated dynamic range enhancement achieved with the former scanning scheme. Note the dark ring (at ring Nº 23) in the images that is due to a mistake in the layout[16] disabling the reset of that ring which its Vs values cancels its Vr values after

[16] VIA23 between metal layer 2 and 3 connecting the reset decoder and the reset ring (ring No 10) was missing.

CDS operation resulting a black ring in the image. Besides, the starry view of the images is caused by the fixed pattern noise due to the dark current of localized lattice defects in the silicon substrate as mentioned in 2.1.6.

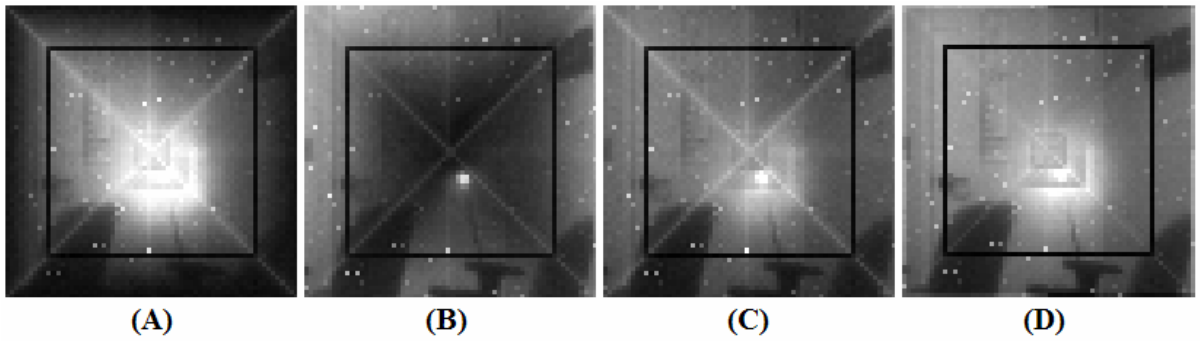


Fig 5.13 Demonstration of foveated dynamic range enhancement at the foveal rings

In Fig 5.13 are shown the inward scanned image (A), the outward scanned image (B), the fused image[17] of (A) and (B) shown in (C) and the rolling scanned image (D). These images were taken with the pyramidal imager sampling the incident scene light intensity of 270 lux. For comparison sake, the fused image (built from inward and outward scanned images) and the rolling scanned image are sampled both at 8 fps (frame per second), which requires the bouncing scanned (the inward and outward) images to be sampled twice as fast as the non bouncing image. There sampling rate of the bouncing scanned images is 20 KHz while that of the rolling scanned images is 10 KHz.

[17] The fused image of (A) and (B) is achieved by averaging their sum.

In the central region of the sampled image (or the imager's fovea) a small light-bulb is set over a chair. The shape of bulb is more visible in the fused image (Fig 5.13.C) than in the rolling scanned image (Fig 5.13.D) that is faster approaching saturation in the bright region. The saturation effect is also visible in the inward scanned image (Fig 5.13.A), whereas its is absent in the outward bounced scanned image (Fig 5.13.B) in which the bright spot of the bulb is seen as small bright point. This is due to integration time profiles of the inward and bounced scanning enabling the latter to sample brighter light intensities than could the former (nor the rolling) scanning sample before reaching saturation. This explains the extended optical dynamic range of the fused image, in this case at the fovea region of the imager as predicted in Chapter 4.

The figures shown in Fig 5.14 were taken at a higher light and sampling rate levels in contrast to the previous set of acquired images, however, they represent the same nature and order of pictures as defined previously but for a different scene. The scene shown in Fig 5.14 is of the author showing is right hand and back-lighted by a bright lamp.

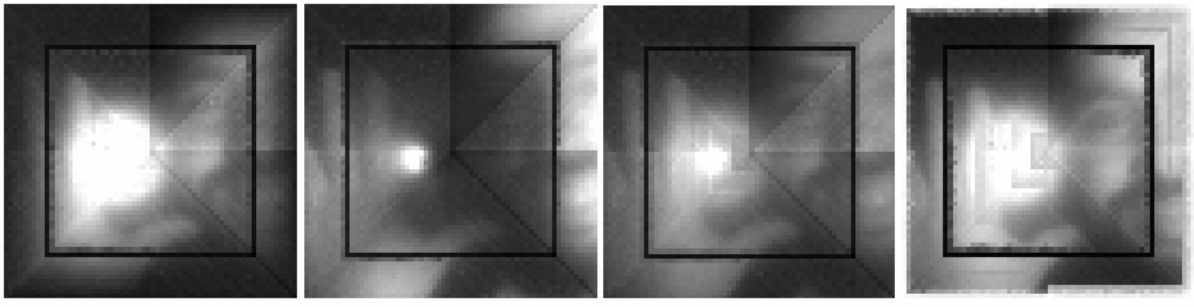


Fig 5.14. Manifestation of foveated dynamic range enhancement at the bouncing rings

The light flux at the imager was measured to be at 402 lux delivering frame rates for both scanning schemes of around 400 fps and corresponding to a sampling rate of 1MHz in the bouncing scanning and 500 KHz in the rolling scanning. Despite the fact of visible noise at the boundaries of the rolling scanned image (Fig 5.14.D) and the missed ring (ring Nº 23) mentioned earlier, the foveated dynamic range enhancement is more visible than in Fig 5.13. This fact is not a surprise recalling that the dynamic range is inversely proportional to the integration time and thus its enhancement is therefore more visible.

The enhancement of range in Fig 5.14 is noticed in both bouncing boundaries, the fovea rings (at the central part of the imager) and the imager's outer rings as anticipated in the previous Chapter 4. For fovea rings, the hand details looks more visible in fused image (Fig 5.14.C) than in the rolling scanned image thanks to the inward scanned image extra sampled details nearby the bright spot. As for the outer rings dynamic range enhancement, it is visible through the details of the left eye noticeable only in the fused image thanks to the same argument mentioned earlier for the inner rings.

It is now concluded that FDR enhancement is visible in measured images, as predicted in the previous chapter's mathematical prediction. It should be noted that the magnitude of the optical dynamic range improvement in this sensor is limited to less than the ideal value by the high dark currents for this fabrication process.

5.5.1 Foveated Dynamic range and Video Communications

Video communication is of crucial importance for current and future image sensors. The ability of the imager to sample an image and communicate it in an optimal fashion, in terms

of transferred data, dynamic range and frame rate, is critical for this kind of application. Video communication has many application areas such as video phones and remote imaging such as space imaging and security systems. Our Pyramidal imager is a good example of such applications, especially videophones because of the limitation of the human eye in temporal sampling enables the use of fast imaging cameras to integrate the image sequences at the eye level. The pyramidal imager as seen earlier has the high frame rate property due to its high degree of parallelism in terms of video channels as well as the sampling dimension. Hence, using this property the bouncing scanning scheme can be adopted for sampling images and integrating inward and bounced images at the human eye. This will certainly enhance the intrascene dynamic range at the human eye without a need of local memories to make the fusion of both acquired images.

5.6 Summary

We have shown in this chapter an indirect method in estimating the pyramidal imager dynamic range without enhancement. In addition, we have shown the calculated foveated intrascene dynamic range enhancement is close to the theoretical foveated dynamic range developed in the previous chapter. Finally, based on the inherent high speed image sampling of the pyramidal image sensor we suggested to use the human visual system at an integrating system for the two frames resulting from the bouncing scanning which will result a foveated dynamic range in the observer visual system.

Chapter 6

Pyramidal Imager Fixed Pattern Noise Low Perception by the Human Visual System

Exploiting the limitations and characteristics of the human visual system has been widely used by image science, image processing and image compression, in order to minimize the complexity and hardware cost of many visual communication and display products and even lighting products. Examples include the persistence of vision to blend consecutive images in one smooth picture for TV and movies. Similar use is made in lighting devices such neon lamps. On the other side, and we shall discuss later that the human visual system is known by its impressive adaptation capability that takes some time to reach its steady state and this fact is actually the key factor the imaging engineers use in order to create the wanted illusion in the human observer's vision.

In this chapter we explore of the human visual system in discriminating oblique patterns (lines tilted by 45° or 135° compared to the cardinal axes at 0° and 90°) for resolving fixed pattern noise FPN. Firstly, we will discuss the topology of fixed pattern noise in classical image sensor architecture and in pyramidal architecture showing that FPN is distributed between columns that are obliquely oriented in the latter architecture. Then, we shall construct a novel spatial filter equivalent to the pattern discrimination filter of the human visual system in the fovea region of the retina based on empirical data. The resulting filter is

applied to the pyramidal image to determine how its FPN noise is perceived by a human observer [94].

6.1 Fixed Pattern Noise

The definition of fixed pattern noise as well as its sources is discussed in section 2.1.6. In the next section the topology of this noise is discusses further in both the classical CMOS imager and the Pyramidal CMOS imager.

6.2 Fixed Pattern Noise Topology in Classical CMOS imagers

The topology of the fixed pattern noise is found by analyzing the path the photo-signal undergoes off the imager chip is many caused by the different mismatches between these different paths used to buffer out the imager photo signals. The mismatch in the photo-signal transportation is due the variations that the generated photo-signal faces on its path off-chip such as those in pixel source follower. This particularly includes the mismatch of the offset and gain of the column amplifiers, double sampling circuits and other column-based systems. In the classical orthogonal CMOS imager architectures these cause a striped noise distribution [95] as shown in the following Fig 6.1.

Fig 6.1.a and Fig 6.1.c are typical images taken by a 64x64 CMOS image sensor before and after correlated double sample denoising respectively [95]. Transferring the previous images into Fourier domain gives a nice picture about the how FPN noise topology is distributed on the classical imager's rows and column. Fig 6.1.b and Fig 6.1.d are the Fourier transform

spectrums of Fig 6.1.a and Fig 6.1.c respectively, plotted in log-10 scale to highlight frequency spectrum details with emphasis on their high values.

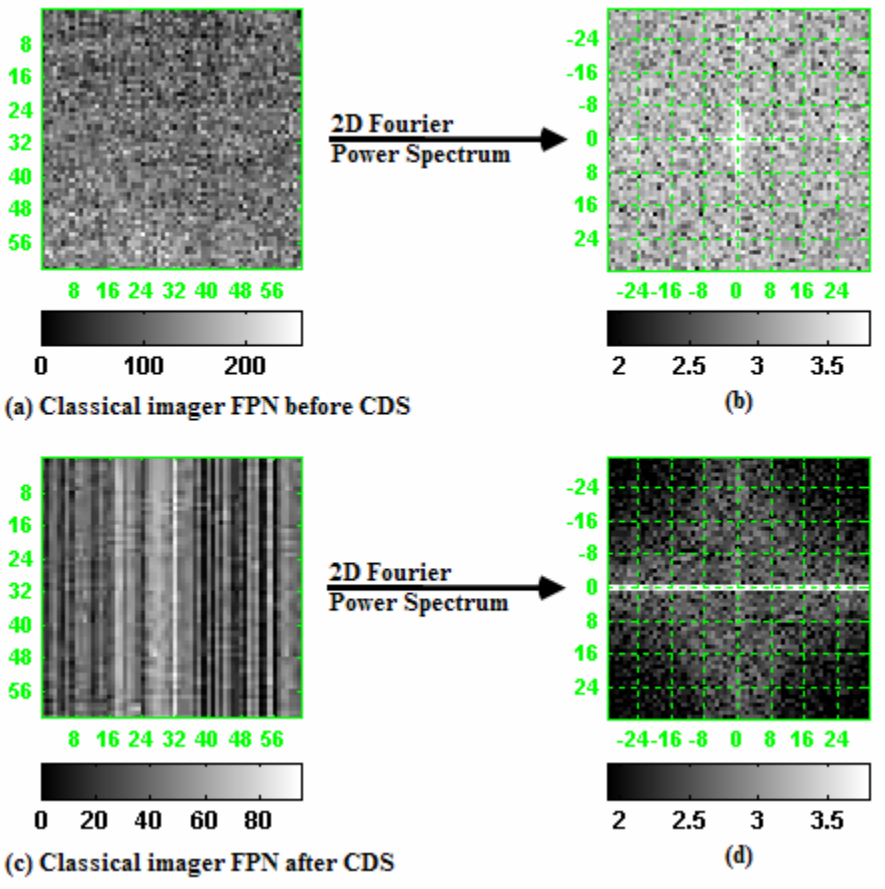


Fig 6.1 FPN noise topology in classical CMOS imager sensor

It is clear that the classical imager output signal without CDS is highly uncorrelated, except for a short-narrow band of frequencies on the vertical axis near the DC region. This implies some correlation on the horizontal directions of Fig 6.1.a which are basically rows. This suggests that there is a short range correlation between pixels on the same row while there is almost no correlation between pixels on the same column. We note that Fig 6.1.a and Fig

6.1.b include all the different noise sources present in the imager before CDS de-noising. The situation after CDS filtering is the reverse, with a higher degree of correlation between pixels on the same column than between adjacent pixels of the same row. This is clearly seen in Fig 6.1.d with the highest frequencies being concatenated on the horizontal axis and those of low energies scattered around it, which is just the reflection of the vertical stripes of the CDS filtered image as shown in Fig 6.1.c. The above frequency analysis is in good agreement with the results mentioned in [95].

