

**Propuesta de Arquitectura y Circuitos para la Mejora
del Rango Dinámico de Sistemas de Visión en un Chip
diseñados en Tecnologías CMOS profundamente
Submicrométricas**

**Proposal of Architecture and Circuits for Dynamic
Range Enhancement of Vision Systems on Chip
designed in Deep Submicron Technologies**



Sonia María Vargas Sierra
Directors: PhD. Gustavo Liñán Cembrano, PhD. Elisenda Roca Moreno

Departamento de Electrónica y Electromagnetismo
Universidad de Sevilla

Instituto de Microelectrónica de Sevilla (IMSE-CNM)
Consejo Superior de Investigaciones Científicas (CSIC)

A thesis submitted for the degree of

Doctor of Philosophy

2012

To David and my family

Acknowledgements

First, I would like to thank both my thesis directors, PhD. Gustavo Liñán Cembrano and PhD. Elisenda Roca Moreno, for the help, encourage and comprehension during difficulties.

I want to express my love and gratitude to my boyfriend David for believing in me much more than I will ever do, and his understanding of how important this work has been for me.

I am so grateful to my beloved family, whose unconditional support, understanding and love have always help me to seek my dreams.

I want to show my appreciation to Prof. Bedrich J. Hosticka and his group of the Fraunhofer Institute for Microelectronic Circuits and Systems (IMS), for accepting me to stay with them during three months, and for their nice treatment during that term.

I am thankful for all the words of encouragement and emotional support from colleagues and friends that I have received. Especially, I would like to thank all the colleagues, that during lunches and breakfasts, have accompanied me during all these years.

Summarizing, I would like to thank all of those who contributed to the success of this work in one way or another.

Resumen

Una cámara es esencialmente un dispositivo para capturar escenas. La información sobre estas escenas se mide mediante la captura de luz, tanto emitida como reflejada por los objetos. Tradicionalmente, la fotografía ha intentado capturar representaciones de las escenas tal y como lo haría el ojo humano. Esto significa capturar los fotones en la banda del llamado espectro visible (típicamente comprendido entre longitudes de onda de 380 a 780 nanómetros). El método empleado para medir la luz ha evolucionado desde los procesos químicos usados al comienzo de la fotografía, hasta los procesos puramente electrónicos utilizados hoy en día, gracias a la aparición de los sensores electrónicos de imagen. Desde su invención en la década de los 60 en el siglo pasado, y especialmente en la primera década de éste, las prestaciones de los sensores electrónicos de imagen han evolucionado de una forma vertiginosa. Esto ha sido motivado, especialmente en los últimos años, por un mercado que crece de manera exponencial. No obstante, las cámaras de hoy en día aún sufren de algunas limitaciones. Una cámara ideal sería aquella en que las imágenes captadas tengan muy poco ruido, un rango dinámico excepcional (o al menos el máximo rango dinámico encontrado en escenas reales), gran número de píxeles, capturadas con un sensor de alta sensibilidad, en un sistema de muy bajo consumo de potencia y todo ello mientras se mantiene un coste de fabricación bajo. Aunque en los últimos años se han realizado grandes esfuerzos hacia la mejora de muchas características, como por ejemplo, el incremento del número de píxeles (sensores de muchos megapíxeles [1]), la reducción del coste, la minimización del consumo de potencia y disminución del ruido (con ruidos de lectura por debajo de un electrón rms [2] en sensores de última generación), otras, como es el caso del rango dinámico, han experimentado un desarrollo más lento. La tendencia evolutiva dictada por los planes de marketing de la mayoría de compañías es generalmente incrementar el número de píxeles mientras se mantiene el tamaño físico del chip. Inevitablemente, esto lleva a una reducción del tamaño del píxel (en muchos casos por debajo del límite de difracción, llamado Disco de Airy). En consecuencia, esto provoca una reducción casi continua del máximo número de cargas fotogeneradas que pueden ser detectadas y almacenadas a nivel de píxel (reducción de la llamada Full Well Capacity), lo que produce una reducción, en la mayoría de los casos, del rango dinámico alcanzable.

En la actualidad, existen dos alternativas tecnológicas en el diseño de sensores electrónicos de imagen para sensar luz en el espectro visible, los denominados Dispositivos de Carga Acoplada (CCD) y los circuitos Complementary Metal Oxide Semiconductor (CMOS), los

cuales han coexistido durante aproximadamente 50 años. En un principio, los sensores de imagen CCD fueron la opción principal para la mayoría de las aplicaciones de imagen debido a sus superiores características de ruido, sensibilidad y también eficiencia en la transferencia de carga. Sin embargo, los sensores de imagen CMOS han experimentado un gran impulso desde finales del siglo pasado. Además, se han desarrollado tecnologías CMOS especialmente dedicadas para la detección de luz visible, lo que hace que tengan una calidad de imagen comparable a la de los dispositivos CCD. Estas tecnologías, usualmente conocidas como tecnologías CMOS de sensores de imagen o tecnologías CIS, han aumentado el número de potenciales aplicaciones para los sensores de imagen CMOS. Hoy en día, los sensores de imagen CMOS han copado la mayoría del mercado, básicamente debido a la posibilidad de añadir etapas de acondicionamiento de señal (o incluso procesamiento de señal) en el mismo chip, permitiendo dimensiones muy reducidas para ser insertadas, por ejemplo, en un teléfono móvil, así como debido al reducido consumo de potencia en comparación con los dispositivos CCD.

El trabajo presentado en esta tesis trata de proponer nuevas técnicas para la expansión del rango dinámico en sensores electrónicos de imagen. En este caso, hemos dirigido nuestros estudios hacia la posibilidad de proveer dicha funcionalidad en un solo chip. Esto es, sin necesitar ningún soporte externo de hardware o software, formando un tipo de sistema denominado Sistema de Visión en un Chip (VSoC). El rango dinámico de los sensores electrónicos de imagen se define como el cociente entre la máxima y la mínima iluminación medible. Para mejorar este factor surgen dos opciones. La primera, reducir la mínima luz medible mediante la disminución del ruido en el sensor de imagen. La segunda, incrementar la máxima luz medible mediante la extensión del límite de saturación del sensor.

Cronológicamente, nuestra primera opción para mejorar el rango dinámico se basó en reducir el ruido. Varias opciones se pueden tomar para mejorar la figura de mérito de ruido del sistema: reducir el ruido usando una tecnología CIS o usar circuitos dedicados, tales como calibración o auto cero. Sin embargo, el uso de técnicas de circuitos implica limitaciones, las cuales sólo pueden ser resueltas mediante el uso de tecnologías no estándar que están especialmente diseñadas para este propósito. La tecnología CIS utilizada está dirigida a la mejora de la calidad y las posibilidades del proceso de fotosensado, tales como sensibilidad, ruido, permitir imagen a color, etcétera. Para estudiar las características de la tecnología en más detalle, se diseñó un chip de test, lo cual permite extraer las mejores opciones para futuros píxeles. No obstante, a pesar de un satisfactorio comportamiento general, las medidas referentes al rango dinámico indicaron que la mejora de éste mediante sólo tecnología CIS es muy limitada. Es decir, la mejora de la corriente oscura del sensor no es suficiente para nuestro propósito. Para una mayor mejora del rango dinámico se deben incluir circuitos dentro del píxel. No obstante, las tecnologías CIS usualmente no permiten nada más que transistores NMOS al lado del fotosensor, lo cual implica una seria restricción en el circuito a usar. Como resultado, el diseño de un sensor de imagen con mejora

del rango dinámico en tecnologías CIS fue desestimado en favor del uso de una tecnología estándar, la cual da más flexibilidad al diseño del píxel.

En tecnologías estándar, es posible introducir una alta funcionalidad usando circuitos dentro del píxel, lo cual permite técnicas avanzadas para extender el límite de saturación de los sensores de imagen. Para este objetivo surgen dos opciones: adquisición lineal o compresiva. Si se realiza una adquisición lineal, se generarán una gran cantidad de datos por cada píxel. Como ejemplo, si el rango dinámico de la escena es de 120dB al menos se necesitarían 20-bits/píxel, $\log_2(10^{120/20})=19.93$, para la representación binaria de este rango dinámico. Esto necesitaría de amplios recursos para procesar esta gran cantidad de datos, y un gran ancho de banda para moverlos al circuito de procesamiento. Para evitar estos problemas, los sensores de imagen de alto rango dinámico usualmente optan por utilizar una adquisición compresiva de la luz. Por lo tanto, esto implica dos tareas a realizar: la captura y la compresión de la imagen. La captura de la imagen se realiza a nivel de píxel, en el dispositivo fotosensor, mientras que la compresión de la imagen puede ser realizada a nivel de píxel, de sistema, o mediante postprocesado externo. Usando el postprocesado, existe un campo de investigación que estudia la compresión de escenas de alto rango dinámico mientras se mantienen los detalles, produciendo un resultado apropiado para la percepción humana en monitores convencionales de bajo rango dinámico. Esto se denomina Mapeo de Tonos (Tone Mapping) y usualmente emplea solo 8-bits/píxel para la representaciones de imágenes, ya que éste es el estándar para las imágenes de bajo rango dinámico.

Los píxeles de adquisición compresiva, por su parte, realizan una compresión que no es dependiente de la escena de alto rango dinámico a capturar, lo cual implica una baja compresión o pérdida de detalles y contraste. Para evitar estas desventajas, en este trabajo, se presenta un píxel de adquisición compresiva que aplica una técnica de mapeo de tonos que permite la captura de imágenes ya comprimidas de una forma optimizada para mantener los detalles y el contraste, produciendo una cantidad muy reducida de datos. Las técnicas de mapeo de tonos ejecutan normalmente postprocesamiento mediante software en un ordenador sobre imágenes capturadas sin compresión, las cuales contienen una gran cantidad de datos. Estas técnicas han pertenecido tradicionalmente al campo de los gráficos por ordenador debido a la gran cantidad de esfuerzo computacional que requieren. Sin embargo, hemos desarrollado un nuevo algoritmo de mapeo de tonos especialmente adaptado para aprovechar los circuitos dentro del píxel y que requiere un reducido esfuerzo de computación fuera de la matriz de píxeles, lo cual permite el desarrollo de un sistema de visión en un solo chip.

El nuevo algoritmo de mapeo de tonos, el cual es un concepto matemático que puede ser simulado mediante software, se ha implementado también en un chip. Sin embargo, para esta implementación hardware en un chip son necesarias algunas adaptaciones y técnicas avanzadas de diseño, que constituyen en sí mismas otra de las contribuciones de este trabajo. Más aún, debido a la nueva funcionalidad, se han desarrollado modificaciones de los típicos

métodos a usar para la caracterización y captura de imágenes.

Esta memoria de tesis doctoral está organizada en cinco capítulos. El capítulo 1 describe los principales conceptos referentes a los sensores de imagen enfocados al alto rango dinámico. El capítulo 2 presenta el estudio de una tecnología CIS para su consideración para técnicas de mejora del rango dinámico. El capítulo 3 describe el nuevo algoritmo de mapeo de tonos desarrollado para optimizar la compresión de imágenes capturadas desde escenas de alto rango dinámico. El capítulo 4 presenta el chip que se ha diseñado y fabricado, el cual incorpora un sensor de imagen que utiliza el nuevo algoritmo de mapeo de tonos. El capítulo 5 describe los resultados experimentales y el comportamiento del chip. Finalmente, se detallan las conclusiones y el trabajo futuro.

[1] Productos de la compañía Samsung: Sensor de Imagen CMOS S5K2P1, 16-Megapíxel y tecnología con píxel de iluminación posterior de 1.34 μ s.

[2] Publicaciones de la compañía Caeleste: "A 0.5 noise electrons CMOS pixel", por Bart Dierickx, Nayera Ahmed, y Benoit Dupont.

Conclusiones

El objetivo de esta tesis doctoral ha sido desarrollar nuevas arquitecturas, algoritmos, y técnicas de diseño de circuitos, para la mejora del rango dinámico de la parte sensora en Procesadores de Plano Focal. El trabajo presentado en esta tesis constituye un avance en el estado del arte de sensores de alto rango dinámico, proponiéndose una solución innovadora para mejorar el rango dinámico en sensores de imagen CMOS mediante la introducción de una técnica de mapeo de tonos. Esta técnica no sólo incrementa considerablemente el rango dinámico, sino que también optimiza la representación final de la imagen mediante una reducida cantidad de bits (siete) por píxel. Esto se consigue generando información acerca de la probabilidad de distribución de potencia de la luz incidente en el anterior fotograma, y usando esta información para ajustar como se comprime y guarda el fotograma actual.

Las aportaciones realizadas en esta tesis pueden agruparse en tres campos diferentes: (1) exploración de tecnologías CIS dedicadas, (2) diseño de algoritmos para el sensado de alto rango dinámico basado en el mapeo de tonos, y (3) diseño de un sensor QCIF de alto rango dinámico, en una tecnología estándar, el cual sirve como demostrador de lo que puede conseguirse con las técnicas desarrolladas. Las principales conclusiones para cada una de estas líneas de investigación se resumen a continuación:

Respecto al trabajo en tecnologías CIS:

- El uso de tecnologías CIS para la mejora del rango dinámico de los sensores indica avances significativos en términos de reducción de corriente oscura (factor 29) e incremento de sensibilidad (factor 10) cuando se compara con el sensor equivalente en una tecnología CMOS estándar. Además, la posibilidad de añadir microlentes encima de los sensores ofrece una mejora aún mayor de la sensibilidad (con un máximo del 80 % en los píxeles de menor tamaño) y mejora el crosstalk, sin impacto en el tamaño del píxel o en el ruido eléctrico a nivel del sensor.
- Las limitaciones tecnológicas asociadas a las tecnologías CIS excluyen la incorporación de transistores PMOS y contactos de sustrato en el área del sensor, reduciendo fuertemente, por lo tanto, la cantidad de “inteligencia” que puede ser incorporada a nivel de píxel.
- Estas limitaciones producen dos importantes inconvenientes para esta alternativa. Primero, esto casi descarta¹ este tipo de tecnologías cuando se diseñan Procesadores

¹Se puede utilizar procesamiento basado únicamente en transistores NMOS, aunque su posible funcionalidad es muy limitada.

de Plano focal (que por definición, deben incorporar algunas estructuras de procesamiento a nivel de píxel), y segundo, esto básicamente limita las opciones de mejora del rango dinámico a aquellas que disminuyen el nivel de ruido, y en el mejor de los casos, al uso de técnicas de muestreo múltiple (donde las imágenes finales son una combinación, realizada fuera del chip, de imágenes tomadas con distintos tiempos de exposición). Esto, igualmente, no es compatible con el enfoque hacia Procesadores de Plano Focal, debido a que la información del píxel es el punto de partida del algoritmo de procesamiento de imagen y una parte significativa de éste está implementado con circuitos dentro del píxel.