6.3 FPN Topology in Pyramidal CMOS Imager

In the pyramidal architecture, diagonals have replaced columns of the classical orthogonal CMOS imager. At the base of these diagonals reside the sample and hold circuits. The different distribution of the FPN from column stripes in the classical CMOS imager to diagonal distribution topology is shown below. We do not show the image of our imager FPN before CDS filtering because we are more interested in noise that remain in final image.

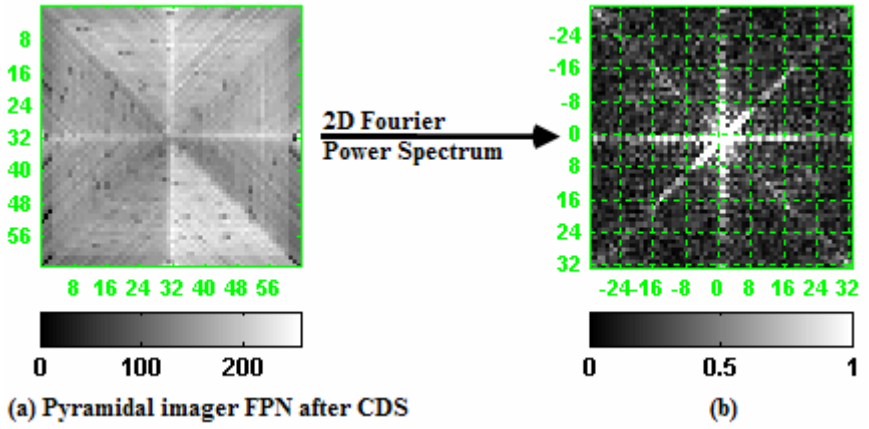


Fig 6.2 FPN topology in a 64x64 pyramidal CMOS image sensor

It is clear from Fig 6.2.a that diagonal stripes remain in the image after CDS even away from segment boundaries. This distribution is also shown by the diagonal Fourier spectra shown in Fig 6.2.b. This is a very interesting result especially if we consider the limitation of the human visual system (HVS) in resolving obliquely distributed contrast, known as the oblique effect [96] (see below).

6.4 HVS Pattern Sensitivity

In this section we develop a new model for the HVS, including the relative insensitivity to obliquely oriented patterns. This is then applied to the FPN images to evaluate how a human see the FPN images of a pyramidal CMOS image sensor. Fig 6.3 shows the oblique effect found in the contrast sensitivity of the HVS at relatively high frequencies [97].

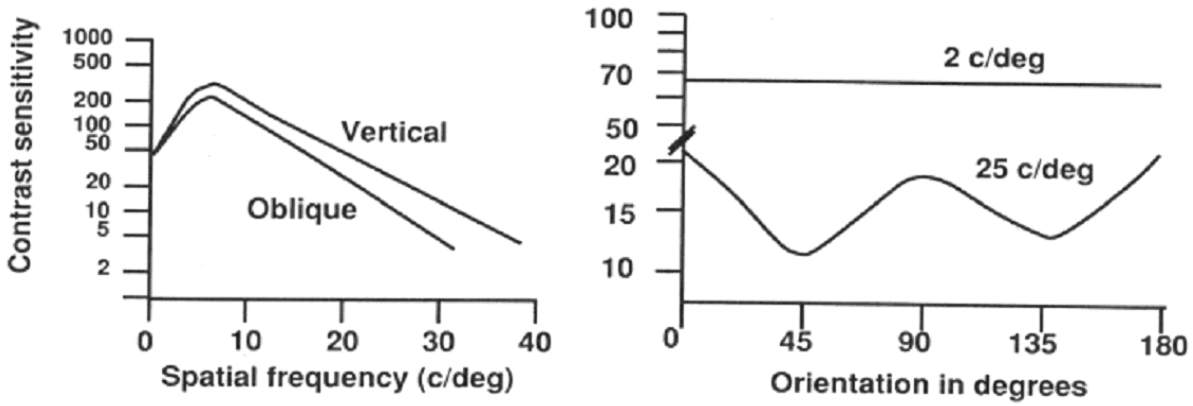


Fig 6.3 Oblique effect in the human visual system

6.4.1 HVS Spatial Filter Model

The human visual system is a complex system with many levels of image acquisition and image processing layers. We have restricted our scope to the spatial image acquisition of the HVS or what is widely known as the spatial vision. We adopted the spatial filtering modeling developed by Hugh R. Wilson [98] in which the HVS is believed to process spatial patterns in parallel, utilizing at least six different ranges of spatial frequencies and perhaps a dozen different preferred orientations. This model was based on the assumption that a spatial filter for the HVS may be thought of as the psychophysical equivalent of a physiological receptive field. Thus, given a psychophysical unit centered at a particular point in the space of vision (or a single cell with a receptive field centered at a particular point on the retina), its 2D spatial filter determines the sensitivity of the unit to image luminance at each point in the visual space. In order to determine the characteristics of the individual filters, which correspond to the tuning properties of underlying visual mechanism, a well established

technique known as pattern masking [98] was used. By applying this psychophysical measurement technique and analysis on human individuals it has been shown that masking data obtained in the fovea over a range of test frequencies from 0.25-22.0 cycle per degree (cpd) were consistent with the operation of the six spatial frequency tuned mechanisms plotted in Fig 6.4.

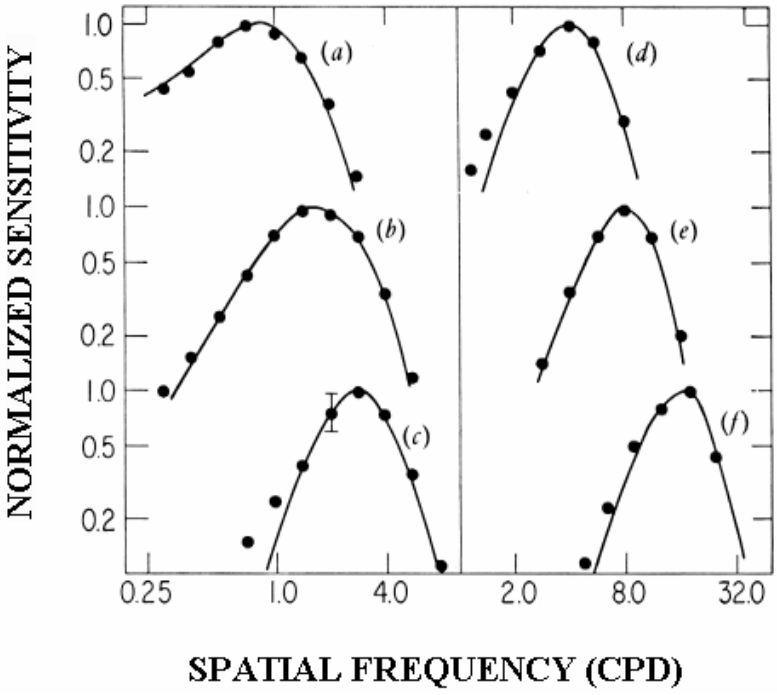


Fig 6.4 The six human spatial pattern sensitivity filters [98]

The above graphs have been well fitted [98] with the following function for the receptive fields (or HVS spatial filters in spatial domain).

$$RF_i(x,y) = A_i\left(e^{-x^2/\sigma_{1i}^2} - B_i e^{-x^2/\sigma_{2i}^2} + C_i e^{-x^2/\sigma_{3i}^2}\right) e^{-y^2/\sigma_{yi}^2} \quad \ldots (6.1)$$

The variables x and y represent spatial coordinates in the fovea in degrees or pixels and A_i, B_i, C_i, $\sigma_{1,2,3i}$ are fitting parameters of the suggested function of the receptive fields response. The above function describes a vertically oriented filter centered at the origin. For filters at different locations tuned to preferred orientations other than the vertical, one simply uses this equation in conjunction with the familiar equations for translation and rotation of coordinates. In order to compare this function with the masking data results, one has to use the Fourier transform of equation (6.1) which is shown with solid lines in Fig 6.4 in good agreement the acquired data. The parameters shown in equation (6.1) have been discussed in [98] and are shown in table 6.1.

Table 6-1 Fitting parameters of the equation (6.1) for some fovea spatial filters [98]

Mechanism (Spatial Filter)	**Peak Frequency (cpd)**	**Peak Sensitivity**	**A**	**B**	**C**	**σ_1**	**σ_2**	σ_3
A	0.8	30.0	123.19	0.267	---	0.198°	0.593°	---
B	1.7	70.0	596.59	0.333	---	0.098°	0.294°	---
C	2.8	140.0	2046.13	0.894	0.333	0.084°	0.189°	0.253°
D	**4.0**	**150.0**	**3141.85**	**0.894**	**0.333**	**0.059°**	**0.132°**	**0.177°**
E	**8.0**	**76.7**	**5129.43**	**1.266**	**0.500**	**0.038°**	**0.060°**	**0.076°**
F	**16.0**	**18.4**	**2457.71**	**1.266**	**0.500**	**0.019°**	**0.030°**	**0.038°**

6.4.2 Oblique Effect

The spatial vision in human visual system as well as in many animal species is not anisotropic. It is more sensitive to stimuli (spatial patterns) that are oriented vertically or horizontally compared to those obliquely oriented. This effect is also defined as by the greater contrast needed by oblique contours to become visible. This reduced effectiveness of oblique contours compared to horizontal or vertical ones is referred as oblique effect [99]. This phenomenon has been originally noted by Ernst Mach [100] and later found in children and in numerous animal species [101]. Although many studies have proven the existence of the oblique effect for both detection and discrimination, its origin remains still a mystery. The reader can refer to [99] to review these attempts that tried to understand the functional origin of the oblique effect beside those authors' own attempt in this regard. Despite the inability to pinpoint the cause of the oblique effect in the HVS, it is almost evident that the oblique effect is a result of an adaptation of the human visual system to match the spatial spectrum of most of the scenes human being sees in his everyday life [102][103]. Fig 6.5 shows how anisotropic the spatial frequency power spectrum average of 500 images of various scenes such as persons, animals and pastoral landscapes [103]. It is clear that the bandwidth of this averaging power spectrum is wider in the cardinal axis (horizontal and vertical) than in its oblique axis (45º and 135º).

Fig 6.5. Average spatial power spectrum distribution of about 500 scenes [103]

6.4.3 HVS Spatial Filter Construction Including Oblique Effect

The model for building the HVS spatial filter is based on the work done by Wilson and published in [98]. It is based on the assumption that a spatial can be thought as the psychophysical equivalent of physiological receptive field. In other words, a given psychophysical unit centered as a particular point in a visual space (or a single cell with a receptive field centered at a particular point in the retina), its two dimensional spatial filter, designed earlier as $RF(x,y)$, determines the sensitivity of the unit to image luminance at each point in visual space. Consequently, the linear response of a unit can be calculated simply by convolving $RF(x,y)$ by the image luminance or equivalently convolving the Fourier transforms of both the filter and the image luminance. Although the response of HVS to contrast is non-linear, this fact can be adequately handled by introducing an appropriate

nonlinearity following the filtering operation. What we are interested on at this stage is the neural image that the human visual cortex receives from the retina where each cell in its central region, called fovea, will exhibit the spatial response to the impinging image luminance based on the receptive field function $RF(x,y)$ shown in equation 6.1 and using the empirical parameters values from in Table 6-1.

The human retina is a multilayered structure on top another starting by ganglion cells that face the incident light and finishing by the photoreceptor cells layer that host the chemical photo conversion process converting incident light to a detectable photocurrent [104] sensed by the connecting neurons that convert it into neural pulses. The retina anatomy reveals that the human visual acquisition system is composed of two categories of cells, the rods and the cones. The distribution of the visual cells (or photocells) is however not uniform, the cones being highly dense on the central region of the retina, named the fovea, and the rods being spread all over around in an almost circular symmetry [4]. Fig 6.6 shows the location of the fovea in the human eye as well as the distribution density of rods and cones in this part of the human eye.

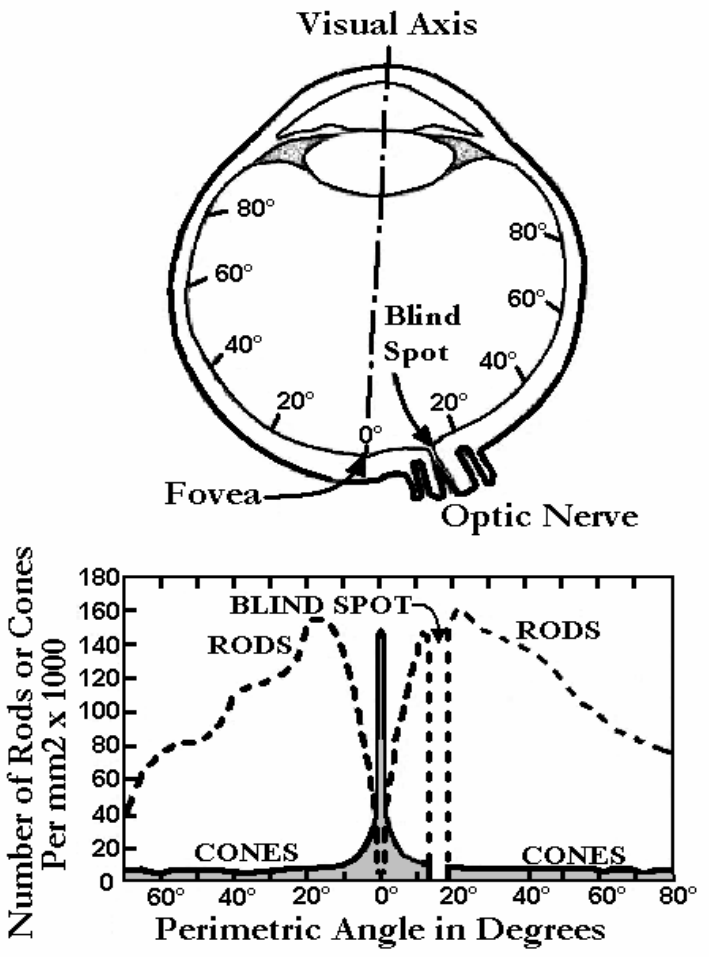


Fig 6.6 The distribution of rods and cones photoreceptors in human eye [4]

Fig 6.7 shows the overall anatomy of the human retina at the fovea region.

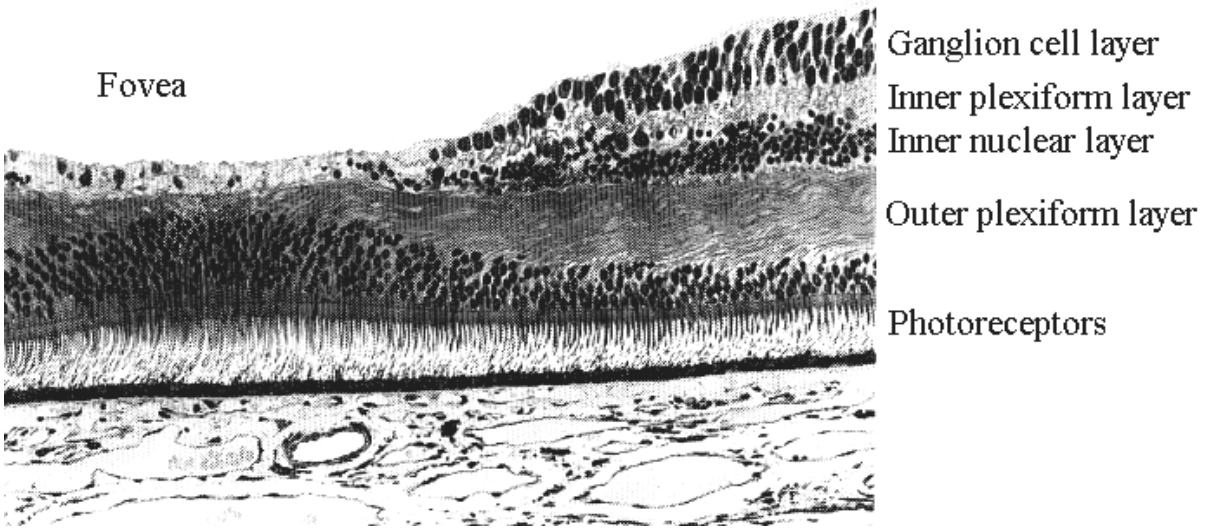


Fig 6.7 Cross section of the human retina near the fovea region [4]

To our knowledge, no spatial filter model exists for the oblique effect. Accordingly, the model described in previous section is used to construct a HVS spatial filter including the oblique effect by combining filters corresponding to filters (D) (E) and (F) from Fig 6.4 using their parameters mentioned in Table 6-1. These filters have been chosen because we have chosen to model the acquisition of the pyramidal FPN image as if it was viewed by an eye at a distance such that it will occupy only a small portion of the fovea that extends only to a 1° field of view around the eye visual axis. This region is also called foveola and it about 350μm wide around the optical axis containing about 120 cones and is shown in Fig 6.7 by the retina region that has no layer on it except the photo receptors [4]. Because our pyramidal image sensor is just 64x64 pixels, we need to expand the size of the image to fit the 120 photocells to cover the 1° extension over the fovea. Hence, the imager picture will be

zoomed to double of its size pitch to (64x2)x(64x2), i.e. 128x128. Besides, to have the pyramidal imager FPN image laid on the foveola over 1º of expansion, the image needs to be put in front of the eye at a distance D and printed with edge equal to H where the two parameters must obey the following formula that simply says that the image expansion on 1º of fovea is equivalent to a viewing angle of 1º and is shown in Fig 6.8 [4].

$$\frac{H}{D} = \tan\left(1^0\right) = 0.017 \ldots (6.2)$$

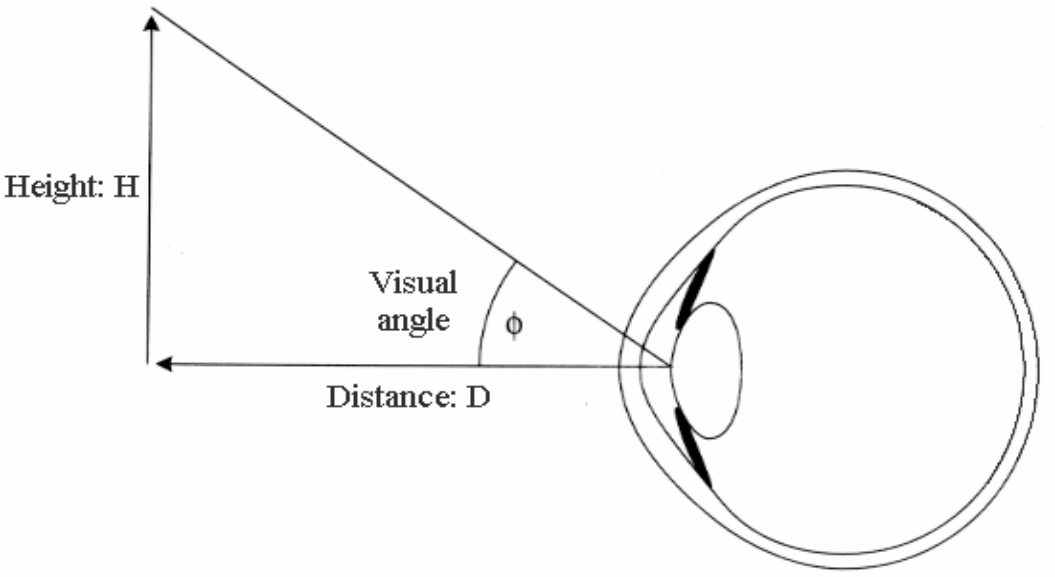


Fig 6.8 Viewing angle calculation for image construction at the retina

The foveola spatial extension will have significant response only from filters (D) (E) and (F) based on their corresponding peak sensitivities. Filters D, E have been rotated by 8 equally spaced directions 0º, 22.5º, 45º, 67.5º, 90º, 112.5º, 135º and 157º before being added in

spatial domain. Furthermore, and in order to have the oblique effect of the human vision included in the human spatial filters, the filter F have been added to the previously constructed filter only on cardinal axes (horizontal and vertical axis) before transforming the resulting filter in Fourier domain. Filter F will extend the final filter bandwidth on the cardinal axes to about two times the bandwidth of the HVS spatial filter at the oblique angles namely at 45° and 135°. This corresponds well to the measured response of the human eye to contrasting strips at various angles [96] (see Fig 6.3). The final filter is shown in Fig 6.9 where ω_x and ω_y represent the spatial frequency in cycles per degree (cpd).

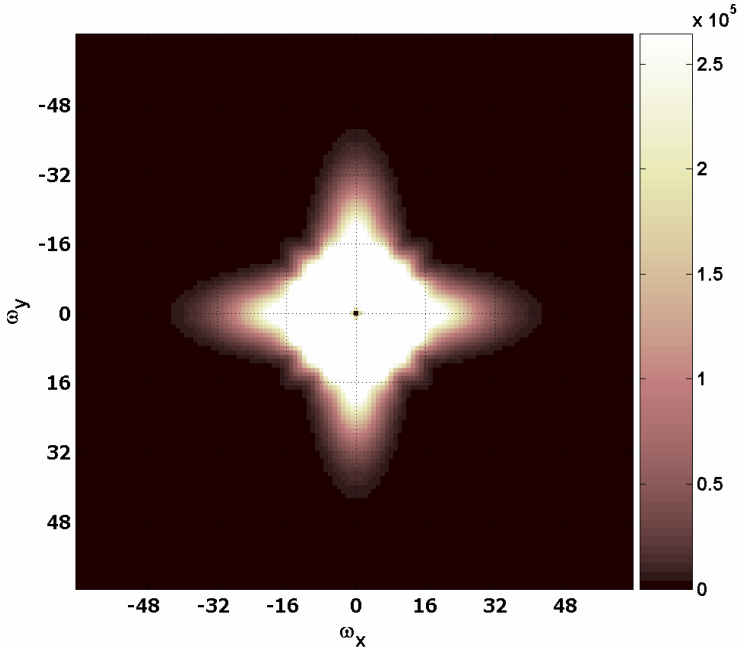


Fig 6.9 Fourier spectrum of the constructed HVS spatial filter

6.5 Pyramidal Imager FPN Perceptibility by HVS

In order to estimate the effect of the HVS spatial filter to the obliquely distributed FPN noise of the pyramidal CMOS imager, a comparison is made between the response of the HVS spatial filter and a the same filter but tilted by 45º. This tilting is made in Fourier domain which is equivalent in the spatial domain as rotating the observed head by the same angle of 45º following the rotation property of the Fourier transformation [105]. These two filters and their spatial equivalents are shown in Fig 6.10. In this tilted arrangement the diagonals of the pyramidal sensor appear to the eye in the same way as a normal raster scan sensor. The reason of creating the second filter (HVS filter tilted by 45º) is to verify whether or not the HVS spatial filter has truly sensed the FPN noise of the pyramidal imager (see assumption below). The tilted HVS filter spectrum as well as its equivalent tilted face are shown in Fig 6.10.b2 and Fig 6.10.b1 respectively, while the normal face up HVS system along with its equivalent filter spectrum are shown in Fig 6.10.a1 and Fig 6.10.a2 respectively.