Respecto al nuevo algoritmo de Mapeo de Tonos:

- Se ha desarrollado un nuevo algoritmo de Mapeo de Tonos orientado al hardware para la mejora del rango dinámico, el cual combina los dos paradigmas típicos para la adquisición de corrientes fotogeneradas. Durante la mayoría del tiempo de exposición codificamos la información visual como el tiempo que tarda el nivel de voltaje en cruzar una tensión de referencia fija (en un píxel de integración de fotocorriente). Al final del tiempo de exposición, medimos el voltaje del píxel para esos píxeles que no han cruzado previamente el valor de referencia. Comúnmente, sólo uno de estos paradigmas es usado, sin embargo, la combinación de ambos es una innovación dentro de este algoritmo.
- El algoritmo genera una curva compresiva de mapeo de tonos desde el histograma de una imagen auxiliar, que se emplea como descriptor global de la distribución de iluminaciones en el actual fotograma, y requiere muy poca computación a nivel de píxel.
- Aunque originalmente no se ha pretendido optimizar la visualización de escenas de alto rango dinámico en monitores de bajo rango dinámico, el algoritmo puede ser usado igualmente para este propósito, debido a que la compresión implementada es mayor a altas iluminaciones, lo cual es consistente con el sistema de visión humano. Las características de la compresión desarrollada, global y monotónica, producen representaciones visualmente apropiadas mientras evita la creación de artefactos visuales.
- Además, el algoritmo implementa una pseudo-ecualización de los detalles de la escena en la imagen final, ya que se evitan espacios vacíos en el uso de códigos de salida digitales cuando se configura apropiadamente.
- El algoritmo es completamente compatible con los típicos recursos computacionales localizados a nivel de píxel en Procesadores de Plano Focal, contribuyendo al estado del arte de técnicas de Mapeo de Tonos, las cuales son raramente orientadas al hardware.

Con respecto al chip demostrador, el cual implementa el algoritmo de mapeo de tonos diseñado, hasta donde conoce la autora, no ha sido reportado ningún otro diseño de Sistema

de Visión en un Chip que implemente un mapeo de tonos simultáneo a la captura dentro del píxel. Adicionalmente, las técnicas de diseño usadas para hacer posible la implementación en un chip del algoritmo constituyen aportaciones al campo de los sensores de alto rango dinámico. Las conclusiones en este ámbito son:

- Se ha diseñado un demostrador en una tecnología estándar de fabricación con sólo dos modificaciones para mejorar las capacidades de detección de la luz: un revestimiento anti-reflexión y un sustrato EPI optimizado. Esto reduce el coste del chip en comparación con el uso de una tecnología CIS, y no introduce ninguna limitación al tipo de circuitos que pueden ser situados cerca del sensor.
- El chip ha probado que el algoritmo funciona apropiadamente. También confirma que el método matemático, usado para su simulación en escenas de alto rango dinámico, es apropiado para predecir la funcionalidad del chip.
- Los píxeles incluyen una técnica de auto cero para minimizar el Ruido de Patrón Fijo. El esquema no requiere ningún condensador de almacenamiento adicional o múltiples lecturas, ya que se introduce como el voltaje de reset efectivo del fotodiodo.
- Hemos diseñado un esquema dinámico para la distribución de la referencia analógica, el cual aprovecha la rápida distribución de carga en señales cuya media es cero en cargas RC ampliamente distribuidas. Esto permite una comunicación precisa de la señales analógicas a un amplio número de nodos de alta impedancia.
- El chip contiene una matriz heterogénea de píxeles, cuya unidad básica consiste en una disposición de 2×2 , la cual permite una fácil implementación de la funcionalidad submuestreada.
- El almacenamiento dentro del píxel de la imagen final en una memoria RAM estática permite tomar fotogramas con exposiciones a la luz muy largas, sin distorsión de los datos causados por fugas o los circuitos circundantes.
- El chip puede ser usado como un sensor normal de adquisición lineal; en tal caso, produce señales con SNR de 53.8dB.
- El uso del algoritmo proporciona más de 126dB, 168dB - 42dB (Mapeo de Tonos de 7-bits - Lineal de 7-bits), de incremento del rango dinámico en comparación con el modo de adquisición lineal.
- Escenas de hasta 168dB de rango dinámico pueden ser mapeadas por el chip usando sólo 7-bits/píxel.
- La comparación con otros tres sistemas comerciales demuestra que el chip diseñado los supera en calidad visual, ya que, aunque muestra más ruido que un sensor CCD comercial, su operación de alto rango dinámico provee detalles no conservados por este sistema comercial. Cuando se compara con un sensor de imagen específico para alto rango dinámico, aunque este sensor de imagen proporciona un rango dinámico

ligeramente mayor, nuestra solución ofrece mucho mejor contraste y mucho menores niveles de ruido.

- El chip es igualmente muy adecuado para futuras evoluciones del sistema usando tecnología de integración vertical (3D-IC).

Preface

In photography, the method employed for light sensing has evolved from chemical processes at the origins to the purely electronic processes employed today thanks to the appearance of the electronic image sensors. Since their invention in 1960s, and especially during the last decade, the performances of these devices have evolved in an amazing way, driven in the last few years, by an exponentially growing market. However, today's cameras still suffer from some old limitations.

An ideal image sensor should exhibit very low noise, extreme dynamic range (or at least the maximum dynamic range found in real scenes), high responsivity, very high number of pixels, very low power consumption, while keeping low manufacturing cost. Indeed, great efforts have been focused towards the increment of the number of pixels (many megapixels sensors [1]), the reduction of the cost, the minimization of power consumption and noise diminution (read noise below the one electron rms [2]). Other features, such as the sensor dynamic range, have experienced a lower development. Furthermore, the evolution trend dictated by most companies' marketing plans is generally to increase the number of pixels while basically keeping the actual size of the chip. Unavoidably, this leads to pixel size reduction (in many cases below the diffraction limit, named Airy Disk). Consequently, it provokes an almost continuous shrinking of the maximum number of photogenerated charges that can be sensed and stored at the pixel level (reduction of the so-called Full Well Capacitance). This produces a reduction, in most of cases, of the attainable Dynamic Range.

The work presented in this thesis proposes new techniques for dynamic range expansion in electronic image sensors. Since Dynamic Range (DR) is defined as the ratio between the maximum and the minimum measurable illuminations, the options for improvement seem obvious; first, to reduce the minimum measurable signal by diminishing the noise floor of the sensor, and second, to increase the maximum measurable light by increasing the sensor saturation limit.

In our case, we focus our studies to the possibility of providing DR enhancement functionality in a single chip, without requiring any external software/hardware support, composing what is called a Vision-System-on-Chip (VSoC). In order to do so, this thesis covers two approaches. Chronologically, our first option to improve the DR relied on reducing the noise by using a fabrication technology that is specially devoted to image sensor fabrication, a so-called CMOS Image Sensor (CIS) technology. However, measurements from a

test chip indicated that the dynamic range improvement was not sufficient to our purposes (beyond the 100dB limit). Additionally, the technology had some important limitations on what kind of circuitry can be placed next to the photosensor in order to improve its performance. Our second approach has consisted in, first, designing a Tone Mapping algorithm for DR expansion whose computational needs can be easily mapped onto simple signal-conditioning and processing circuitry around the photosensor, and second, designing a test chip implementing this algorithm in a standard CMOS technology.

This thesis is organized in five chapters. Chapter 1 describes the main concepts involved in image sensors focusing in High Dynamic Range (HDR) operation. Chapter 2 presents the study of an image sensor optimized technology in order to be considered for dynamic range improvement techniques. Chapter 3 describes an innovative tone mapping algorithm used to optimize the compression of HDR scenes. Chapter 4 introduces the image sensor chip that has been designed and fabricated, which implements the new tone mapping algorithm. Chapter 5 shows the experimental results and evaluation of the performance of the chip. Finally, conclusions and future work are drawn out.

[1] Samsung enterprise products: CMOS Image Sensor S5K2P1, 16-Megapixel and 1.34 μs backside illumination pixel technology.

[2] Caeleste enterprise publications: "A 0.5 noise electrons CMOS pixel", by Bart Dierickx, Nayera Ahmed, and Benoit Dupont.

Contents

List of Figures	XIX
List of Tables	XXIII
1. Background	1
1.1. Historical Evolution of CMOS Imagers	1
1.1.1. Types of Pixels	2
1.1.1.1. Passive Pixel Sensors	3
1.1.1.2. Active Pixel Sensors	4
1.1.1.3. Digital Pixel Sensors	5
1.1.2. CIS Technologies	5
1.1.3. Trends on Image Sensor Technologies	6
1.1.3.1. Backside Illumination vs. Frontside Illumination	7
1.1.3.2. Deep Trench Isolation (DTI)	8
1.1.3.3. 3D Integration	9
1.2. Photodetection in CMOS Technologies	11
1.2.1. PN Photodiode	11
1.2.2. Pinned Photodiode	12
1.3. Noise	14
1.3.1. Thermal Noise	15
1.3.2. Photon Shot Noise	15
1.3.3. Reset Noise	16
1.3.4. Dark Current	16
1.3.5. Fixed Pattern Noise	17
1.3.5.1. Photo Response Non Uniformity	17
1.3.5.2. Dark Signal Non Uniformity	17
1.3.6. Flicker Noise	17
1.3.7. Quantization Noise	18
1.3.8. Partition Noise	18
1.3.9. Random Telegraph Signal Noise	18

CONTENTS

1.4. Basic Concepts in CMOS Image Sensors	18
1.4.1. Radiometry vs. Photometry	19
1.4.2. Spectral Response	19
1.4.3. Sensitivity	19
1.4.4. Quantum Efficiency	19
1.4.5. Full Well Capacity	19
1.4.6. Microlenses	19
1.4.7. Color Filter Array	20
1.4.8. Airy Disk	21
1.4.9. Modulation Transfer Function	22
1.4.10. Electronic Shutter Types	23
1.4.10.1. Rolling Shutter	23
1.4.10.2. Global Shutter	24
1.4.11. Methods for Noise Reduction	24
1.4.11.1. Correlated Double Sampling	24
1.4.11.2. Double Delta Sampling	24
1.4.12. Frame Rate	24
1.4.13. Crosstalk	25
1.4.14. Shading	25
1.4.15. Signal-to-Noise Ratio	26
1.4.16. Dynamic Range	26
1.5. High Dynamic Range Imagers	26
1.5.1. Companding Mode Sensors	28
1.5.2. Multimode Sensors	29
1.5.3. Well Capacity Adjusting Sensors	29
1.5.4. Pulse Modulation Sensors	30
1.5.5. Multiple Sampling Sensors	31
1.5.6. Sensors with In-pixel Exposition Control	32
1.6. Conclusions	32
2. SCU: A Test Chip in CMOS Image Sensor Technology	35
2.1. Introduction	35
2.2. Selected CIS Technology	35
2.3. Evaluation Chip	37
2.3.1. Sensors Array	39
2.3.2. Pixels Circuitry	41
2.3.3. Variations in Pixels Parameters	43
2.3.3.1. Pixel Size	43
2.3.3.2. Layout of Active Diffusion	43
2.3.3.3. Threshold Voltage of Source Follower Transistor	45

2.3.3.4. Use of Microlenses	46
2.3.4. Analog Output Buffer	47
2.4. Control of Operation	47
2.4.1. Image Capture Operation	47
2.4.2. Operation Modes	51
2.4.3. Timing Specifications	52
2.5. Measurements	53
2.5.1. Experimental Setup	53
2.5.2. Spectral Sensitivity	57
2.5.3. Dark Current	59
2.5.4. Dynamic Range	61
2.5.5. Performance Comparison	62
2.5.6. Microlenses Effect	63
2.5.7. Crosstalk	64
2.6. Conclusions	64
3. Tone Mapping Algorithm	67
3.1. Introduction	67
3.2. Background	68
3.2.1. Human Vision	68
3.2.1.1. Dynamic Range Adaptation	68
3.2.1.2. Tone Mapping Techniques: Basic Concepts	70
3.2.2. Dynamic Range of Reproduction Devices	71
3.3. Tone Mapping Techniques	73
3.3.1. Global Operators	73
3.3.1.1. Miller’s Operator	74
3.3.1.2. Tumblin-Rushmeier’s Operator	74
3.3.1.3. Ward’s Scale Factor	74
3.3.1.4. Ferwerda Visual Adaptation Operator	74
3.3.1.5. Logarithmic and Exponential Operators	74
3.3.1.6. Drago Logarithmic Operator	75
3.3.1.7. Reinhard-Devlin Photoreceptor Operator	75
3.3.1.8. Ward Histogram Adjustment	75
3.3.1.9. Schlick’s Rational Operator	75
3.3.2. Local Operators	76
3.3.2.1. Chiu’s Operator	76
3.3.2.2. Rahman Retinex Operator	77
3.3.2.3. Pattanaik Multiscale Observer Model	77
3.3.2.4. Ashikhmin’s Operator	78
3.3.2.5. Reinhard et al. Photographic Tone Reproduction	78