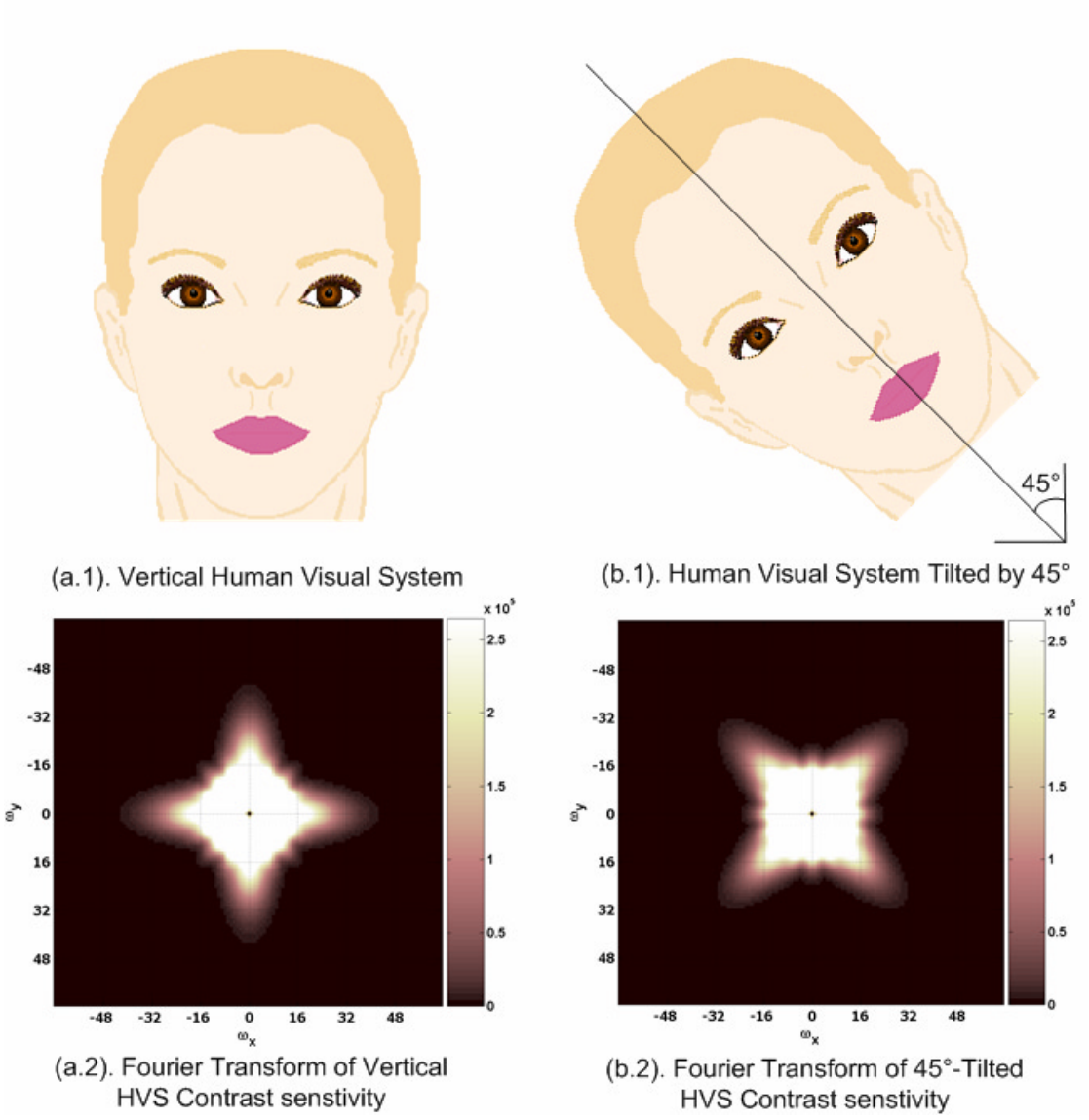


Fig 6.10 Pyramidal FPN noise perception by HVS experiment

It is assumed that differentiating the response of the original normal HVS filter to pyramidal FPN image from that of the tilted HVS filter will show in principle whether or not the oblique FPN noise of the pyramidal CMOS imager is perceived less by the HVS.

6.5.1 Results and Conclusion

Fig 6.11 below shows the steps taken to demonstrate how FPN based on the diagonal columns is perceived at a reduced level by the normal HVS. The tilted HVS is used for comparison because it effectively acts as if the columns are vertical with respect to the eye. Fig 6.11.A is the pyramidal FPN image and both Fig 6.11.D and Fig 6.11.G are its Fourier spectrum. Fig 6.11.E and Fig 6.11.I are the HVS normal vertical and 45° tilted filters respectively. Fig 6.11.F is the product of normal HVS filter and the FPN image.

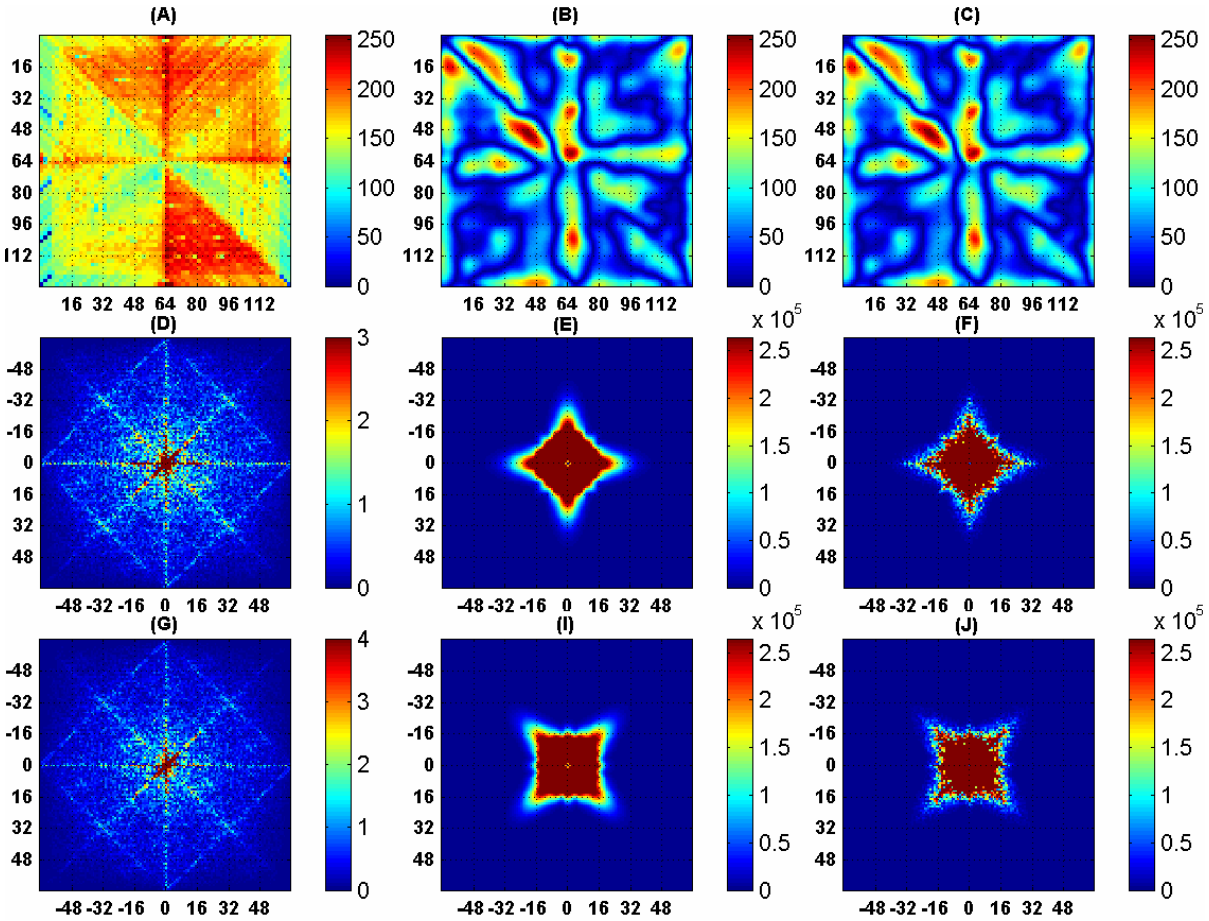


Fig 6.11 HVS perception verification steps operations.

Fig 6.11.J is similarly the product of the tilted HVS filter and the pyramidal FPN image. Finally, the figures Fig 6.11.B and Fig 6.11.C are the reverse Fourier transforms of Fig 6.11.F and Fig 6.11.J respectively. At first sight, both images Fig 6.11.B and Fig 6.11.C look very similar however if we subtract the normal HVS filtered image (Fig 6.11.B) from the tilted HVS filtered image (Fig 6.11.C) we get the following image in Fig 6.12.

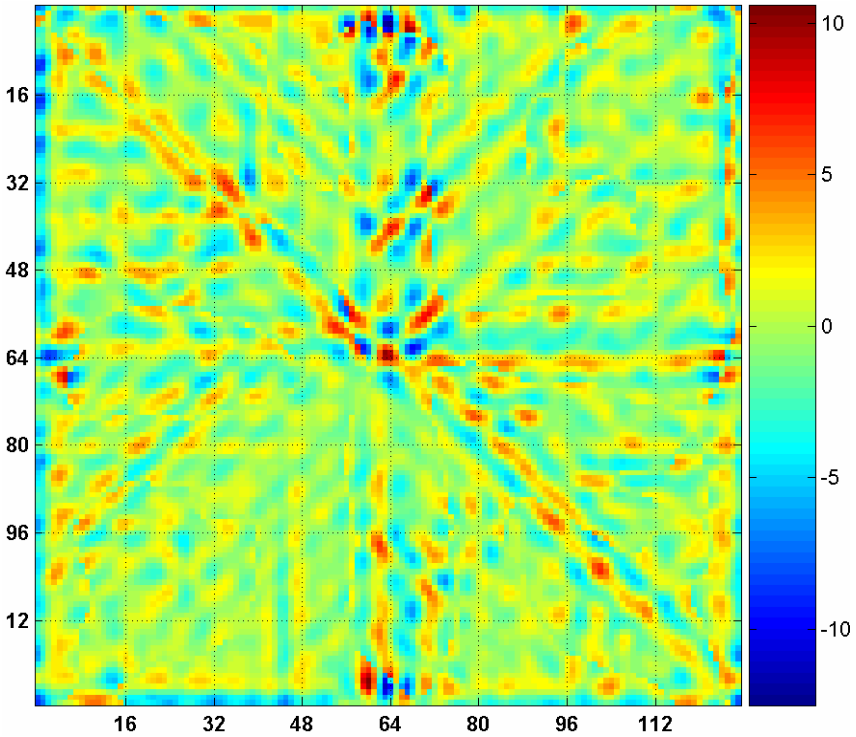


Fig 6.12 The subtraction image of Fig 6.11.B from Fig 6.11.C.

Fig 6.12 shows many positive stripes that are obliquely distributed either along 45° or along 135° axis that are existing on the tilted HVS filtered imager and absent on the normal HVS filtered imager.