CONTENTS

3.3.3. Frequency Domain Operators	78
3.3.4. Gradient Domain Operators	79
3.4. Proposed Algorithm	79
3.4.1. Analog Compression	80
3.4.2. Digital Compression	82
3.4.2.1. Time Stamp Code	82
3.4.2.2. Tone Mapping Code	84
3.4.3. Some Alternatives for Level per Bin Assignment	90
3.5. Simulation Results	91
3.5.1. Composition of HDR Image using Multiple LDR frames	91
3.5.2. Mathematical Simulations	92
3.6. Algorithm Comparison	98
3.7. Conclusions	101
4. TVHC: A HDR Tone Mapping Imager in Standard CMOS Technology	105
4.1. Introduction	105
4.2. Fabrication Process	106
4.3. Architecture of the TVHC Chip	108
4.4. Pixels	109
4.4.1. Pixel Digital Circuitry	112
4.4.2. Auto-Zeroing Reset Technique	115
4.4.3. Pixels Physical Characteristics	118
4.5. Digital to Analog Converter	119
4.5.1. Dark Signal Contribution Attenuation	120
4.6. V_{ref} Distribution Scheme	122
4.7. Code Generator	131
4.8. Pixels Control and I/O Interface	131
4.9. Attainable Accuracy	133
4.10. Control of the Operation	138
4.10.1. Image Capture Operation	138
4.10.2. Timing Requirements	141
4.11. TVHC Photograph and Layout	141
4.12. Other Tone Mapping Hardware Implementations	141
4.13. Conclusions	144
5. Experimental Results	147
5.1. Introduction	147
5.2. Experimental Setup	147
5.2.1. TVHC PCB Host	147
5.2.2. Optical Setup	149

5.2.3. Lens and Optical System	151
5.2.4. Operation Modes	152
5.2.4.1. Tonemapping Operation Modes	153
5.3. Measurements	153
5.3.1. Spectral Response	153
5.3.2. Dark Discharge	154
5.3.3. Sensitivity	156
5.3.4. Dynamic Range Measurements	156
5.3.5. Noise Measurements	159
5.3.6. Shading Effects due to Heterogeneous Pixels Layout	162
5.4. Captured Tonemapped Images	164
5.5. Summary of TVHC Characteristics	172
5.6. Conclusions	172
6. Conclusions and Future Work	175
6.1. Conclusions	175
6.2. Future Work	177
A. TVHC Time Requirements	181
References	185

CONTENTS

List of Figures

1.1. Passive Pixel Sensor.	3
1.2. Active Pixels Sensors schematics.	4
1.3. Digital Pixel Sensor schematic.	5
1.4. Pixel shrinking tendency.	6
1.5. Frontside (with and without light guides) vs. Backside Illumination.	7
1.6. Pixel Isolations.	8
1.7. Illustration of 3D integration idea.	11
1.8. PN Photodiode layers.	12
1.9. Pinned Photodiode layers.	13
1.10. Microlens intended performance.	20
1.11. Bayer CFA structure.	21
1.12. Irradiation pattern in observation plane.	22
1.13. Sampling of an ideal point of light source.	23
1.14. Logarithmic pixel schematic.	28
1.15. Linear-Logarithmic pixel schematic.	29
1.16. LOFIC pixel schematic.	30
1.17. Pulse Modulation pixel schematics.	31
2.1. UMC CIS Technologies for Broad Applications.	36
2.2. Color Filter and Microlenses characteristics.	37
2.3. Block Diagram of SCU.	39
2.4. Layout of SCU Chip.	40
2.5. Pixel schematic.	41
2.6. Pixel operation phases.	42
2.7. Pixels layout of performance analysis arrays.	44
2.8. Pixels layout of crosstalk analysis arrays.	45
2.9. Microlenses effect on light.	46
2.10. Microlenses 3D illustration.	47
2.11. Schematic of the Analog Output Buffer.	48
2.12. Control timing illustrating data capture in performance arrays.	50

LIST OF FIGURES

2.13. Control timing illustrating data capture in crosstalk arrays.	50
2.14. Electrical simulation of pixels data retrieval.	51
2.15. Timing illustration.	52
2.16. SCU PCB test board blocks.	54
2.17. SCU PCB test board schematic.	55
2.18. SCU PCB test board.	56
2.19. SCU measurement setup.	56
2.20. Newport 6334 lamp spectral irradiance.	58
2.21. Spectral Response.	59
2.22. Conversion gain curves from <i>VDDPIX</i> to out pad for all the arrays.	60
2.23. Discharge in darkness.	60
2.24. Signal to Noise Ratio.	62
2.25. SNRxSensitivity comparison	63
2.26. Relative sensitivity gain by microlenses.	64
2.27. Crosstalk influence comparison.	65
3.1. Organization of the retina.	69
3.2. Light levels effect on photoreceptors visual function.	70
3.3. Intersection of discharge signals with ideal analog reference.	80
3.4. Assignment of digital codes at intersection time.	83
3.5. Example TMC curve.	86
3.6. Example of the distribution of TMC codes all in one bin.	86
3.7. Example of TMC codes when levels per bins are not submultiple of total subdivisions.	87
3.8. Look Up Table of code decrements for TSC representation of 3-bits.	88
3.9. Look Up Table of code decrements for TSC representation of 7-bits.	88
3.10. TMC code distribution with zero levels assigned to first bin.	89
3.11. TMC code distribution with zero levels assigned to a middle bin.	90
3.12. LDR frames used for HDR composition.	93
3.13. Photocurrent image.	94
3.14. Photocurrent image histogram.	94
3.15. Time Stamp Image.	95
3.16. TSI Histogram of photocurrent image.	95
3.17. Tone Mapping Curve vs. evaluation subdivisions for photocurrent image.	96
3.18. Final Tone Mapped Image.	97
3.19. Histogram of the Tone Mapped Image.	97
3.20. Results from Luminance HDR.	99
3.21. Results after the application of CLAHE.	100
3.22. Results of applying the different levels per bin assignation modes (a-i), and (j) their corresponding tone-mapping curves.	102
3.23. Results of CLAHE in comparison with mode 6.	103

4.1. Photodiode layer arrangement.	107
4.2. TVHC: High-Level Block Diagram.	108
4.3. Block Diagrams of the Basic and Time Stamp Pixels.	110
4.4. 5T Amplifier in the Pixels.	111
4.5. 5T Amplifier Gain Plots.	111
4.6. 5T Amplifier Comparator Simulations.	112
4.7. Pixel SRAM Cell.	113
4.8. Pixels Digital Control.	114
4.9. Offset of the combination of the buffer, comparator and digital circuitry.	114
4.10. Simulation of auto-zero improvement.	118
4.11. Pixels Group.	118
4.12. Pixel Area Organization.	119
4.13. Voltages generated by DAC block.	120
4.14. DAC Characteristics.	121
4.15. Voltage range for Dark Signal elimination.	122
4.16. Charge Injection Amplifier Cell Schematic.	126
4.17. Charge Injection Amplifier Operation.	127
4.18. Schematic of the Amplifier in the row-wise V_{ref} distribution scheme	128
4.19. Buffer Gain Plots.	129
4.20. Folded Cascode Buffer Errors.	129
4.21. Effect of buffers offset in the distributed V_{ref} signal for $T_{step} = 1.2\mu s$	130
4.22. Operation of the V_{ref} distribution scheme for $T_{step} = 1.2\mu s$	130
4.23. Settling Error in the V_{ref} distribution scheme for $T_{step} = 1.2\mu s$	131
4.24. Sense Amplifier 1-bit Cell.	133
4.25. Exemplary Temporal Configuration for 40ms Exposure	136
4.26. Standard deviation of intersection time (T_{cross}) for $V_{ref} = 0.7V$	137
4.27. Reset voltage versus photogenerated current.	138
4.28. TVHC Operation.	139
4.29. TVHC Layout.	142
4.30. TVHC Photograph.	142
5.1. TVHC PCB general scheme.	148
5.2. TVHC PCB schematic.	150
5.3. TVHC board photograph.	151
5.4. Spectral Response.	154
5.5. Dark discharge.	155
5.6. Dark discharge after calibration.	155
5.7. Calibration SNR comparative.	156
5.8. Response to white light.	157
5.9. Dynamic range measurements.	158

LIST OF FIGURES

5.10. Photon Transfer Curve.	161
5.11. Illustration of nearest metals distribution in horizontal axis.	163
5.12. Metal shadows image capture.	163
5.13. Pixel vertical lines zoom micrograph.	164
5.14. Tungsten lamp.	166
5.15. Halogen lamp.	167
5.16. Fluorescent lamp.	168
5.17. Led torch.	169
5.18. Ceiling Lamp.	170
5.19. Natural light through window.	171
6.1. 3D-IC layers.	178
6.2. 3D-IC pixel schematic.	179
A.1. Pixel memories erase operation.	181
A.2. Reset and exposition operation.	182
A.3. Image download operation → 1 row.	182

List of Tables

1.1. Ambient luminance levels for some common lighting environments.	27
2.1. SCU Chip Main Characteristics	38
2.2. Summary of the arrays included in the chip.	40
2.3. Analog and Digital Control SCU Pins.	48
2.4. Row and column addressing of the arrays.	49
2.5. Timing description.	52
2.6. SCU PCB board devices.	57
2.7. Sensitivities at 550nm.	59
2.8. Discharge slope in darkness.	61
3.1. Example of bins distribution.	84
3.2. Levels per bin calculus.	85
3.3. Subdivisions with decrement in the case of 7 levels per bin in 8 levels.	87
3.4. Levels per bin calculus for photocurrent image.	96
4.1. Some typical values for $\rho_{k\sigma}$	137
5.1. TVHC Chip Main Characteristics	173
A.1. Control signals time requirements	183

LIST OF TABLES

Chapter 1

Background

1.1. Historical Evolution of CMOS Imagers

The evolution of electronic image sensors, or simply imagers, has been tightly bonded to that of the semiconductor industry. It could not be otherwise, since the performance of the image sensing process is absolutely dependent on the technology that is employed to fabricate the sensor. Electronic image sensors have been fabricated since the early stages of the electronics. The poor performance due to the limitations of the first fabrication processes slowed down the incorporation of these devices into the, by that time, emerging market of electronic imaging. However, the situation has changed dramatically in the last decade, where the development of new technological processes has allowed the boost of imagers in a myriad of application fields.

Depending on the technology used for its fabrication, the most common and developed electronic image sensors are:

- Charge Coupled Devices (CCD) image sensors: where the photogenerated charges are captured in potential wells in the semiconductor that are generated under Metal-Oxide-Semiconductor (MOS) structures by means of appropriate biasing. Once collected, these carriers are moved pixel-to-pixel serially, by proper clocking, to a specific readout circuitry. Carrier collection and transfer processes can be executed with extremely high rates of efficiency (in other words, with extremely low losses), and so these devices are generally considered as having excellent optical performances. However, due to the nature of the sensing method and process, it is very difficult to insert circuitry inside the pixels in order to improve its performance. Other drawbacks are the need of high voltages to create the potential wells and the inherently destructive readout scheme.
- Complementary Metal Oxide Semiconductor (CMOS) image sensors: imagers that originally¹ exploited the light sensing capability of some parasitic devices fabricated using standard CMOS mixed-signal technologies. Using standard CMOS technologies is a tremendous advantage since

¹Today, we have specific imager technologies that are primarily oriented to the optimization of the light sensing process.

1. BACKGROUND

it allows the integration of additional blocks together with the image sensor, in the same chip. Moreover, certain processes allow for the integration of circuitry inside the pixels. Last, but absolutely not least, these devices typically have a low fabrication cost¹, low power, and low voltage operation when compared to CCDs.

Both technologies have coexisted during a long time as their developments started almost in parallel. The invention of the first MOS² image sensor by Morrison was in 1963 [1] and the first report of an image sensor device in CCD technology was released in 1970 [2]. However, early MOS imagers exhibited very poor performance mainly due to excessive noise levels. Meanwhile, advances in the fabrication technology made that CCD devices experimented a significant improvement in its performances by the same time, being CCD the selected option for almost all imaging applications.

Despite the good optical performance of CCDs, power consumption, among other reasons, made necessary the evolution of the MOS imagers in the 90s. The first CMOS imager with a performance comparable to that of CCD imagers was developed by the Jet Propulsion Laboratory (JPL) in 1995 [3], which included the image sensor with in-pixel amplification (these pixels received the name of active pixels) plus exposure time control and noise suppression blocks. Since the development of these new CMOS imagers in the 90s, they have experimented a fast increment in the percentage of the imager market, especially since 2000, when their low power consumption and low cost made it possible the inclusion of cameras in mobile phones. In order to improve the performance of CMOS imagers, special technologies have been developed making the image quality comparable to CCD devices. However, certain cameras are still developed using CCD processes, such as in astronomy, where very low noise levels are required, while CMOS imagers embrace the majority of the mass market applications, also including very lately the high-end digital photography market. This is caused by the higher number of applications for CMOS imagers in comparison with CCDs, with applications such as mobile phones, laptops, surveillance cameras, high speed, scientific and industrial applications.

CMOS imagers have evolved regarding, not only the process technology, but also concerning chip architectures and circuitry within the pixels. The following subsections will briefly describe this in order of appearance.

1.1.1. Types of Pixels

Three elementary modes of pixel operation can be distinguished:

- Charge mode: The capacitor of the photodetector (integration node) is first charged to a certain voltage, which is discharged by the current produced by the incoming light. It is also named integration mode.

- Current mode: The photocurrent is directly processed, usually by means of current mirrors.

¹When using mainstream standard CMOS, or even (though less) when using special CMOS technologies.

²Complementary MOS was introduced later.

- Voltage mode: The photocurrent is converted into voltage usually based on an ohmic resistance. The typical resistance is realized by means of a transistor in subthreshold regimen, which performs a logarithmic encoding of the illumination information.

Current mode and voltage mode circuits are used for continuous time pixels whereas charge mode pixels are used in non-continuous approaches, where the photodetector has first to be precharged to a certain voltage level that will be reduced after a certain period of time (the exposition time). Charge mode circuits are the most commonly used structure due to lower noise and better linear behavior. Consequently, the following paragraphs describe the evolution of this type of pixels.

1.1.1.1. Passive Pixel Sensors

First developed CMOS image sensors included passive pixel sensors (PPS), which incorporate a photodetector and only one transistor within each pixel. Figure 1.1 shows the typical schematic of this kind of pixels. The photodetector is connected to the column line through the transistor, which is working as a simple switch. Therefore, PPS pixels are usually named 1T pixels. The column line connects the pixel to a column output amplifier. Incident light in the photodetector will generate a current that will discharge the integration node V_{ph} . Previously to the light exposition, the V_{ph} will be set to a certain voltage through the column line and ROW selection switch. This operation is named reset of the pixel. The readout will be realized by charge sensing, which introduces an error due to the high capacitance of the column line compared to the capacitance in the integration node V_{ph} .

Due to the simple circuitry besides the photodetector, PPS have a simple addressing scheme and a high ratio between the photodetector and the total area of the pixel, named Fill Factor. Nevertheless, they exhibit a small output swing and high noise, which imply a low Signal to Noise Ratio (SNR). Additionally, they have a slow readout speed due to the high capacitance of the output node when connected to the column line.

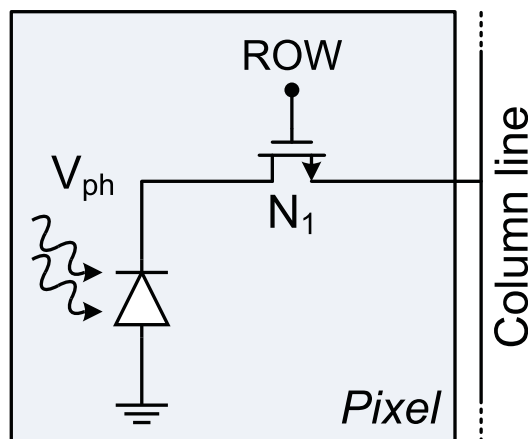


Figure 1.1: Passive Pixel Sensor.

1. BACKGROUND

1.1.1.2. Active Pixel Sensors

In Active Pixel Sensors (APS), the circuitry inside the pixel implements an amplification stage added to the selection transistor. The amplification stage is a source follower buffer, which allows for faster readout compare with PPS.

Mainly two APS pixels architectures can be distinguished, whose schematics are shown in figure 1.2:

- 3T APS: They include a source follower (transistor N_3) and a reset transistor (N_2) besides the row selection transistor (N_1). The reset transistor is used to set the initial voltage previously to the exposition to the light. The source follower replicates the voltage in the integration node V_{ph} minus the threshold voltage V_{th} of transistor N_3 . The column line provides the polarization of the source follower by a current source when the transistor N_1 is activated.
- 4T APS: They include, as a photodetector, a pinned photodiode, which will be explained in a following section. The photogenerated charge in the photodetector must be transferred via the transfer gate transistor (N_4) to the floating diffusion (FD) node in order to be read by a source follower similar to the 3T APS structure. The reset operation is performed by the reset transistor (N_2) in combination with the transfer gate. Summarizing, this photodetector needs an extra transistor inside the pixel. In this design, the photodetection and photoconversion regions are separated. Therefore, the relationship between incident light and generated signal is determined by the value of the FD capacitance.

Comparing with PPS, 3T APS have higher SNR but smaller fill factor. Moreover, the addressing is more complicated and the source follower adds noise plus a threshold voltage shift V_{th} .

4T APS have higher SNR and lower noise compared with 3T APS. Noise figures are comparable to those of CCDs. However, they have smaller fill factor due to the added transistor, more complicated

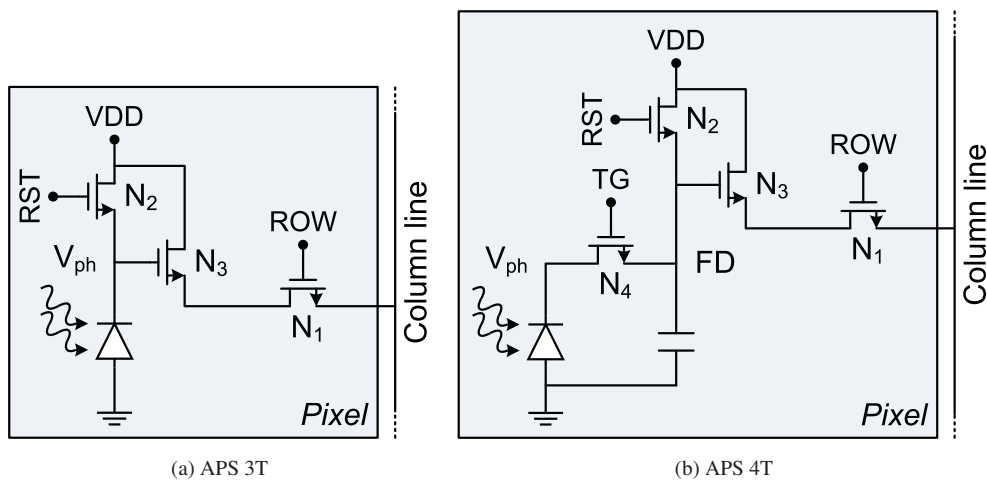


Figure 1.2: Active Pixels Sensors schematics.

addressing and similar source follower noise and V_{th} drops. However, the enhanced detector (pinned photodiode) needs a dedicated optically enhanced technological process and careful design for high performance, which imply a higher cost and effort compared with 3T APS and PPS.

1.1.1.3. Digital Pixel Sensors

In the Digital Pixel Sensors (DPS), Analog-to-Digital (AD) conversion is performed locally inside every pixel instead of outside the pixel array. Figure 1.3 shows a typical schematic of a DPS pixel. The reset of the integration node can also be performed by the digital circuitry of the pixel instead of externally, e.g. a self-resetting scheme.

The signals are digitally converted directly from the integration node. This minimizes the readout path noise producing higher SNR compared to APS. The AD conversion and storage is performed fully parallel, which allows for high speed readout as the digital data is read from the pixel array, similarly to a digital random access memory (RAM). Moreover, depending on the design, they can achieve low power operation and focal plane processing. DPS gives improved scaling with CMOS technologies due to the minimized analog circuitry. The main drawback of DPS is the increment on pixel area and low Fill Factor due to the high amount of circuitry besides the photodetector. However, there is a limitation of pixel scaling much beyond $2\mu\text{m}$ due to the nature of light and imaging optics (see subsection 1.4.8), which will benefit this kind of pixel from CMOS process scaling.

1.1.2. CIS Technologies

First CMOS image sensors were developed using standard CMOS technologies [4] and exploited the light sensing capabilities of the inherently available diodes. However, the performance of these image sensors in certain applications is poor (e.g. low sensitivity or large noise), especially as technologies scaled down. Therefore, foundries have incorporated additional steps in their standard CMOS processes to improve sensors performances, developing what are presently known as CMOS Image Sensors (CIS) Technologies. Additionally, these technologies incorporate post CMOS processing steps that are useful

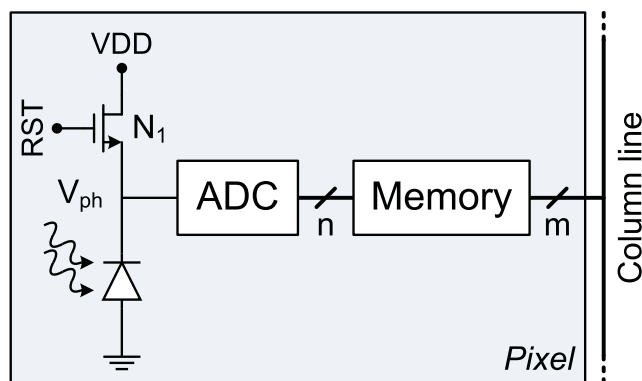


Figure 1.3: Digital Pixel Sensor schematic.

1. BACKGROUND

for imaging applications like microlenses (which focus the light directly into the detector area of the pixel) and color filters (that allow color imaging), which will be explained in following sections. During the last few years, different CIS technologies have been developed that include different optimized sensors. A first process variation consists of modifying the doping profiles of the photodiode to improve sensitivity and noise in darkness (dark current), keeping the traditional 3T-APS architecture. Another modification introduces the use of pinned photodiodes in a 4T-APS architecture with transfer gate.

It is important to remark that in these technologies, the circuitry surrounding the pixels is limited. Consequently, usually PMOS devices are not allowed because the additional well will affect the proper performance of the photodetector. Special doping profiles, including different doping levels for the substrate in the pixel area will also affect the analog circuitry in the pixels. Therefore, the development of Digital Pixel Sensor schemes is very limited in these technologies.

1.1.3. Trends on Image Sensor Technologies

As it is the case in standard CMOS technologies, where technology scaling is leading the technology evolution, pixel shrinking [5] is the common trend also in image sensor technologies. Figure 1.4 shows the general tendency of smallest pixels during recent years. Pixels below $2\mu\text{m}$ usually share circuitry, such as the select transistor and a source follower transistor but with different TG signals. This shrinking tendency is slowing down due to increasing technological difficulties at so small dimensions, which is noticeable observing the pixels below 1.4 microns.

Besides the modifications of doping profiles and the improvement of photodetectors, several additional steps are being included in order to improve the optical performance of CIS technologies, most of them already used in the past in CCD technology. These innovations in CIS technologies are explained in the following subsections.

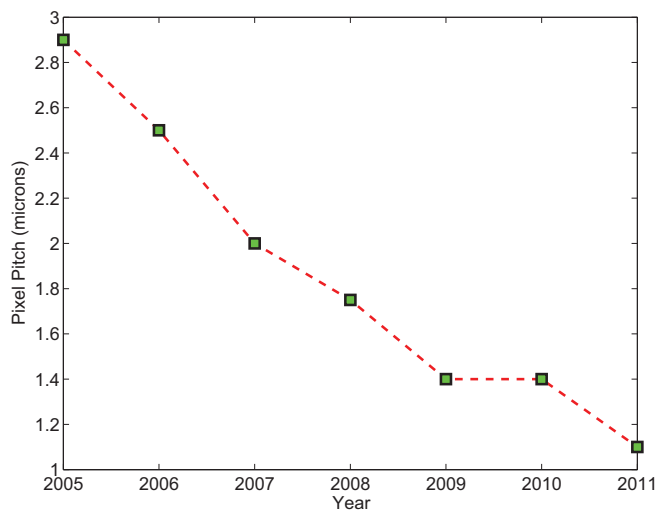


Figure 1.4: Pixel shrinking tendency.

1.1.3.1. Backside Illumination vs. Frontside Illumination

Traditionally, CMOS Image Sensors have been fabricated/ designed for Front Side Illumination (FSI) operation [6], where the light arrives to the detector after crossing through different structures. The photons travel through layers (such as microlenses, color filters, passivation layers, oxide layers, antireflective coatings) between the aperture of the metal lines of the pixels, finally creating photogenerated charges in the silicon. Figure 1.5 shows the crosssection of two pixels, where (a) shows a typical organization of a FSI scheme, where the light path can be clearly observed and photogenerated charges are indicated as red dots (e⁻):

- Microlenses: They focus the light in the area of the photodetector.
- Planarization: Layer added in the process step to create a smooth flat surface on top via chemical-mechanical methods.
- Color Filters: They filter the wavelength of the light to the required bandwidth of the spectrum. They are usually red, green and blue allowing for color CMOS imagers.
- Passivation: It protects the surface against mechanical damages and contamination by losing its chemical reactivity, avoiding effects such as oxidation due to the reaction with surrounding air.
- Oxide Layers: The light travels through these layers in the areas where no metal lines are present, although it can be reflected and refracted by them.

Camera manufacturers have aimed their efforts toward the increment of spatial resolution. This is commonly known as the race toward megapixels, which is mainly based in the reduction of pixel size. If the shrinking pixel tendency [7] continues, the main drawback of this scheme will increase: if the ratio between the pixel stack height and the dimension of the aperture is too low, the optical performance is

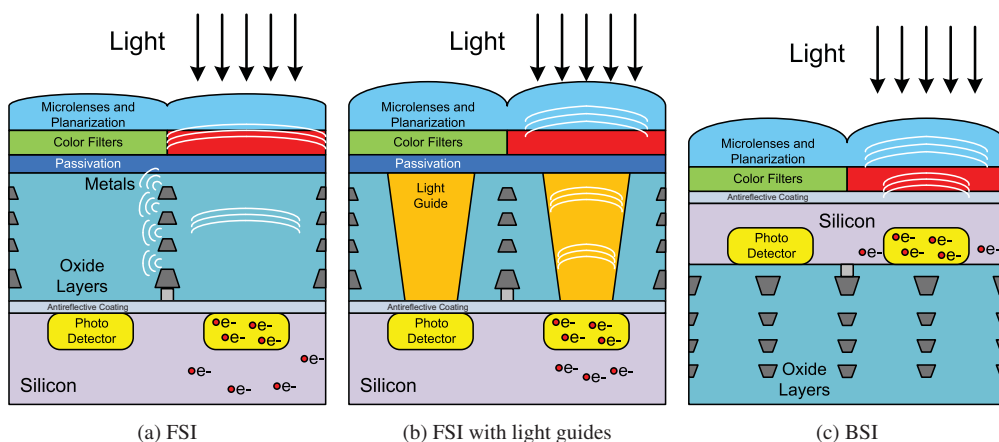


Figure 1.5: Frontside (with and without light guides) vs. Backside Illumination.

1. BACKGROUND

”poor”. In order to solve the loss of optical performance due to the low amount of light arriving to the photodetector two schemes have been developed. One variation preserves the FSI scheme but includes a light enclosing path from the passivation layer to the photodetector. Thus, it eliminates the light lost in oxide layers and scattering between metals. An illustration of this method is shown in figure 1.5(b) [8]. The second alternative is the development of CMOS imagers under a scheme inherited from CCD [9] called Back Side Illumination (BSI). In this design, the wafer is inverted and the light does not arrive to the photodetector through the top layers but through the substrate. Now, the metal lines are under the photodetector in the light path making unnecessary to keep an aperture of metal lines. Figure 1.5(c) shows the position of the typical layers in BSI scheme. Therefore, the light arrives to the photodetector through only microlenses, planarization, color filters, and passivation.