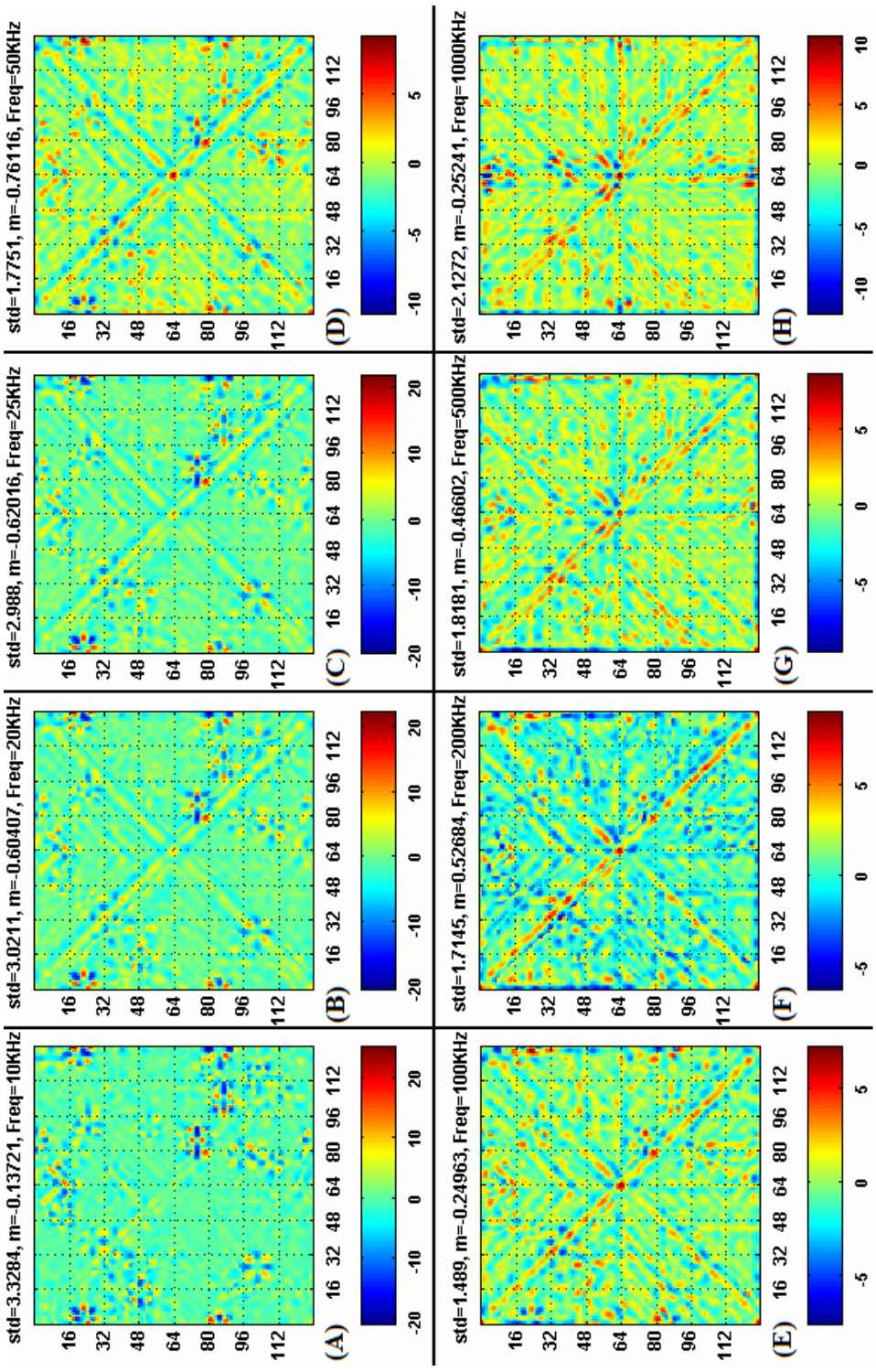


Fig 6.13 Frequency analysis of oblique FPN suppression by HVS

To further analyse this result the above methodology has been applied over images of fixed pattern noise acquired at various frequencies and the results are displayed in Fig 6.13. As the integration time decreases with increasing sampling frequencies, more FPN noise starts to be apparent because of the creasing SNR. The topology of the FPN noise also increases its manifestation from vague random at 10 KHz to obliquely distributed above a sampling frequency of 50 KHz. In Fig 6.13 for each image the standard deviation **std** as well as the mean **m** of the image has been calculated. Following the evolution of these characteristic numbers, it is clear that very little correlation is observed between the mean or standard deviation and the image noise. This low correlation holds true in both the noise amount, as in between images of Fig 6.13.A and Fig 6.13.D, and noise topology, as in between Fig 6.13.B and Fig 6.13.F. This is to show the image quality quantification is still far from being fully solved problem and that confirms why human observation is an important part of image processing task [62]. Finally, to further demonstrate the reduced perception of the pyramidal CMOS image FPN, one can notice the positive oblique noise residues remaining after subtracting the non tilted HVS spatial pattern filter from the oblique counter part and the close-to-zero[18] values of the subtraction elsewhere especially visible in Fig 6.13.G and Fig 6.13.F. This shows a clear evidence of the absence of the obliquely distributed FPN noise in the HVS filtered image and its presence in the tilted filter and hence proves the low perception of the pyramidal CMOS image FPN noise by a human observer [106].

[18] This region corresponds to the similarity in the low frequency content present in both filtered images which is expected from the shape of both filters.

6.6 Summary

This chapter is a good example of an interdisciplinary research case namely between the electrical engineering responsible of design CMOS image sensors and human vision that interacts with the display of the acquired images. The discussion about the pyramidal CMOS imager has been developed in previous chapters and in this chapter we mainly discussed human vision spatial pattern filtering property. After discussing the physiology of the retina we used psychophysical analysis that lead us to the construction of the HVS spatial filter including the oblique effect. This inclusion served us later to demonstrate qualitatively the low perception of the human visual system to the obliquely distributed FPN noise of out pyramidal CMOS imager sensor.

Chapter 7

Multiresolution CMOS Image Sensor

7.1 Introduction

Images consist of data representing spatial information (and temporal information in the case of video images) which can be analysed by software programs to extract certain information for clustering and classification purposes. These image processing tasks go slower and consume larger electrical and computational power with increasing image resolution and therefore ultimately result in a bottleneck for high speed imaging applications. Consequently, data reduction sampling architectures are needed to facilitate image processing without compromising the useful information required by image processing algorithms. One promising method for achieving selective data reduction is the concept of a multi-resolution sensor in which the sampled image contains regions of the highest possible resolution covering the region(s)-of-interest (ROI); other regions of less importance can be sampled at lower resolutions. Similar vision architectures are found in biological vision systems [4]. We think that CMOS image sensors are the only candidate for such vision systems due to their architectural flexibility compared with their CCD counterparts.

7.2 Multiresolution CMOS Imagers

A particularly interesting image sampling technique dedicated for low power and high-speed imaging, known as multiresolution image acquisition, has recently attracted the interest of

many researchers ([107], [108], [109], [110]). In the previous implementations of multiresolution image sensors, averaging of the photo-signals within a cluster of pixels, or kernel, was used to reduce resolutions in regions of non-interest while maintaining the maximal imager resolution in the area(s) of interest. Multiresolution CMOS image sensors were inspired by the pixel binning technique used in CCD technology to enhance the photo-signal to noise ratio and frame speed of the camera at the expense of reduced spatial resolution. Thus, multiresolution imaging is meant for enhancing the temporal resolution at the expense of spatial resolution. Adding the advantage of random accessibility found in CMOS image sensors, programmable pixel binning (or averaging) over the whole CMOS image sensor becomes a relatively straightforward implementation. Multiresolution CMOS imaging has been so far implemented at the column and chip (frame) levels, requiring relatively large control circuitry. The reason for the high silicon area consumption is because pixel averaging is performed outside the pixel, and requires analogue memories, such as capacitors, to store pixel-sampled voltage.

In our implementation, the pixel binning is realized inside the image sensor array in between the pixels themselves. This allows the multiresolution decoders to have higher controllability over the region-of-interest size and location (instead of using the banks of shift registers as in previous designs). The schematic and layout of the suggested multiresolution pixel are shown in Fig 7.1 and Fig 7.2 respectively.

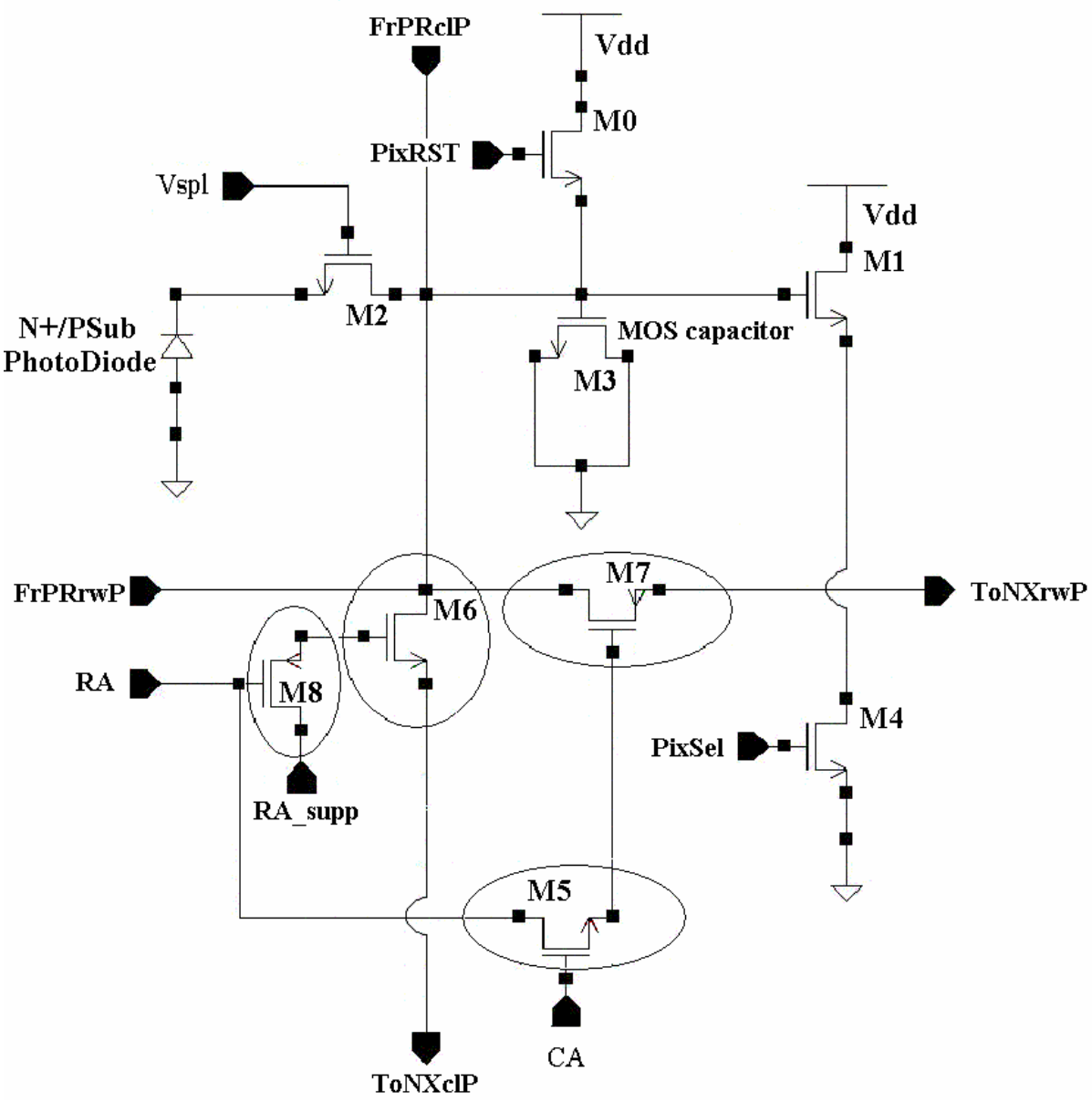


Fig 7.1 Schematic of the proposed multiresolution pixel[19]

[19] For the description of the multiresolution pixel input and output signals please refer to section 7.4

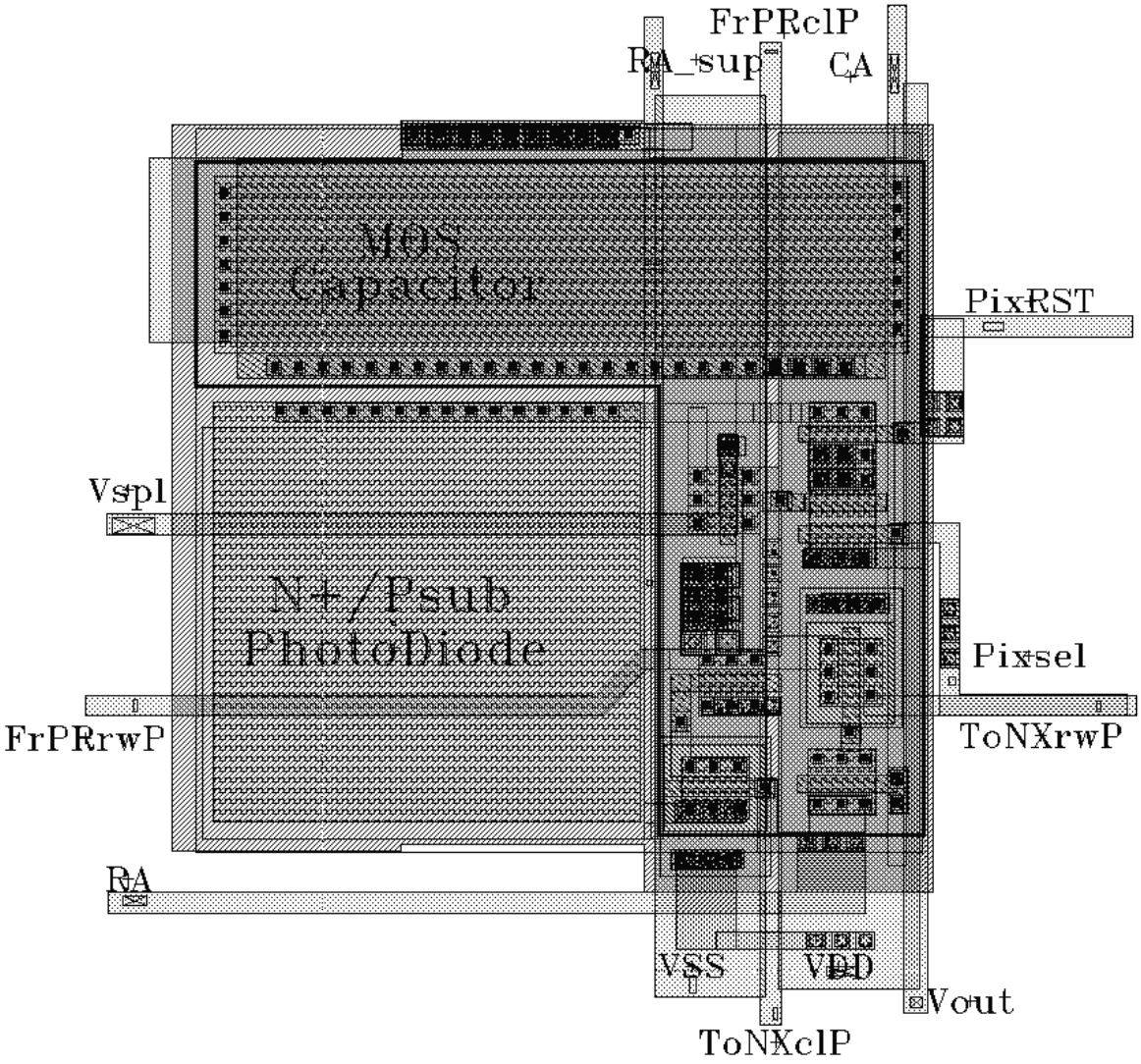


Fig 7.2 Layout of the multiresolution pixel

7.3 Multiresolution Pixel Structure

As shown in Fig 7.1, the pixel structure comprises eight MOSFETs, two PMOS and six NMOS, as well as an NMOS capacitor and an N+/Psub photodiode. Transistor M0 is the reset transistor for both the photodiode and the sampling NMOS capacitor M3. The