FSI has the advantage of lower cost and the maturity of the technology. Moreover, light guides achieve low crosstalk that is an issue in BSI due to the device structure. However, BSI achieves higher efficiency in the capture of incident light and the optical path can be optimized independently to the electrical one. Therefore, FSI is the dominant technology due to its lower cost but the use of BSI is increasing in applications with very small pixels (1.4 microns pitch pixel and beyond).

1.1.3.2. Deep Trench Isolation (DTI)

Lately, due to the low pixel dimensions, the unwanted effects of the neighbor pixels are larger. These effects are denominated crosstalk, whose cause will be explained in section 1.4.13. In order to reduce crosstalk, several technological schemes have been developed apart from the typical Shallow Trench Isolation (STI). Shallow Trench Isolation (STI) is the mainly used isolation technology for CMOS technologies under $0.5\mu\text{m}$. Regarding CMOS image sensor, STI prevent crosstalk between neighbor pixels via diffusion. A shallow trench is etched in the surface of the wafer and then filled with a dielectric serving as the Field Oxide (FOX) positioned as shown in figure 1.6 (a).

The first modification to STI added P+ implants creating an isolation wall in the side of the STI

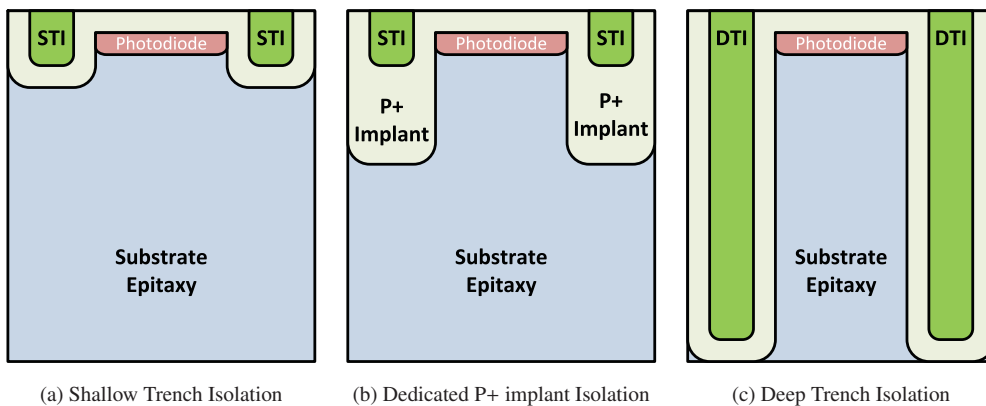


Figure 1.6: Pixel Isolations.

structure as shown in figure 1.6(b). Other solution is the use of adapted pinned deep photodiode on n-substrates also with isolation wall and STI. However, the recent trend is to substitute the STI for Deep Trench Isolation (DTI) structures [10] depicted in figure 1.6 (c). In this method, the Si-SiO₂ interface is extended as a deep wall in the substrate with high aspect ratio such as 1:25, achieving higher depth than in the other solutions. DTI is done during the same processing step of the STI formation for the external area of the pixel array. However, it is needed only one extra mask. A passivation step is done in order to avoid degradation on dark current and other failures in pixels due to the increment of defects in the larger Si-SiO₂ interface.

DTI high aspect ratio deep isolation¹ provokes a high charge collection volume, due to the created barriers that avoid losing some carriers, and therefore the improvement of collection performance. Moreover, the structure performs a sort of light guide structure due to the different refractive index between Si and SiO₂. Therefore, it improves the performance for off-axis illumination, which arrives with an inclination.

1.1.3.3. 3D Integration

Recently, the dimensions of the smallest pixels have arrived to the physical shrinking limits of imaging lenses, as will be explained in section 1.4.8. In order to solve this issue, two different options can be considered. First, the surrounding circuitry can be increased taking advantage of the shrinking of standard CMOS Image Sensors processes. Second, instead of increasing the amount of circuitry horizontally, it can increase vertically.

Microelectronic systems have typically consisted of multiple integrated circuits (ICs). The connection of these systems has been traditionally done by 2D wiring in a Printed Circuit Board (PCB), where every chip has its package. To improve performance, several chips can be enclosed in a single package side by side, named Multi-Chip Package (MCP). The natural evolution of this scheme must take advantage of the vertical dimension to achieve a smaller footprint. This leads to 3D integration, where vertically stacked chips will connect.

3D integration approaches can be divided in these categories [11][12]:

- Packaging-based: it is applied over completed chips by means of the pads [13]:
 - Package-on-Package (PoP): multiple packaged chips are vertically stacked [14]. The chips can be wire-bond [15] or flip-chip [16] type [17]. An advantage is the possibility of IC packages of being fully tested with high yield prior to stacking, whereas bare die stacking suffer from yield issues [11].
 - System-in-Package (SiP): multiple naked dies are vertically stacked in a single IC package. It is usually fabricated connecting thinned chips and their connection is achieved via the wire bonding of the enclosing package and large dimension Ball Grid Array (BGA), which are an array of solder balls.

¹Large height and narrow width.

1. BACKGROUND

- **Assembling-based:** different levels of functionalities are created with normal integration processes, each with a different substrate, and connected to each other. The main representative techniques are Microbumps and Through Silicon Vias (TSV), which is a passing through vertical electrical connection in a wafer or die with copper or tungsten. The way in which the chips are connected varies depending on the technology:
 - **Die-to-Die (D2D):** individual dies are glued together to form the 3D-IC. It is made with die-to-die bonding (Microbumps) and Through Silicon Vias (TSV). TSV are performed by laser drill or deep reactive-ion-etching and later filled with copper. The interconnect pitch is large because it is necessary to compensate for misalignment of dies. Fabrication involves several extra manufacturing steps, such as alignment, resulting in potential yield reduction.
 - **Die-to-Wafer (D2W):** individual dies are stacked on top of dies that are still not cut from a wafer. First, each individual die must be aligned to the base wafer. Interchip electrical connection can be formed by post-bond via formation or die bonding process. The alignment of each die takes about as long as the alignment of an entire wafer. The alignment process is typically the most expensive equipment used in the 3D assembly process. The advantages of these processes are that the wafers on which the different layers of the 3D stack are produced can be of different size. Another advantage is that the individual dies can be tested before the stacking process and only “Known-Good-Dies” (KGD) can be used, thereby increasing the yield of the 3D-IC.
 - **Wafer-to-Wafer (W2W):** it is realized with full semiconductor wafers, which are aligned, bonded, and diced. Therefore, they must be of same size and similar material for match of the thermal coefficient of expansion. The vertical interconnection is usually achieved using TSV. Wafer-level handling and processing allows hundreds or thousands of devices to be created at once in comparison with dies handling. Consequently, it results in a more cost effective approach.
- **Monolithic:** this approach is the vertical extension of normal standard integration technologies. Each layer is processed sequentially starting from the bottom layer. Therefore, devices are built on a substrate wafer by mainstream process technology. After proper isolation, a second device layer is formed and devices are processed by conventional means on the second layer. This sequence of isolation, layer formation, and device processing is repeated to build a multi-layer structure. This option allows for lower power consumption and faster speed than using separate circuits. Thermal cycling can degrade underlying devices. Due to the sequential processing of layers, the manufacturing throughput is low.

3D integration results in a chip that is a combination of multiple tiers (wafer layers) of thinned 2D integrated circuits. The tiers are stacked, bonded, and vertically connected, as illustrated in figure 1.7. It is possible to choose independently the best technology for each layer based on the functionality, performing a heterogeneous 3D integration. It may result in a reduction of costs and an increase in performance. First tier can be implemented using a technology aimed for image sensors. Second tier

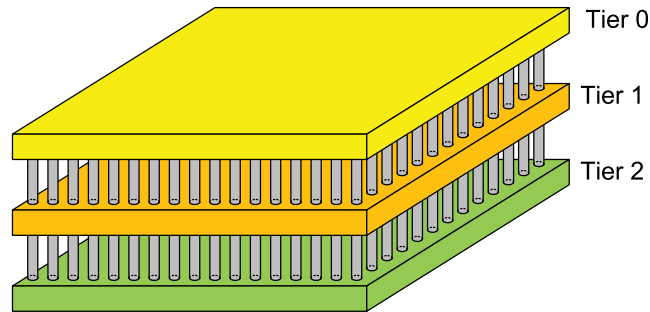


Figure 1.7: Illustration of 3D integration idea.

could be developed in a technology for high analog performance. Finally the last tier, for example including the digital processing, can take advantage of a very deep sub micron technology [18].

The more mature 3D technologies are related to packaging. However, due to the high possibilities of 3D integration, these technologies are under high research nowadays. The main advantages of 3D approaches are: more suitable form factor, cost, integration of heterogeneous technologies, performance, improvement of interconnection time delay (shorter connections), possibility of high speed and low power consumption. The main disadvantage is the thermal management and the technical difficulties in the process development.

1.2. Photodetection in CMOS Technologies

Incident light on a semiconductor generates electron-hole pairs or photogenerated carriers, due to the photons of energy greater than the bandgap energy. In order to sense these charges, the most used structure in CMOS image sensors are photodiodes [19][20]. A photodiode is based on a junction of P and N semiconductor material. In the surroundings of the junction, the area is depleted of free charge carriers, forming a depletion zone with an electric field. When photons of energies greater than the bandgap energy are absorbed in or close to a PN junction create electron/hole pairs, which are pulled by the electric field of the contact potential across the PN junction. Inserted in a pixel, the photodiodes are reverse biased creating a larger depletion zone in order to sense more photogenerated carriers.

Mainly, two types of photodiodes are used in CMOS imagers: PN photodiodes and pinned photodiodes, which will be explained in the following subsections.

1.2.1. PN Photodiode

PN photodiode [21] has been the most used photodetector in early generations of CMOS image sensors. They are already present in standard CMOS process technologies, thus enabling image sensor design within a general-purpose IC design environment. Therefore, the PN photodiode pixel is a cost effective solution.

PN photodiodes use the electric field of reverse-biased PN junction to separate the photogenerated

1. BACKGROUND

electron-holes pairs. These photogenerated carriers are created by the incident light absorbed in the depletion region. The electric field moves the carriers towards the edges of the depletion region, generating a reverse photogenerated current I_{ph} , which is (approximately) directly proportional to the incident light.

The typical P and N layers used in order to create PN photodiodes are:

- n+ diffusion/p-well (n+/PW).
- n-well/p-type substrate (NW/PSUB).
- n+ diffusion/p-type substrate (n+/PSUB)¹.

In an n+/PW photodiode, a highly doped n+ region is formed in a PW region. The increased doping concentration and reduced depth of the PW region in recent highly integrated CMOS technologies reduces the thickness and depth of the depletion layer and affects sensitivity. In terms of realistic sensitivity, n+/PW photodiode is suitable for above 0.5-0.8 μ m rule CMOS technology [20]. NW/PSUB photodiodes have a larger and deeper depletion zone compared to n+/PSUB photodetectors. Therefore, a larger photoconversion volume can be obtained, even with deep submicron processes. Therefore, NW/PSUB photodiodes usually implies better sensitivity. However, due to layout rules, well formation usually implies the need of larger pixels than the n+/PSUB photodiode option.

1.2.2. Pinned Photodiode

The layer scheme of an n+/PSUB photodiode with a 3T APS pixel circuitry is shown in figure 1.8. The dotted line indicates the boundaries of the depletion zone. The junction capacitor is charged to

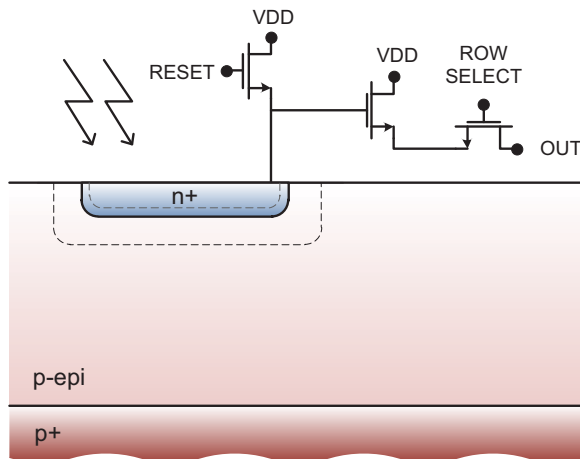


Figure 1.8: PN Photodiode layers.

¹This photodiode is obtained in CMOS twin-well technologies when in the region under the photodiode (in this case, p-type substrate) a p-well doping profile is not applied.

VDD through the RESET transistor, resulting in a wider depletion zone. The reset transistor is turned OFF and the photogenerated current discharges the sensing node which is connected to a source follower amplifier, whose voltage will be measured when the transistor ROW SELECT is ON. This output is connected to the column selection circuitry, which adds the polarization of the source follower of the pixel.

The main drawback of PN photodiodes are the noise due to dark current, explained in section 1.3.4, due to surface generation and thermal noise associated with the photodiode reset (initial charge).

In order to solve the main drawbacks of the PN photodiode, it was developed the pinned photodiode [22]. This structure is somehow a photodiode structure that has been buried in order to isolate the behavior from the noise created by the surface. Figure 1.9 shows an illustration of the typical structure of a 4T APS pixel circuitry. The pixel consists of a pinned photodiode and four transistors that include a TRANSFER GATE, a RESET transistor, an amplifier transistor (source follower), and a selection transistor (ROW SELECT). Regarding the photodiode, the top layer is a thin p+ layer under which are stacked an n layer and the p-substrate (P+NP structure). Using appropriate doping profiles, both P+N and NP junction depletion region can be joined forming a depletion zone dragged away from the surface. The p+ layer has the same potential as the p-substrate and the accumulation region is separated from the surface, where the trapped states are located producing dark current levels.

At the beginning, the floating diffusion n+ is reset by the *RESET* transistor. Then the pixel output is read as a reset reference signal, which includes the pixel offset of the source follower and the reset noise. Then, the photogenerated charge is integrated in the pinned photodiode potential-well and transferred via the *TRANSFER GATE* to the floating diffusion n+ to be sensed. The conversion charge to voltage is realized by the capacitance of the floating node n+, which is connected to the source follower. The output of the pixel is read and subtracted by the reset reference in order to obtain the final measure with compensated noise.

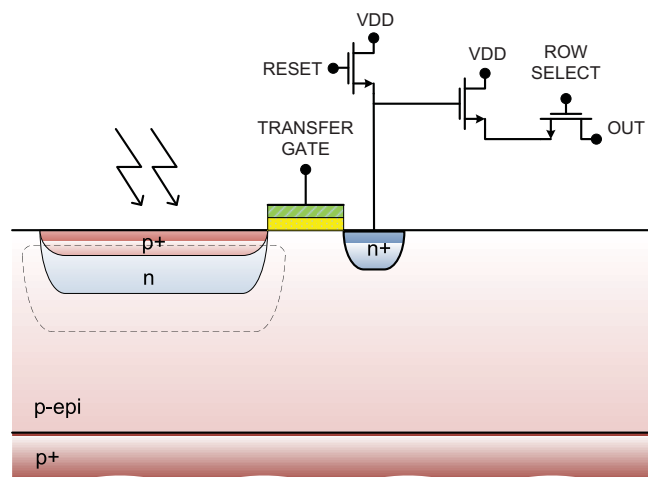


Figure 1.9: Pinned Photodiode layers.

1. BACKGROUND

There are several advantages in this structure. The photodetection and the photoconversion area are separated. The pinned photodiode has less dark current because the p+ layer masks the traps from the surface, which are one of the main sources of dark current. Therefore, the noise performance of PN photodiode is commonly worse than that of the pinned photodiode. They achieve high sensitivity and good image quality. Therefore, they can compete with CCD imagers.

Compared to the PN photodiode, it offers a smaller full well capacity, which is the maximum quantity of charge that can be storage in a pixel (see subsection 1.4.5), and higher sensitivity. In PN photodiodes, the full-well capacity can be increased through the photodiode capacitance, but in a pinned photodiode it is limited to the pinned potential of a photodiode. If the pixel is working in charge mode, this feature enhances the dynamic range. However, the main drawback of this structure is the cost and difficult design. A proper depletion zone requires not only well-formed thin p+ layer in the surface but also an elaborate design of the potential profile by precise fabrication process control. Incomplete charge transfer from the photodiode to the floating node node causes excess noise, image lag, and a nonlinear response. Therefore, the pinned photodiode and the TG should be optimized to avoid this issue; this must be performed also in CCDs but they operate in high voltage, which simplify the problem. However, in pinned photodiodes (due to the low voltage operation) very accurate dopant control is required in fabrication. High transfer efficiency requires optimizing the region between the photodiode and the transfer gate while optimizing photodiode depth. Furthermore, forming an L-shaped transfer gate will improve transfer efficiency due to a three-dimensional potential effect [20]. Additionally, the reset transistor still causes some noise. They also have lower fill factor than PN photodiode, due to the inclusion of an additional transistor (the transfer gate) and the charge capacity is limited by the pinned potential [20].

1.3. Noise

Noise in image sensors can be defined as any sensed variation from the ideal scene to be captured [20] and therefore it deteriorates imaging performance. The imperfections found in sampled images are originated in the optics, sensor and post-processing operation. However, electronic noise is related to the last two path of the signal while optical imperfections are named optical distortion and aberrations, which will be described in section 5.2.3. The noise can be presented as fixed offsets, random noise, etc. with statistical variation with time, location, temperature, signal level, position, and/or the design of the system.

At low illumination, the dominant source of noise is reset and readout transistors, which is related to the electronics. At high illumination, the dominant source of noise is the photodiode shot noise and PRNU, which is caused by the quantized nature of light. However, there are several noise components in CMOS image sensors [23]:

- Thermal Noise: It is caused by the effect of temperature in transistors.
- Photon Shot Noise: It is provoked by the quantized nature of light.

-
- Reset Noise: It originates at the reset switch of the pixel.
 - Dark Current: It is the electric current that flows through the photodetector in the absence of light.
 - Flicker Noise: It is caused by the defects in the lattice structure.
 - Fixed Pattern Noise: It is provoked by the variations caused by technological process. It is fixed in position.
 - Photo Response Non Uniformity: It is caused by the pixels different responses to light.
 - Dark Signal Non Uniformity: It is caused by the pixels different dark signal.
 - Quantization Noise: It is caused by the Analog-to-Digital conversion of the pixel signal.
 - Partition Noise: It is caused by the reset operation.
 - Random Telegraph Signal Noise: It is caused by the trap and emission of carriers in the Si-SiO₂ interface of the channel and discrete fluctuation of the dark current.

1.3.1. Thermal Noise

Thermal noise [24] is caused by the effect of temperature in the carriers, i.e by random thermally induced motion of carriers in resistive regions. It is also named Johnson noise because J.B. Johnson discovered it in 1928. This noise exists in any semiconductor. There is a constant thermal motion of carriers in any direction of semiconductor. They frequently collide with the material molecules and change the direction of motion. Therefore, in a long interval, it has a zero mean value, but in a short interval, these random fluctuations form a noise current. Noise will increased with temperature due to its effect over the mean square speed of current carriers. Thermal noise has constant power over all frequencies, i.e. white noise. Therefore, the bandwidth of circuit determines the noise value. The expression of thermal noise voltage is [25]:

$$\overline{V_n^2} = 4kTR\Delta f \quad (1.1)$$

where k is Boltzmann constant, T is absolute temperature, R is equivalent resistance and Δf is bandwidth of measuring amplifier. As thermal noise is often associated with the actual circuit bandwidth, the design of the input stage of amplifier is particularly important.

1.3.2. Photon Shot Noise

Photon Shot Noise (PSN) [26] is related to the quantum nature of light. There is an uncertainty in the number of photons collected in the photodetector. The number of photons incident in a photodetector illuminated by a perfectly uniform source of light follows a Poisson distribution. The magnitude of PSN is the square root of the mean number of electrons stored in the photodetector. Therefore, the root mean square noise voltage due to this noise is:

1. BACKGROUND

$$\overline{V_{PSN}} = CG \cdot n_{photon} = CG \cdot \sqrt{N_{stored}} = \frac{q}{C_{ph}} \sqrt{N_{stored}} \quad (1.2)$$

CG denotes Conversion Gain of the pixel, which is the relationship between number of photons and the voltage that they generate. C_{ph} capacitance of the sensing node and q is the charge of an electron. N_{stored} is the number of electron stored in the photodetector and n_{photon} is the photon shot noise.

1.3.3. Reset Noise

Reset Noise [26] is generated by the reset transistor of the pixel. In charge mode pixel, previously to the light acquisition, the pixel must be charged to a reference voltage. This operation introduces a noise related to the thermal noise of this transistor. During the reset phase, as the transistor is “on”, the reset transistor can be considered as a resistance and therefore with an associated thermal noise, which is sampled in the integration node. The root mean square voltage of the noise is:

$$\overline{V_n} = \sqrt{\frac{kT}{C_{ph}}} \quad (1.3)$$

where k is Boltzmann constant, T is the temperature and C_{ph} of the integration node.

1.3.4. Dark Current

The dark current [26] is the current present in photodetectors even in the absence of light. It is a negative effect because it is added to the photogenerated current. Thus, it reduces the dynamic range of the sensor.

The dark current is related to the pixel circuit design and technology process. There are several sources, which contribute to the total dark current in photodiodes [19], such as:

- Diffusion current: The density of minority carriers in the neutral region is under the level of equilibrium. Therefore, the system will tend to restore the equilibrium by diffusing carriers.
- Tunnel current: Under certain circumstances, the valence band carriers can reach the conduction band without the necessity of acquiring the energy to overcome the bandgap. It is denominated band-to-band tunneling and trap-assisted tunneling. These phenomena appear more frequently in narrow depletion region and therefore it takes place in highly doped junctions.
- Generation-Recombination current: In the depletion region of a p-n junction, the minority carriers are depleted. Therefore, the generation overcomes the recombination of carriers by thermal generation. The rate of electron-hole pair generation under reverse bias condition is calculated by the Shockley-Read-Hall theory.
- Impact ionization current: Under high electric field in the depletion region, free carriers are accelerated acquiring sufficient energy to break covalent bonds and produce electron-hole pairs.

-
- Surface leak current: At the surface in the Si-SiO₂ interface, the structure of the silicon becomes discontinuous introducing new energy states. These states facilitate the generation of carriers, which can be minimized by adequate technological process.

1.3.5. Fixed Pattern Noise

Fixed Pattern Noise (FPN) [27] is a spatially distributed noise, which does not change over time. It causes a measurement mismatch in the value of pixels that conforms a pattern. This noise is usually compensated because this pattern can be subtracted from the sampled image taking a reference image.

In an image sensor, spatially fixed variations of the output signal are of great concern for image quality and therefore its compensation is usually performed in image sensors. Regarding types of FPN, regular variations can be perceived by humans more easily than random variations. Therefore, column FPN (caused by column amplifiers) are more easily perceived than pixel variable FPN patterns. There are two types of FPN: (1) Dark Signal Non Uniformity, which takes place in the absence of light, (2) Photo Response Non Uniformity, which takes place in the presence of light.

1.3.5.1. Photo Response Non Uniformity

Photo Response Non Uniformity (PRNU) [28] is caused by the sensor manufacturing process variations, which made the pixels to have different responses to light. It is considered as a non-uniformity in gain. The rms value of this source of noise is proportional to the light:

$$\overline{V_{PRNU}} = K \cdot S \quad (1.4)$$

where K is a constant and S is the mean signal generated by light. The PRNU noise implies a pattern, which is unique for an image sensor. Therefore, it can be used as sensor fingerprint and it is commonly employed to solve the problem of digital camera sensor identification.

1.3.5.2. Dark Signal Non Uniformity

Dark Signal Non Uniformity (DSNU) [20] occurs because the dark current generation centers are statistically distributed throughout the semiconductor material. Moreover, the generation rate of each center varies between them. When a pixel has very high dark current in the statistical distribution of DSNU, they are referred as "hot pixels".

1.3.6. Flicker Noise

Flicker noise or 1/f noise [29] is caused by the defects in the lattice structure in the interface of the Si-SiO₂ of the MOS transistor channel, which trap and de-trap carriers producing a random current. The process of this noise is not well determined and therefore there exist several models. However, it is generally considered independent of temperature and current. The easiest model can be expressed as [30]:

1. BACKGROUND

$$\overline{V_n^2} = \frac{K}{C_{ox}WL} \cdot \frac{1}{f} \quad (1.5)$$

where K is a process-dependent constant on the order of $10^{-25} \text{V}^2\text{F}$, W and L are the width and length of the transistor, C_{ox} is the gate capacitance and f is the frequency.

1.3.7. Quantization Noise

Quantization noise [20] is generated by the Analog-to-Digital (AD) conversion performed to the analog pixel signal. An analog signal can take any value within its range. However, it is usually assigned a discrete digital bit value depending on the difference with a set of threshold values. Due to the nature of the conversion, an error is introduced which is named quantization noise.

This error will depend on the circuit used as the AD converter, such as in-pixel or column converter. It will be more noticeable in a digital image representation of low number of bits. However, the level of tolerance to this error will depend on the final application.

1.3.8. Partition Noise

Partition noise [31] takes place at the integration node due to the reset operation. It increases as the reset switch fall time of the reset pulse decreases. It is caused by the uncertainty of channel charge moving to the integration node. In order to model this noise, the transistor turn off transients must be considered plus drift and diffusion components of current.

1.3.9. Random Telegraph Signal Noise

Random Telegraph Signal (RTS) noise [32][26][33][34] is caused by two phenomena. First phenomenon is caused by the temporal fluctuations in the channel conductance of the source follower, which are caused by the trap and emission of carriers in the Si-SiO₂ interface of the channel. Second phenomenon is a discrete fluctuation of the dark current. These variations are not caused by trapped carriers, but by generation/recombination centers, which are metastable. The generation rate change instantaneous and randomly originating discrete leakage current fluctuations.

1.4. Basic Concepts in CMOS Image Sensors

Following subsections are presented in order to have a general knowledge regarding CMOS image sensors and cameras. This will lead to understand the concepts, which will be introduced in following chapters.

1.4.1. Radiometry vs. Photometry

Radiometric measurements are purely physical while photometric measurements depend on how the light is sensed by the human vision system. The radiometric unit is Watt per Square Meter, which indicates Irradiance, and the photometric unit is Lux, which indicates Illuminance [35].

Illuminance is the amount of luminous flux falling on a unit area of surface. Therefore, the combination of standard response of the human eye to wavelengths and the spectrum of the radiation will result in the photometric value. Consequently, non-visible radiation will not contribute to this value as humans perceive wavelengths mainly from 380 to 780 nanometers (photopic, see more information in Chapter 3). Within this range, the average human eye response varies depending on wavelength. The curve of response is standardized by the Commission Internationale de l'Eclairage (CIE) [36] and is known as the $V(\lambda)$ curve, or CIE Spectral Luminous Efficiency Function for Photopic Vision.

1.4.2. Spectral Response

The spectral response [20] of an image sensor is the output per incident light energy depending on the wavelength in the operating bandwidth of the sensor.

1.4.3. Sensitivity

Sensitivity [20] is the temporal change in the output caused by the incident light. It is usually expressed in volts per lux-second. Therefore, the pixel is considered to work in charge mode.

1.4.4. Quantum Efficiency

Quantum Efficiency (QE) [26] is the sensitivity to photons of an image sensor as a function of the wavelength. It is the percentage of photons of incident light in the photodetector, which produce an electron-hole pair.

1.4.5. Full Well Capacity

During the reset operation, a limited amount of electrons can be stored in the integration node in order to be discharged by the photogenerated current. Once the node is discharged, the pixel is saturated. Therefore, full well capacity is the amount of electrons that can be stored and it will determine the range of measurable light in linear sensors.