photodiode is separated from the sampling capacitor M3 by a shuttering NMOS switch M2 that is locally controlled by one of the lateral multiresolution decoders. We revisit these decoders later when we discuss the whole system operation. The last classical components of the standard CMOS photodiode pixel are M1 and M4, which represent the pixel buffer (source follower) and select transistors respectively. Finally, we get to the structure of the multiresolution at the pixel level comprising the transistors M5, M6, M7 and M8. The NMOS transistor M5 and the PMOS transistor M7 are responsible for column averaging, while the PMOS transistor M8 and the NMOS transistor M6 are responsible for row averaging.

7.4 Pixel Averaging and Readout

Before going further into detail, it is worth mentioning that the photodiode cannot be reset above V_{DD}-V_T due to V_T drop across M0, as can be concluded from the diagram shown in figure 4. The multiresolution structure is controlled by three digital signals namely, row average (RA) signal, row average support (RA_supp) and column average (CA) signal as shown in figure 4. The photodiode and the sampling capacitor M3 are reset first at the beginning of the integration time through transistors M2 and M0. Then the shuttering transistor M2 is opened to allow the photodiode to integrate the incident light while the NMOS capacitor M3 is being reset. At the end of the integration time, the reset transistor M0 is open while the shuttering transistor M2 is closed, for a short period to allow the sampling of the integrated photo signal by the NMOS capacitor. At this stage, the sampled photo charge is averaged through the switch network as follows. The kernel to be averaged is

programmed though the multiresolution decoders that generate the control signals RA, RA_supp and CA as shown in figure 6. When RA signal is high PMOS transistor M8 is OFF regardless the signal RA_supp. When RA becomes low and RA_supp is high (VDD), M8 is ON and the gate of the NMOS M6 is at VDD enabling it ON, and the photo charge is mixed between current pixel NMOS capacitor and the next pixel in the same column (ToNXclP). This effect does not happen when RA and RA_supp are both low. This concludes the row averaging mechanism. For the column averaging to occur, only the photo-charge of the current pixel and that of the next pixel in the same row (ToNXrwP) have to be mixed. This is only possible when the NMOS transistor M7 is ON, in which case it can only happen when the CA is high and the RA is low. This is a very important result in building the charge mixing, and hence voltage averaging, mechanism leading to the implementation of any arbitrary kernel size and resulting in a highly flexible multiresolution configuration.

7.5 Pixel Multiresolution Implementation

The proposed multiresolution technique has been implemented in 0.18 μm Salicide CMOS technology (briefly described in section 3.1) with 64x64 pixels and three fundamental kernels namely 2x2, 4x4 and 8x8. We call them fundamental kernels because of their square shape, but this is not a limitation of the proposed multiresolution scheme. Hence, through the multiresolution decoders, one can program kernels such as 2x4, 4x2, 2x8, 8x2, 4x8 and 8x4 of rectangular shapes either horizontally or vertically. This is a specific advantage of the present suggested multiresolution scheme over the one suggested earlier [107]. This particular advantage allows different resolution and hence different spatial filtering in the

horizontal and the vertical image axes of the sampled image. The high-speed imaging property of the multiresolution CMOS image sensor is realized through the sub-sampling of the output of the kernels. Therefore, only one pixel of a given averaged kernel is readout in the sub-sampling mode without having spatial aliasing in high frequency regions, eliminated through the spatial frequency low-pass filtering property of spatial averaging.

A physical description of the multiresolution CMOS image sensor layout is illustrated in the following Table 7-1.

Table 7-1 Physical characteristics of the multiresolution pixel

PIXEL CHARACTERISTICS	PHYSICAL VALUE
Pixel pitch	**15.292 μm**
Pixel area	**233.845 μm^2**
Active area	**79.644 μm^2**
Fill factor	**34.06 %**
Active area perimeter	**35.698 μm**
Power supply voltage	**3.3 V**

One notable characteristic of the multiresolution APS is its low fill factor due the implementation of charge averaging (in addition to electronic shuttering) circuitry in pixel. The multiresolution CMOS imager has been under an extensive test for a long period of time

however, no satisfactory photo-signal has been recorded to-date. The reason for this is still unclear.

The methodology to test our multiresolution CMOS imager was based increasing functionality complexity step by step. This means to acquire the whole sampled image at the highest spatial resolution and disabling the use of in-pixel sample and hold circuitry by keeping the shutter transistor open all time. With this setting the multiresolution imager was exposed to variable lighting conditions at different biasing levels of the imager source followers. No dependence between the acquired electrical signal and the incident light intensity was found. The nest step that was investigated was to get back to the layout and re-simulate its constituent blocks and verify their interconnection. This step did not show any inconsistencies. Back to the testing setup, all the Printed Circuit Board (PCB) (mounting the multiresolution imager) connections were tested with the board input and output ports. Yet again, no anomaly was found and all the connections were consistent. The last component of the testing setup namely the testing pattern was verified and the signals were exactly what they were intended to control in imager under test. All these steps were carried many times and unfortunately, no mistake has been found the testing setup and thus the only possibility remaining to us was a mistake in the fabrication process of the multiresolution imager. It was only lately suggested to send the imager to Ottawa-based professional laboratory known as MuAnalysis Incorporation[20] specialised failure analysis of integrated circuit using optical

[20] http://www.muanalysis.com/

techniques. This remains a future work due to the time and cost limitation of the present research programme.

Nevertheless, some simulated results based on the functionality of the imager have been achieved and are shown in the following figures.

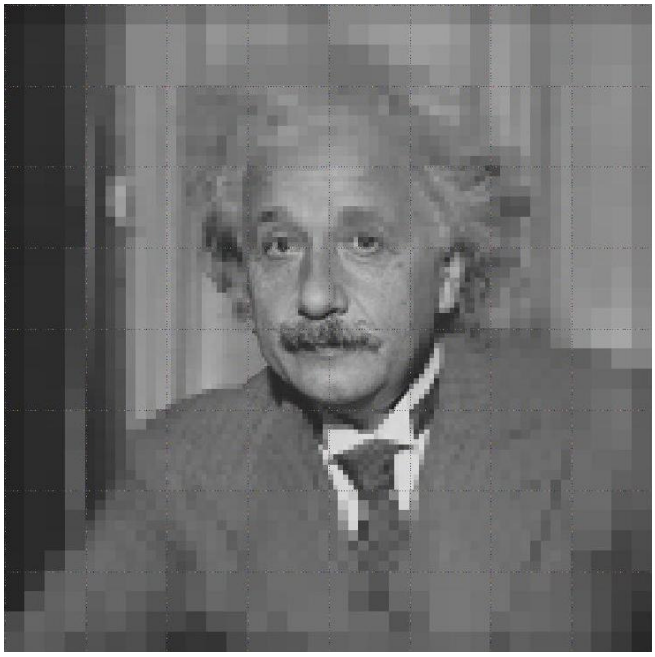


Fig 7.3 A multiresolution image of centric foveation

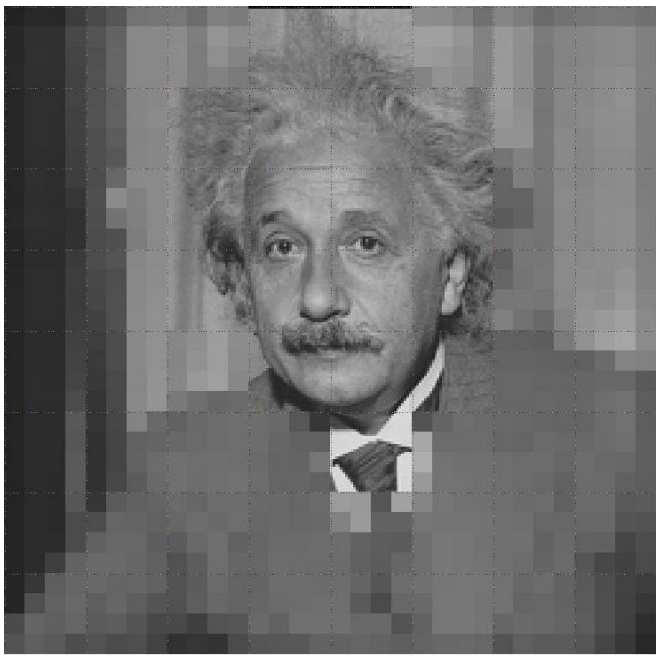


Fig 7.4 A multiresolution image of random foveation

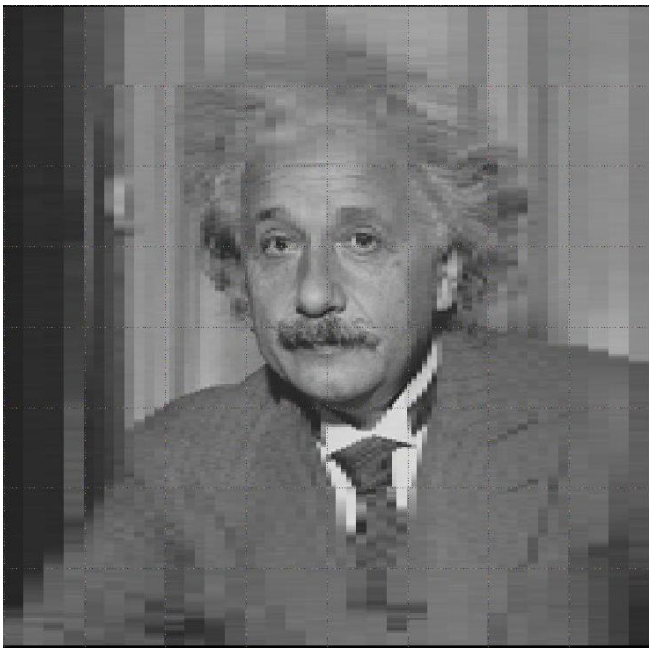


Fig 7.5 A multiresolution image with a horizontal kernel averaging

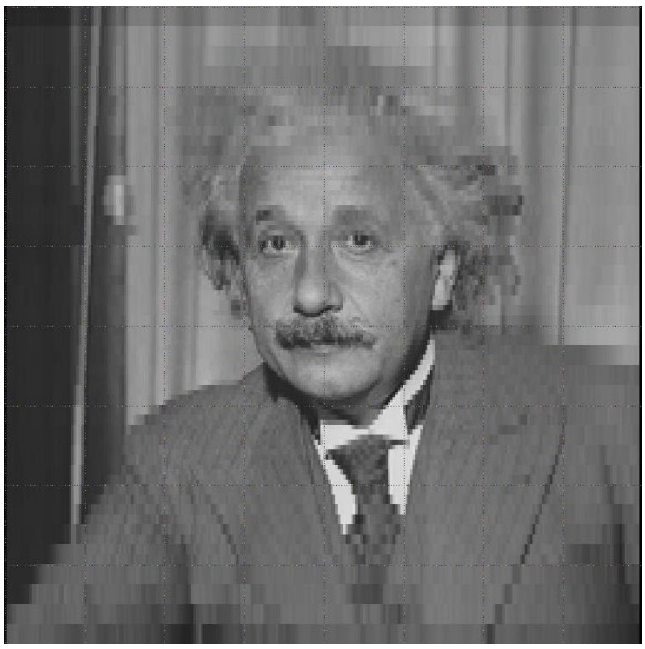


Fig 7.6 A multiresolution image with a vertical kernel averaging

Fig 7.3 shows simulation results of the multiresolution acquisition with all the available averaging kernels of square shapes with decreasing sizes from the outer area towards the inner region that does not suffer from any averaging. This is similar to the human fovea that is centric in the human vision where charge sharing mimicking in interconnectivity configuration of the ganglion cells hooked to the peripheral rod photocells [4]. The next figure, Fig 7.4 shows the programmable foveation capability of the multiresolution imager to any region of the imager's focal plane. The last figures, Fig 7.5 and Fig 7.6 shows the importance of averaging kernel shape in preserving or destroying the spatial frequency content of the resulting image. Thus, for example, the tie shape was clearer with the vertical kernel than with the horizontal kernel averaging because of its vertical shape.

Chapter 8

Conclusion

With increasing market demand for CMOS image sensors with higher resolutions and frame rates, data throughput can reach a bottleneck. New sampling architectures and scanning schemes are potential methods for minimizing the transferred data without affecting (or with minimal effect on) the information carried by this data. Towards this end, benefiting from architectural flexibility of CMOS imaging technology, a new sampling architecture called the pyramidal architecture has been suggested. Inspired by biological vision, the pyramidal CMOS imager was implemented on the basis of 2D ring sampling and diagonal output buses instead of the 1D row sampling and vertical buses used in classical CMOS image sensors. This hardware transformation of the sampling process has been further supported by a new scanning scheme called bouncing scanning to replace the raster scanning widely used in classical CMOS imagers.

This change of sampling paradigms has many consequences that are mathematically analysed in Chapter 4. It has been found that the inherent parallelism in pyramidal image sampling results in a high frame rate capability compared to that of the classical CMOS imagers. On the other hand, bouncing scanning and fusing the two resulting frames provides a dynamic range enhancement that is higher in the inner ring (at central part) of the pyramidal imager and decreasing outwards. This shape of dynamic range enhancement is analogous to that of the human fovea, so the enhancement is called foveated DR.

Practically, fusing two frames to extend the dynamic range is costly in terms of memory to store the two frames. A remedy to this problem would be to include a local memory (analog or digital) and an adder locally at the pixel level which will somewhat decrease the fill factor and the imagers spatial resolution. Further development of CMOS technology will allow in future integrating more functionality at the pixel level.

Chapter 5 presents the experimental verification of the foveated dynamic range. It did not include the noise analysis to determine the dynamic range of the imager either with or without the bouncing scanning. Instead an indirect method has been used to estimate the dynamic range and its foveated enhancement (when using the bouncing scanning). While this is an acceptable method in CMOS imagers' characterization, the explicit inclusion of the noise would bring more useful insight about the impact of the bouncing scanning on the imager noise performance, especially at the bouncing edges. Despite this fact, experimental results have shown acceptable matching between theoretical and experimental foveated dynamic range enhancement profiles.

The interaction between the pyramidal imager output and the human visual system has been introduced in Chapter 6. The fixed pattern noise topology of the pyramidal CMOS imager was analysed and been demonstrated to follow an oblique distribution. The human visual system is known to be less sensitive to oblique patterns compared to the cardinally (i.e. horizontal and vertical) distributed spatial frequencies. A model of the HVS pattern sensitivity based on empirical data has been constructed and applied on an FPN dominated pyramidal CMOS imager sampled data. This analysis was carried out on images acquired at

increasing sampling frequencies, with which FPN is known to increase as well. The qualitative comparison between the filtering performance of the normal and the tilted (by 45º) HVS filter was used to verify the presence (or absence) of the FPN noise in the HVS filtered image. The analysis has shown that a certain amount of the pyramidal FPN noise was present in the tilted HVS spatial filter and absent in the tilted filter, which implies that the oblique FPN noise was filtered out by the human vision system spatial pattern filter. The reduced perception of the pyramidal imager FPN noise by a normal human observer was therefore evident. The present application of HVS modeling to the design of CMOS image sensors is, to our knowledge, the first of its kind. By incorporating an understanding of how the image is perceived by the viewer, system resources can be deployed optimally.

While the pyramidal CMOS image sensor was mainly a system level approach to design a "more-human-like" imager to maximize the information over data ratio, the multiresolution imager was primarily a device level approach to this goal, as discussed in Chapter 7. This architecture, in contrast to previous attempts, was based on implementing the multiresolution mechanism at the pixel level. This is to ensure programmability and expandability of the multiresolution functionality of the imager for high resolutions. Experimental data analysis of this architecture would have brought more useful information regarding the charge sharing mechanism used to decrease resolution on programmed areas of less-interest. It would also help to get a better picture about how fast (frame rates) this architecture can achieve including the dependence of frame rates on averaging kernels and its impact on image

quality. Unfortunately, the chip did not pass the testing phase for unknown reasons and thus it remains a future work to be carried on.

8.1 Future Work and Perspectives

More work need to be done to benefit from the architecture flexibility of CMOS imaging technology either to minimize the drawbacks of using this imaging technology (compared to the CCD imaging technology) especially the relatively higher FPN noise, or exploit its integration capabilities. Biological vision systems can give a great help in synthesising novel architectures exhibiting more adaptability, to light intensity for higher ranges and spatial resolving power. The Pyramidal CMOS imager and the multiresolution CMOS imager are first steps towards this vision.

Bibliography

[1] W. Boyle and G. Smith, “Charge coupled semiconductor devices,” Bell Systems Tech. J., vol. 49, pp. 587–593, Apr. 1970.

[2] E.R. Fossum, "Active Pixel Sensors: Are CCD's Dinosaurs?", SPIE vol. 1900, pp. 1-14,1993

[3] R. Guntupalli and J. Grant, “CCD Cameras Tune in to Scientific Imaging “, Photonics Spectra, April 2004, pp. 62-64

[4] B. A. Wandell, “Foundations of Vision”, Sinauer Associates, Inc. Publishers. Sunderland, Massachusetts, 1995

[5] G. Weckler, “Operation of p-n junction photodetectors in a photon flux integration mode,” IEEE J. Solid-State Circuits, vol. SC-2, 1967, p. 65

[6] E. Fossum, “CMOS image sensors: electronics camera on a chip,” IEDM, vol. 44, no. 10, pp. 17-25, 1995.

[7] P. Noble, Self-scanned image detector arrays, IEEE Trans. Electron Devices, vol. ED-15, p. 202, 1968.

[8] F. Andoh, K. Taketoshi, J. Yamazaki, et al., "A 250,000 pixel image sensor with FET amplification at each pixel for high-speed television camera", 1990 Int. Solid State Circuits Conference Dig. of Tech. Papers, 1990, pp. 212-213.

[9] Eric Fossum, "CMOS active pixel image sensors," Nucl. Instr. and Meth. in Phys. Res., A 395 (1997) 291-297.

[10] A. El Gamal et al. "How Small Should Pixel Size Be?" Proceedings of the SPIE Electronic Imaging '2000 conference, Vol. 3965, San Jose, CA, January 2000

[11] S. Mendis, S. Kemeny, E. R. Fossum, "A 128x128 CMOS active pixel sensor for highly integrated systems," 1993 IEEE internat. Electron Devices Meeting Tech. Dig., pp. 583-586.

[12] R. H. Nixon, S. E. Kemeny, B. Pain, C. O. Staller, E. R. Fossum, IEEE J. Solid State Circuits 31(12), pp. 2046, 1996.

[13] G. P.Weckler, “Operation of p-n junction photodetectors in a photon flux integration mode”, IEEE J. Solid-State Circuits, vol. SC-2, pp. 65–73, 1967

[14]M. White et al., “Characterization of surface channel CCD image arrays at low light levels”, IEEE Journal of Solid-State Circuits, Vol. SC-9, No. 1, 1977

[15]S. K. Mendis, S. E. Kemeny, R. C. Gee, B. Pain, C. O. Staller, Q. Kim, and E. R. Fossum, “CMOS active pixel image sensors for highly integrated imaging systems,” IEEE J. Solid-State Circuits, vol. 32, pp. 187–197, Feb. 1997.

[16]James R. Janesick, Scientific Charge-Coupled Devices, SPIE Press, 2001

[17]Sandor L. Barna, “Sensors Support Low-Light Imaging”, Photonics Spectra, November 2003, pp. 100-101

[18]http://dictionary.reference.com/search?r=67&q=scan

[19]http://inventors.about.com/library/inventors/blcathoderaytube.htm

[20]G. P.Weckler, “Operation of p-n junction photodetectors in a photon flux integration mode”, IEEE J. Solid-State Circuits, vol. SC-2, pp. 65–73, 1967

[21]Hanan Samet, “The Design and Analysis of Spatial Data Structures”, Addison-Wesley, 1990.

[22]Artyomov, E.; Rivenson, Y.; Levi, G.; Yadid-Pecht, O, “Morton (Z) scan based real-time variable resolution CMOS image sensor”, Proceedings of ICECS 2004, pp. 145-148, December 2004.

[23]A. Ansari and A. Fineberg. “Image data compression and ordering using Peano scan and lot”. IEEE Transactions on Consumer Electronics, 38(3):436–445,August 1992.

[24]Nasir Memon, “An analysis of some common scanning techniques for lossless image coding”, IEEE transactions on image processing, Vol. 9, No. 11, pp. 1837-1848, November 2000.

[25]P.T. Nguyen and J. Quinqueton. “Space filling curves and texture analysis”. IEEE Intl. Conf. Pattern Recognition, pages 282–285, October 1982.

[26]Dave Litwiller, “How an Imaging Sensor Feeds off Light”, Photonics Spectra, pp. 96 May 2004

[27]Peter B. Catrysse, Xinqiao Liu, and Abbas El Gamal: QE Reduction due to Pixel Vignetting in CMOS Image Sensors”, Proceedings of SPIE, vol. 3965 (2000).

[28]Yadid-Pecht, B. Pain, C. Staller, C. Clark, E. Fossum , “CMOS active pixel sensor star tracker with regional electronic shutter”, IEEE J. Solid State Circuits, Vol. 32, No. 2, pp. 285-288, February 1997.

[29]O. Yadid-Pecht, R. Ginosar and Y. Shacham-Diamand, "A random access photodiode array for intelligent image capture", IEEE Trans. on Electron Devices, Vol. 38, No. 8, pp. 1772 - 1781, Aug.1991.

[30]T. Nezuka et al. "A smart image sensor with novel implementation of quad-tree scan", Proceedings of 2nd IEEE Asia-Pacific conference on ASICs, pp135-138, October 2000.

[31]http://www.answers.com/main/ntquery;jsessionid=eqcft1m1t6sol?method=4&dsid=2222&dekey=Tree+data+structure&gwp=11&curtab=2222_1&sbid=lc01b

[32]http://www.answers.com/main/ntquery?method=4&dsid=2222&dekey=Tree+%28graph+theory%29&gwp=11&curtab=2222_1

[33]Alireza Moini, Vision Chips, Kluwer Academic Publishers, 1999

[34]Fayçal Saffih, Richard Hornsey, "Biomimetic Sampling Architectures for CMOS Image Sensors" IS&T/SPIE Symposium on Electronic Imaging 2004, 18-22 January 2004, San Jose, CA USA

[35]Ferwerda James, "Fundamentals of spatial vision", Applications of visual perception in computer graphics, Program of Computer Graphics, Cornel University, 1998.

[36]Hecht Eugene, Optics, 2nd Ed, Addison Wesley, 1987

[37]Hon-Sum Wong, "Technology and Device Scaling, Consideration for CMOS Imagers", IEEE transactions on Electron Devices, Vol. 43, No. 12, December 1996.

[38]A. El Gamal et al. "How Small Should Pixel Size Be?" Proceedings of the SPIE Electronic Imaging '2000 conference, Vol. 3965, San Jose, CA, January 2000

[39]S. Xia, R. Sridhar, P. Scott and C. Bandera, "An All CMOS Foveal Image Sensor Chip", 11th IEEE International ASIC Conference, Rochester, NY, September 1998

[40]E. R. Fossum et al. "Multiresolution Image Sensor", IEEE transactions on circuits and systems for video technology. Vol. 7, No. 4, August 1997

[41]J. Coulomb, M. Sawan and C. Wang. "Variable Resolution CMOS Current Mode Active Pixel Sensor", ISCAS 2000 - IEEE International Symposium on Circuits and Systems, May 28-31, 2000, Geneva, Switzerland. pp. 293-296

[42]B. Dierickx, D. Scheffer, G. Meynants, W. Ogiers, and J. Vlummens, ‘‘Random addressable active pixel image sensors,’’ Proc. SPIE 2950, pp.2–7 (1996)

[43]http://www.fillfactory.com/htm/technology/htm/Logresfaq.htm

[44]Chamberlain S. G., Lee J. P., “A Novel Wide dynamic range silicon photodetector and linear imaging array”, IEEE journal of Solid-State Circuits, Vol. SC-19, No 1, February 1984.