1.4.6. Microlenses

As the pixel has shrunk to dimensions comparable with the height stack of the pixel, its performance has been reduced due to the loss of angle of incident light. In order to compensate this loss, microlenses [37] have been introduced. They collimate incident light to the photodetector redirecting the light to the active area in the pixel. Otherwise, this light will hit in non-sensing areas or it will be reflected and/or

1. BACKGROUND

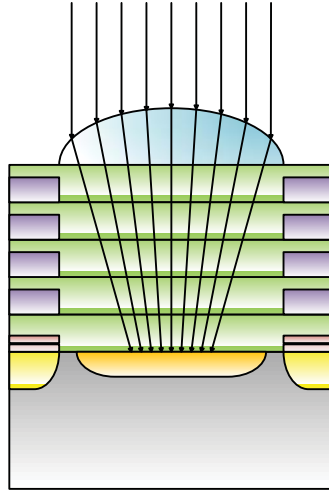


Figure 1.10: Microlens intended performance.

scattered through metals. An illustration of the purpose of the microlenses is shown in figure 1.10. The microlenses will be further explained in subsection 2.3.3.4.

To increase the light-collection efficiency even further, the gap between each microlens can be reduced introducing the “gapless” microlenses, whose base is squared instead of rounded. Additionally, double-stacked microlenses of different sizes can be introduced to further improve the angular response [38].

1.4.7. Color Filter Array

A CMOS image sensor is a monochrome sensor, which is sensitive to light wavelength that are within its spectral range. Therefore, in order to allow for color imaging, color filters must be applied to the image sensor. Color Filter Arrays (CFA) [39] are placed in top of the pixels, but usually under microlenses and planarization, to filter the wavelength of light to color components. The incident white light are usually bandpass filtered to decompose in three colors: red (R), green (G) and blue (B).

However, these filters do not have a flat response (due to their small dimension) and they overlap with each other. These color filters cover the pixel array with different filters in a regular mosaic pattern. The most used is the Bayer mosaic pattern, which is shown in figure 1.11. It double samples the green color because the human eye is more sensitive around the center wavelength of green (555nm).

The data obtained in these pixels do not represent the final color image. The image must be reconstructed from these values, this process is known as Demosaicking [40]. The image captured with the CFA is spatially undersampled in color channels. Therefore, the RGB data on every pixel must be interpolated. As an example, the interpolation can take place in a set of 3x3 neighbors. Therefore, the red channel of G pixels will be interpolated using two red values. However, if we consider B pixels, the interpolation of its red channel will be performed with four values of the neighborhood.

However, as the spatial sampling of different colors happens at different positions, there will always

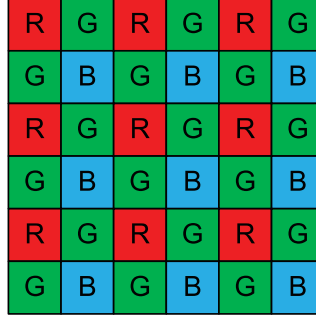


Figure 1.11: Bayer CFA structure.

be spatial artifacts at the object edges and boundaries of different colors. Moreover, as color sampling will be different for every color, after image capture some operations must be performed, such as white balancing, color interpolation (demosaicking), color correction, adjust gain and correction of boundaries.

1.4.8. Airy Disk

A point light emitting body, such as a star, captured by means of an aberration-free lens with a circular aperture will not focus in a perfect point but in a bright disc surrounded by concentric dark circles. This phenomenon is named diffraction [41], and will be present even in perfect lenses [42]. When light waves meets an obstacle, such as a small circular aperture (in this case the aperture through the optical lens), its spreading characteristic is altered bending around the obstacle. If the aperture is illuminated by a very far source of light, the light can be considered as plane wave fronts. However, this ideal source of light will be projected at the plane of observation as a spatial pattern described by equation:

$$I(\theta) = I(0) \left(\frac{2 J_1 [k a \sin(\theta)]}{k a \sin(\theta)} \right)^2 \quad (1.6)$$

where $I(0)$ is the irradiance at the center, which depends on the source strength per unit of area, the area of the aperture and the distance between the aperture and the plane of observation. J_1 is the Bessel function of the first kind. The variable k is $\frac{2\pi}{\lambda}$, where λ is the wavelength of the light, a is the radius of the aperture and θ is the angle with the horizontal optical axis of the line from the center of the aperture to the point in the observation plane. The pattern projected in the observation plane is shown in figure 1.12(a), where the axis are expressed in a relative values. Z-axis is divided by $I(0)$. X-axis and y-axis are expressed relative to the angle θ , as the real distance from the center of the pattern will be $R \cdot \sin(\theta)$, where R is the distance between the aperture and the plane of observation. Figure 1.12(b) shows a cross-section of the pattern.

The pattern has an axial symmetry, where the central maximum is surrounded by a dark ring that is created by the first zero of the function $J_1 [k a \sin(\theta)]$. This defines a center bright region, known as

1. BACKGROUND

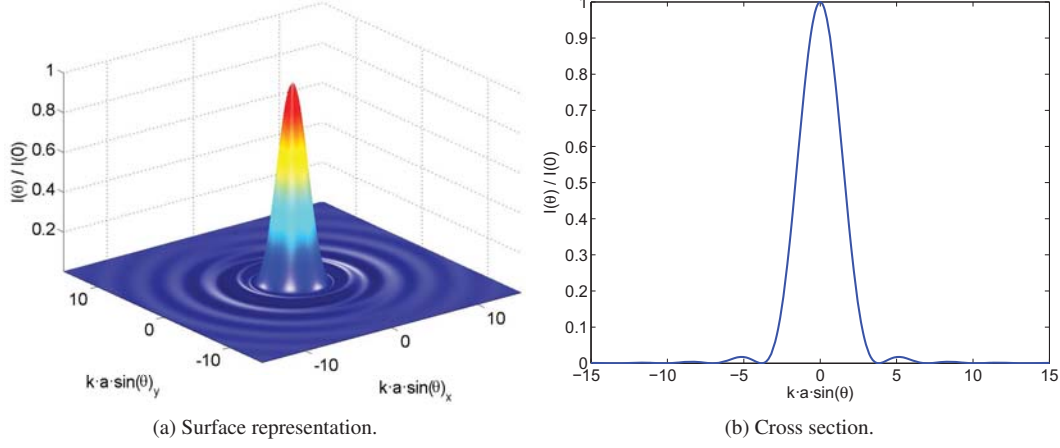


Figure 1.12: Irradiation pattern in observation plane.

Airy Disk. The radius of the first dark ring is given by the equation:

$$r_1 = 1.22 \frac{R\lambda}{2a} \approx 1.22 \lambda \frac{f}{D} = 1.22 \lambda f\# \quad (1.7)$$

where f is the focal length of the lens which acts as aperture, assuming $R \approx f$, and D is the diameter of the aperture. The figure $f\#$ is typically used in the description of lenses, which correspond with the focal length divided by the diameter of the aperture. It must be noticed that, the dimension of the first dark ring will determine the maximum resolution achievable in pixels sampling this image.

As an example, if we consider r_1 for green light (555nm) with a typical f-number such as 2.8, the radius of the airy disk is $1.9\mu\text{m}$. Therefore, if we sample the image (in the focal plane) with pixels with dimensions below this $1.9\mu\text{m}$, the captured image will be affected by the diffraction pattern causing an error. Figure 1.13(a) shows the position of too small pixels in the pattern, where an error is induced as the image should contain a single pixel receiving light. Figure 1.13(b) shows the effect of pixels with dimensions that enclose the airy disk, causing a proper sampling of the projected image. Therefore, nowadays pixel sizes below $2\mu\text{m}$ are around the physical limits of pixels shrinking. However, as explained in section 1.4.7, the color pixels are really captured by four pixels due to color channels, which performs a spatial and color oversampling of the airy disk that still maintain good performance for pixels below the boundary. In the future, approaches shrinking much further the pixels will not be necessary, as the oversampling will be insufficient, and therefore the circuitry surrounding the photosensor can experiment a boost inside the pixels.

1.4.9. Modulation Transfer Function

Modulation Transfer Function (MTF) [43] indicates the performance of the system in capturing information of an object as a function of the spatial frequencies. Therefore, it indicates the image sharpness and resolution. The MTF is influenced by, not only the image sensor, but also the optics.

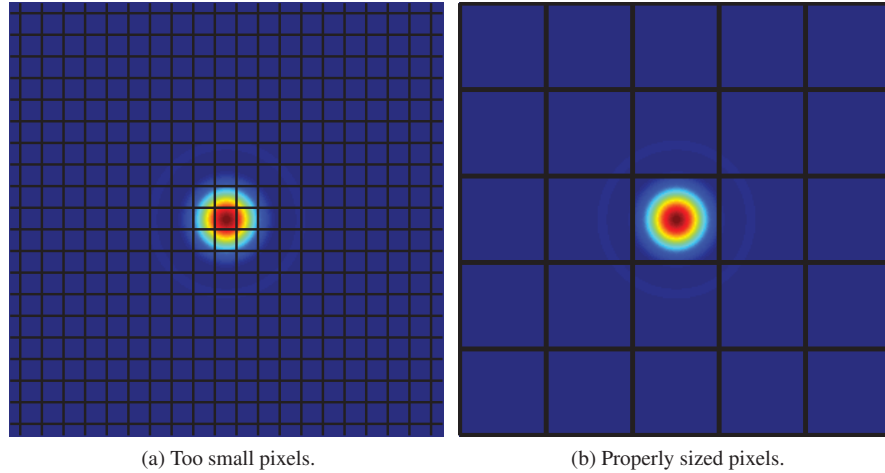


Figure 1.13: Sampling of an ideal point of light source.

It shows the number of stripes of light-dark pattern that can be photographed per millimeter in the image plane. The MTF indicates that the higher the spatial frequency of details, the greater the decrement of performance. However, the shape of the decrement indicates the MTF of the sensor.

1.4.10. Electronic Shutter Types

In photography, a shutter is a mechanism to control the exposition to the light for a period of time, which allows capturing an instantaneous scene. Traditionally, mechanical shutters have been used, which were placed on top of the image sensor or lenses. This mechanical shutter will allow passing the light through an aperture, and after an amount of time, it will completely block it by physical means. However, it is also possible to control the exposition time by means of electronic circuitry, which is named electronic shutters. CMOS cameras usually use global or rolling shutter type.

1.4.10.1. Rolling Shutter

In pixels without in-pixel memory, the time between the reset operation and the readout operation determines the exposure time. Therefore, since read-out is sequential, a gradient in the output voltages can be observed due to the differences in the effective exposure time. The integration time for pixels that are read later is augmented by the read-out time of previous pixels. This is specially noticeable when the readout time is high and/or arrays are large [44]. In the rolling shutter mechanism [45], the start and end of the light collection for each row is slightly delayed from the previous row. This minimizes the gradient, which will be present when applying simultaneous reset but sequential read out. However, it leads to image distortion when capturing scenes in motion. This will be further explained in section 2.4.2.

1. BACKGROUND

1.4.10.2. Global Shutter

The global shutter technique [46] uses a memory element inside each pixel to provide the functionality of a mechanical shutter. This allows a simultaneous reset, exposure and store (in the in-pixel memory) of the whole array. Therefore, the data is not affected by the delay between the readouts of the pixels.

1.4.11. Methods for Noise Reduction

1.4.11.1. Correlated Double Sampling

Correlated Double Sampling (CDS) [47] is a method that suppresses kTC noise from pixel reset, 1/f noise from the in-pixel source follower and Fixed Pattern Noise (FPN) (mainly originated by pixel-to-pixel variation in source follower threshold voltage). In this method, two samples from the pixel are taken, one with the data at the chosen exposition time and other during or just after reset. The reset signal is then subtracted from the signal at exposition time (read signal), resulting in a pixel signal with noise reduction.

It must be remarked that, the CDS method is effective only if the pixel signal is captured linearly and the mismatch has mainly an offset component. These assumptions are not always true and more elaborate CDS circuits, which can compensate for gain mismatch and nonlinearity in the circuits, are required.

1.4.11.2. Double Delta Sampling

Although the CDS reduces noise to a large extent, it also includes a component of FPN due to mismatch in the CDS circuits. These circuits are usually shared by columns of pixels and therefore introduce column-FPN. Double Delta Sampling (DDS) method is used to reduce column-FPN in the CDS method [48][49]. This method includes a column circuitry to remove the offsets due to column drivers and hence reduce column-FPN. The DDS circuit calculates the difference between the voltages from two consecutive read-out operations per column channel. During the read operations, the actual voltage of the CDS data (reset and read signals) are read out and stored in analog memories (capacitors). Then, the capacitors storing these two signals are short circuited. The two outputs of the DDS circuit are the difference of the reset and read signals minus the short-circuited value. The two output signals are sent out of the chip and subtracted (read-reset) by the subtraction circuit in the off-chip data acquisition system. Thus, the complete operation of the double delta sampling circuit is performed.

1.4.12. Frame Rate

The output of the majority of image sensors is an image with the value of the output of each pixel, i.e. a frame. The time of delay between the generation of a new frame determines the frame rate [50], which is expressed in frames per second.

There are several applications, such as automotive or surveillance, which needs a high frame rate to respond as fast as possible. However, the processing system should have sufficient computational power to be fast enough to utilize the high amount of data produced by high frame rate sensors. Moreover, it will need more resources in image sensors with large bit count representation.

1.4.13. Crosstalk

The influence that neighbor pixels have on a pixel is named Crosstalk [51]. It implies an error in the output of the pixel as it includes an amount of signal that does not belong to the pixel, since light was not directly impinging on it. Therefore, it degrades the spatial resolution, reduces overall sensitivity and induces poor color separation.