[45]http://micro.magnet.fsu.edu/primer/lightandcolor/humanvisionintro.html

[46]O. Yadid-Pecht, “Wide dynamic range sensors”, Opt. Eng., vol. 38, no.10, pp. 1650–1660, Oct. 1999.

[47]Anton, P., Granger, R. & Lynch, G. (1992). Temporal information processing in synapses, cells, and circuits. In T. McKenna, J. Davis & S. F. Zornetzer (Eds.), Single Neuron Computation (pp. 291–313). Boston: Academic Press.

[48]Yang W., "A wide-dynamic-range, low power photosensor array", IEEE ISSCC Vol. 37, 1994.

[49]http://www.cs.yorku.ca/~visor/pdf/Xiuling_Thesis.pdf

[50]D. E. Caudle, ‘‘New generation gated intensified camera,’’ Proceedings of SPIE 569, pp. 107–113, 1985

[51]T. Kinugasa et al., ‘‘An electronic variable-shutter system in video camera use,’’ IEEE Trans. Consum. Electron. CE-33, pp. 249–258, Aug. 1987

[52]Lauxtermann, S.; Schwider, P.; Seitz, P.; Bloss, H.; Ernst, J.; Firla, H., “A high speed CMOS imager acquiring 5000 frames/sec”, Electron Devices Meeting, 1999. IEDM Technical Digest. International 5-8 Dec. 1999 Page(s):875 – 878

[53]Martin Wäny and Georg P. Israel, "CMOS Image Sensor with NMOS-Only Global Shutter and Enhanced Responsivity", IEEE Transactions on electron devices, Vol. 50, NO. 1, pp. 57-62, January 2003

[54]R. Miyagawa and T. Kanade, ‘‘Integration time based computational image sensor,’’ presented at 1995 IEEE Workshop on CCD’s and Advanced Image Sensors, April 20–22, 1995, Dana Point, CA.

[55]V. Brajovic, R. Miyagawa, and T. Kanade, "Temporal Photoreception for Adaptive Dynamic Range Image Sensing and Encoding", Neural Networks, Vol. 11, No. 7-8, October, 1998, pp. 1149-1158

[56]O. Yadid-Pecht, ‘‘Widening the dynamic range of pictures,’’ Proceedings of SPIE 1656, pp. 374–382, 1992

[57]http://www.answers.com/main/ntquery?method=4&dsid=2222&dekey=Exposure+%28photography%29&gwp=12&curtab=2222_1

[58]http://www.answers.com/main/ntquery?method=4&dsid=2222&dekey=Reciprocity+%28photography%29&gwp=12&curtab=2222_1

[59]S. K. Nayar and T. Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 2000.

[60]A. J. P. Theuwissen. "Solid State Imaging with Charge-Coupled Devices", Kluwer Academic Press, Boston, 1995.

[61]Gonzalez, Rafael C., "Digital image processing", Prentice Hall, 2002.

[62]Umbaugh, Scott E, "Computer imaging: digital image analysis and processing", CRC Press, 2005.

[63]Brian C. Madden, "Extended Intensity Range Imaging", Technical Report MS-CIS-93-96, Grasp Laboratory, University of Pennsylvania, 1993.

[64]http://www.schorsch.com/kbase/glossary/contrast.html

[65]http://www.answers.com/main/ntquery?method=4&dsid=2222&dekey=Cathode+ray+tube&gwp=12&curtab=2222_1

[66]http://www.answers.com/gamma+correction?gwp=12&method=2

[67]S. K. Mendis, S. E. Kemeny, R. C. Gee, B. Pain, Q. Kim, and E. R. Fossum, "CMOS active pixel image sensors for highly integrated imaging systems," IEEE J. Solid-State Circuits, vol. 32, pp. 187–197, Feb. 1997.

[68]O. Yadid-Pecht, E. R. Fossum, "Wide Intrascene Dynamic Range CMOS APS Using Dual Sampling," IEEE Trans. Electron Devices, vol. 44, no 10, pp. 1721-1723, October 1997.

[69]S. Kleinfelder, S.H. Lim, X.Q. Liu and A. El Gamal, \A 10,000 Frames/s 0.18 um CMOS Digital Pixel Sensor with Pixel-Level Memory," In Proceedings of the 2001 IEEE International Solid-State Circuits Conference, pp. 88-89, San Francisco, CA, February 2001.

[70]http://www.foveon.com/article.php?a=67

[71]Perry, J.S. and Geisler, W.S. (2002) Gaze-contingent real-time simulation of arbitrary visual fields. In B. Rogowitz and T. Pappas (Eds.), Human Vision and Electronic Imaging, SPIE Proceedings, San Jose, CA.

[72]Geisler, W.S. and Perry, J.S. (2002) Real-time simulation of arbitrary visual fields. ACM Symposium on Eye Tracking Research & Applications 2002.

[73]Geisler, W.S. and Perry, J.S. (1999) Variable-resolution displays for visual communication and simulation. The Society for Information Display, 30, 420-423.

[74]Geisler, W.S. and Perry, J.S. (1998) A real-time foveated multi-resolution system for low-bandwidth video communication. In B. Rogowitz and T. Pappas (Eds.), Human Vision and Electronic Imaging, SPIE Proceedings, 3299, 294-305.

[75]Neil H. E. West, Kamran Eshraghian, Principles of CMOS VLSI Design, A System Perspective, 2nd ed, 1992.

[76]http://www.semiconductorglossary.com/default.asp?searchterm=SALICIDE+process

[77]Hsiu-Yu Cheng; Ya-Chin King, "A CMOS image sensor with dark-current cancellation and dynamic sensitivity operations", IEEE Transactions on Electron Devices, Vol. 50, issue 1, Jan. 2003.

[78]John P. Huber, Successful ASIC design the first time through, New York : Van Nostrand Reinhold, 1991.

[79]Charles Kittel, Introduction to solid state physics, New York: Wiley, 1996

[80]H. Tian, X.Q. Liu, S.H. Lim, S. Kleinfelder and A. El Gamal, Active Pixel Sensors Fabricated in a Standard 0.18 um CMOS Technology, Proceedings of the SPIE, Vol. 4306, p. 441-449, January 2001

[81]D. Buchanan and S. H. Lo, "Growth,characterization and the limits of ultrathin SiO2-based dielectrics for future CMOS applications," The physics and chemistry of SiO2 and the Si-SiO2 interface –3, Electrochemical Society Meeting Proceedings 90(1), pp. 3–14, 1996.

[82]Johns David, Ken Martin, Analog integrated circuit design, New York: Wiley, 1997

[83]Jan M. Rabaey, Digital integrated circuits : a design perspective, Prentice Hall, 1996

[84]Martin Wäny, Georg P. Israel, "CMOS Image Sensor With NMOS-Only Global Shutter and Enhanced Responsivity", IEEE Transactions on electron devices, Vol. 50, No. 1, pp. 57-62, January 2003.

[85]A. V. Oppenheim and R. W. Schafer, Digital Signal Processing, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975.

[86]http://www.kodak.com/global/plugins/acrobat/en/digital/ccd/applicationNotes/ShutterOperations.pdf

[87]S. Lauxtermann and G. P. Israel, "Cellule Active Avec Mémoire Analogique Pour un Capteur Photosensible Réalisé en Technologie CMOS,", Demande de Brevet en France No. 0 004 494.

[88]S. Lauxtermann et al., "A mega-pixel high speed CMOS imager with sustainable gigapixel/sec readout rate," in Proc. 2001 IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors, Lake Tahoe, NV, June 7–9, 2001, pp. 48–51.

[89]S. E. Kemeny, R. Panicacci, B. Pain, L. Matthies and E. R. Fossum, "Multiresolution Image Sensor", IEEE Transactions on circuits and systems for video technology, vol. 7, no. 4, august 1997, pp. 575.

[90]Z. Zhou, B. Pain, and E. R. Fossum, Frame-Transfer CMOS Active Pixel Sensor with Pixel Binning, IEEE Transactions on electron devices, VOL. 44, NO. 10, OCTOBER 1997, pp. 1764.

[91]Stefan L. et al., "A high speed CMOS imager acquiring 5000 frames/sec", Electron Devices Meeting, 1999. IEDM Technical Digest, pp. 875-878, 5-8 Dec. 1999

[92]M. Hillebrand et al., "High Speed Camera System using a CMOS Image Sensor", Proceedings of the IEEE Intelligent Vehicles Symposium 2000, pp. 656-661, Dearborn (MI), USA. October 3-5, 2000

[93]Fayçal Saffih, Richard Hornsey, "Pyramidal Architecture for CMOS Image Sensors", Proceedings of IEEE International Workshop on CCD and Advanced Image Sensors, May 15-17, 2003, Schloss Elmau, Germany

[94]Fayçal Saffih, Richard Hornsey, "Human Perception of Fixed Pattern Noise in Pyramidal CMOS Image Sensor", SPIE-Photonics North 2004, Ottawa, Canada, 27 - 29 September 2004.

[95]A. El Gamal, B. Fowler, H. Min, and X. Liu, "Modeling and Estimation of FPN Components in CMOS Image Sensors", Proceedings of SPIE, Vol. 3301, pp. 168-177, April 1998.

[96]A. Watanabe, T. Mori, S. Nagata, and K. Hiwatashi, "Spatial sine-wave responses of the human visual system," Vision Res., vol. 8, pp. 1245–1263, 1968.

[97]Graham, Norma Van Surdam, "Visual pattern analyzers", pp. 506 New York, Oxford University Press, 1989

[98]Hugh R. Wilson, Spatial Vision, chapter 3: “Psychophysical Models of Spatial Vision and Hyperacuity”, edited by D. Regan, The Macmillan Press, 1991.

[99]Matthew J. McMahon, Donald I. A. MacLeod, “The origin of the oblique effect examined with pattern adaptation and masking”, Journal of Vision (2003), vol. 3, 230-239.

[100] Mach, E. (1861). Über das Sehen von Lagen und Winkeln durch die Bewegung des Auges. Sitzungsberichte der Kaiserlichen Akademie derWissenschaften, 43(2), 215-224.

[101] Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the "oblique effect" in man and animals. Psychological Bulletin, 78(4), 266-278.

[102] David et al. “The distribution of oriented contours in the real world”, Proceedings of National Academy of Science. USA, Vol. 95, pp. 4002-4006, March 1998.

[103] T. Yamada et al. “A Progressive Scan CCD Image Sensor for DSC Applications”, IEEE Journal of Solid-State Circuits, Vol. 35, No 12, December 2000.

[104] Baylor, D. A. 1987. Photoreceptor signals and vision. Inv. Opthlamol. Vis. Sci., 28:34-49

[105] Ronald N. Bracewell, “Two-dimensional Imaging”, Prentice Hall signal processing series, 1995

[106] Fayçal Saffih, Richard Hornsey, “HVS low perception of FPN Noise in Pyramidal CMOS imagers”, Proceedings of IEEE International Workshop on CCD and Advanced Image Sensors, June 9-11, 2005, Nagano, Japan

[107] S. E. Kemeny, R. Panicacci, B. Pain, L. Matthies and E. R. Fossum, Multiresolution Image Sensor, IEEE Transactions on circuits and systems for video technology, vol. 7, no. 4, august 1997, pp. 575.

[108] Z. Zhou, B. Pain, and E. R. Fossum, Frame-Transfer CMOS Active Pixel Sensor with Pixel Binning, IEEE Transactions on electron devices, VOL. 44, NO. 10, OCTOBER 1997, pp. 1764.

[109] J. Coulomb, M. Swan and C. Wang, Variable resolution CMOS current mode active pixel sensor, ISCAS 2000 – IEEE International Symposium on Circuits and Systems, May 28-31, 2000, Geneva, Switzerland, pp. II-293

[110] Fayçal Saffih, Richard Hornsey, “Multiresolution CMOS image sensor”, Technical Digest of SPIE Opto-Canada 2002, Ottawa, Ontario, Canada 9 – 10, May 2002, p.425.