The crosstalk can originate by three main causes:

- Spectral: It is due to the effect of the color filters. It is caused by the non-zero transmission of the filters in wavelengths, which they should be blocking.
- Optical: Part of the light coming with a certain angles reaches a neighbor pixel. It is caused by the photons coming from neighbors before electron/hole generation. It is caused by reflection, refraction, or diffraction in top layers or even substrate. This type of crosstalk can be minimized by the use of microlenses.
- Electrical: It is generated by photogenerated carriers in a photodetector, which moves to a neighbor pixel. These carriers are generated deep in the photoconversion area by long wavelength light and they may diffuse. It happens in the silicon due to electrical mechanisms such as diffusion and drift. This type of crosstalk depends on the pixel design.

1.4.14. Shading

Shading [20] is a slowly varying or low spatial frequency output variation in an image. Usually, the center of the image appears brighter while a gradient is present toward darkening the periphery. The causes of shading can be:

- Dark Current: Local heat sources that exist introduce a thermal distribution, producing dark current gradients.
- Microlens: The efficiency of a microlens is diminished at the periphery of an imaging array due to the higher angle of incident in light. Therefore, the output of pixels at the periphery decreases.
- Electrical: Non-uniform biasing and grounding may cause shading.
- Optical: Even in the absence of microlenses, the angle of the incident light is higher in the periphery. Due to the metals stack height, the periphery pixels will loss performance because they receive light with higher angles. The aperture in the optics can increase the difference of angles between the center and the periphery. Optical shading can be minimized by proper design of the

1. BACKGROUND

microlenses taking into account the difference in angles and therefore with non-heterogeneous array of microlenses.

1.4.15. Signal-to-Noise Ratio

The Signal-to-Noise Ratio (SNR) [20] is the ratio between the signal and the noise in the sensor. This determines the quality of a certain image and it will depend on the condition of the capture, such as the exposition time. Noise in image, such as dark current discharge, will increase with exposition time. Therefore, in linear image sensors higher exposition time decreases the SNR of the image.

1.4.16. Dynamic Range

The Dynamic Range (DR) [20] is the ratio between the largest and the smallest signal level that can be measured. The smallest signal to be measured is the noise floor. In other words, the noise equivalent illumination is the illumination that will generate a signal equal to the signal generated by the noise. The largest signal to be measured in simple APS pixels will be determined by the full well capacity as it will determine the saturation point.

$$DR(dB) = 20 \log_{10} \left(\frac{Light_{max}}{Light_{min}} \right) \quad (1.8)$$

Most sensors have a DR around 60-70dB, which is mainly determined by the integration node capacitance. However, for some applications, this DR is insufficient. Applications, such as automotive, require 100dB or more of dynamic range.

1.5. High Dynamic Range Imagers

Real world scenarios contain a large range of illuminations. The dynamic range of the natural world is high where the difference in contrast is over 100,000,000:1 as indicated by table 1.1 [52][53]. We have to distinguish two kind of High Dynamic Range (HDR) situations: (1) inter-frame HDR, when the HDR illumination does not take place simultaneously (scenarios depicted in table 1.1), (2) intra-frame HDR, when the high difference in illuminations takes place in the same frame or scene. For example, a large intra-frame HDR scene can occur if it includes both an outdoor area illuminated by direct sunlight and an indoor area illuminated by interior light or even in shadows. Inter-frame HDR captures can be performed controlling the exposure time. However, intra-frame HDR scenarios imply more complex mechanisms in order to be properly captured.

Therefore, typical cameras are limited in intra-frame dynamic range and they are not able to properly capture high dynamic range scenes. Some applications need cameras that are prepared to capture images in HDR scene environments in order to work properly. Therefore, DR is a very important figure of merit in CMOS Image Sensors.

Condition	Illuminance(lux)
Clear night sky	0.001
Quarter moon	0.01
Full moon	0.1
Late twilight	1
Twilight	10
Heavy overcast	100
Overcast sky	1,000
Full daylight	10,000
Direct sunlight	100,000

Table 1.1: Ambient luminance levels for some common lighting environments.

The linear response of APS image sensors is very limited in dynamic range and therefore variations in image sensor designs are introduced to create High Dynamic Range image sensors. In order to improve this characteristic, there exist two options: (1) reduce the noise floor or (2) extend the saturation point toward higher light intensities. Leaving aside solutions consisting on cooling the sensor, as it is commonly done in astronomy applications, the reduction of the noise floor level has almost unavoidably led to introducing great modifications in the silicon fabrication process, and so related to the use of CIS technologies.

Regarding the extension of saturation point, two main options arrive immediately: linear or compressive acquisition. If the pixel is designed with linear acquisition, the binary image representation will need at least 17 bits/pixel for a DR greater than 100dB, $\log_2(10^{\frac{100}{20}}) = 16.6$. This high amount of data necessary for linear acquisition defines two main tasks in HDR imaging, understood as the overall task of producing an optimum HDR scene representation:

- Image capture: to capture the scene information by means of sensing light. Unavoidably, this task must be performed at pixel level.
- Compression: to reduce the amount of data necessary in the HDR image representation. It can be performed in-pixel (non-linear light acquisition), system level, or post-processing by software in a computer. This fact proposes a classification of pixels depending on a linear or compressive acquisition.

Regarding possibilities not related to technology, and so devoted to the extension of the saturation point, different mechanisms have been proposed in the literature to improve the DR of image sensors, which will be described in the following subsections [54]:

- Companding Mode Sensors: They perform an analog compression, usually due to logarithmic behavior of a load.
- Multimode Sensors: They have several modes of operation. Usually, a linear response is applied to low illumination levels and a logarithmic response is applied to high light levels.

1. BACKGROUND

- Well Adjusting (Clipping) Sensors: The capacity of the well is adapted to extend the DR.
- Pulse Modulation Sensors:
 - Pulse Width Modulation Sensors: They measure the time to reach the saturation signal expressing the pixel signal by a pulse width.
 - Pulse Frequency Modulation Sensors: The pixel signal is converted to frequency.
- Multi-sampling Imagers: Several images are taken with different exposition times and later composed in an HDR image.
- Sensors with local exposition control: Every pixel decides its exposition time in order to improve DR.

1.5.1. Comanding Mode Sensors

Comanding sensors [55] compress the pixel signal instead of having a linear representation. The compression increases toward the higher illumination levels, and it is usually done by a logarithmic transformation. The logarithmic conversion from the photogenerated current to a voltage with logarithmic compression is realized via a non-linear load. The load is a MOS transistor that operates in the sub-threshold region, which shows a logarithmic I-V characteristic. For typical photocurrents in the pA to nA range, which by nature will not be much higher, the load transistor will operate in sub-threshold. Figure 1.14 shows a typical pixel schematic.

These pixels work in continuous time operation mode (non-integrating). The main drawbacks are slow response time for low light levels and small voltage swings. Their main drawback is the large FPN. Several noise reduction techniques, pixel or chip level, can be used in order to achieve a noise performance similar to integration mode pixels.

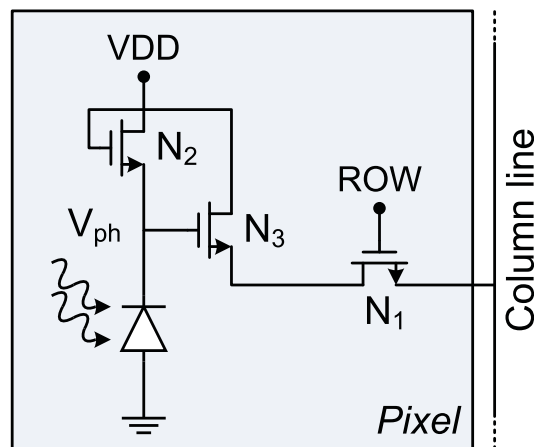


Figure 1.14: Logarithmic pixel schematic.

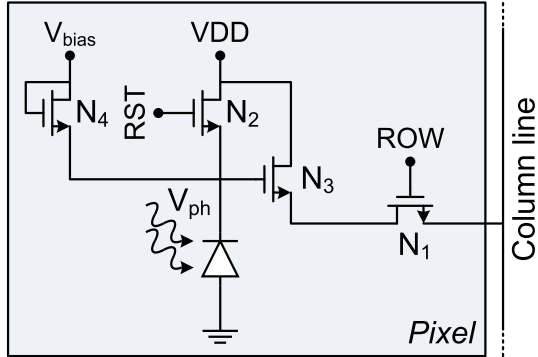


Figure 1.15: Linear-Logarithmic pixel schematic.

1.5.2. Multimode Sensors

Multimode sensors have multiple modes of operation and combine them in order to improve DR. An example of such sensor is linear-logarithmic sensors [56]. The schematic is shown in figure 1.15. At low illumination, they operate as conventional linear pixels and, at high illumination, logarithmic compression is applied.

Usually, the criterion for choosing one operating mode is derived from the characteristic of the output signal at that illumination. Therefore, the logarithmic compression is used in the higher illumination in order to extend the point of saturation.

1.5.3. Well Capacity Adjusting Sensors

The well capacity adjusting (clipping) sensors adapt the maximum charge that can be accumulated during integration, in order to overcome the limitation of the capacitance of the integration node. Usually, the charges that exceed the capacity of the integration node is derived or drained to another node achieving an extended non-linear photoresponse curve.

A typical approach of these sensors is the use of a Lateral Overflow Integration Capacitor (LOFIC) [57]. Figure 1.16 shows an example of this type of pixels [58]. The scheme is realized by a pinned photodiode with transfer gate TG and a floating diffusion FD used to convert charge to voltage. The switch S is used to drain the overflow charge into the capacitance C_S .

First, the floating diffusion FD and the capacitor C_S are reset through the transistors N_5 and N_2 . During integration time, if the photodiode is saturated the additional charge will overflow through the transistor N_4 . Consequently, this charge will be collected by the capacitor C_{FD} and C_S , which are connected through transistor N_5 . After integration, switch N_5 is turned off. Then, the signal in the photodiode is read activating the switch N_4 . Finally, N_5 is again connected to distribute the charges stored in both capacitors depending on the ratio of C_{FD} and C_S . The voltage at FD node minus the source follower threshold voltage of N_3 is read out.

The advantage of this approach is the possibility to minimize the capacitance in the FD node and

1. BACKGROUND

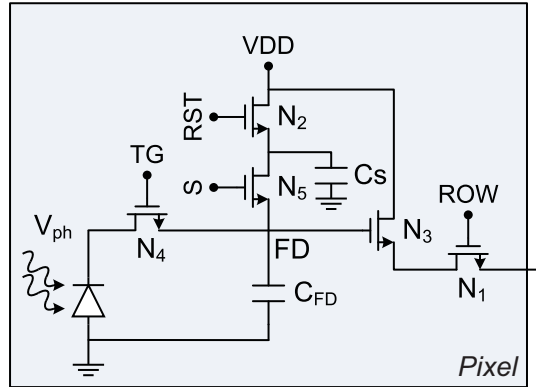


Figure 1.16: LOFIC pixel schematic.

therefore it achieves higher sensitivity. However, the loss of full well capacity is compensated by the overflow capacitance for high saturation. Summarizing, it achieves high SNR, high sensitivity, high Full Well Capacity, and low dark current. The drawback of this sensor is the decrement of fill factor due to the capacitor and extra switch.

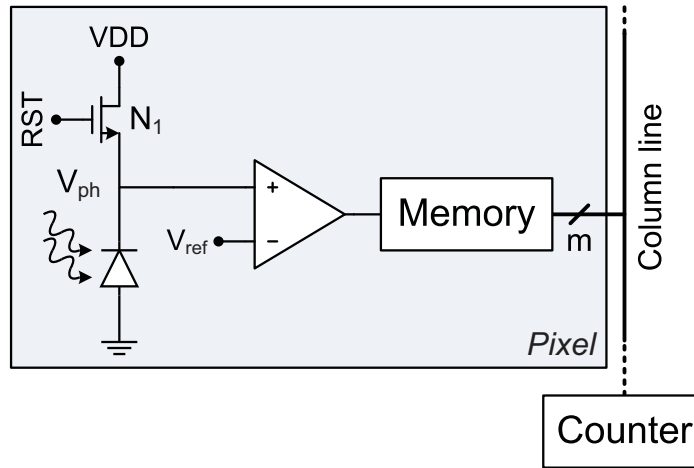
1.5.4. Pulse Modulation Sensors

Pulse Modulation sensors [59] use the time that takes the integration node to reach a certain reference analog voltage. Therefore, instead of express the light intensity as a value of charge, voltage, or current, it is expressed as time. If reference voltage is not fixed, it will perform an on the flight compression. It also can perform an in-pixel conversion and therefore this type of sensors is usually used to perform DPS pixels.

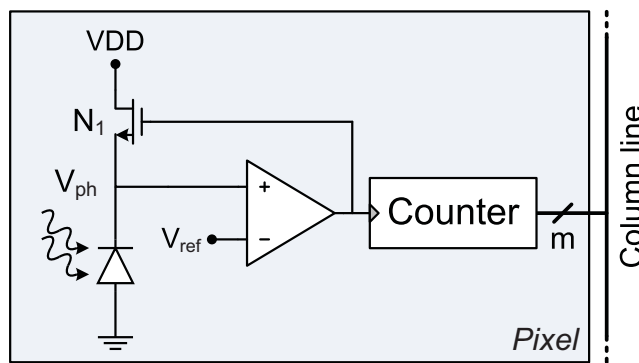
These imagers can be divided in two main categories:

- Pulse Width Modulation (PWM): The time between the reset and crossing the reference voltage is measured. The width of the pulse is inversely proportional to the photogenerated current. A typical pixel is shown in figure 1.17(a).
- Pulse Frequency Modulation (PFM): After reset, when the integration node crosses the reference level an event is triggered, which will be counted usually in pixel-level, and performs a new reset. This operation is continuously repeated performing a photocurrent controlled oscillator, where the number of pulses in a period of time is a linear function of the incident illumination. The Dynamic Range depends on the number of bits of the counter. A typical pixel is shown in figure 1.17(b).

The readout schemes of these sensors can be perform in a frame-based or event-based way. The frame-based structure reads out the content of the pixels row by row in a sequential manner at the end of the exposition or several times in order to determine the pulse characteristics. The most typical event-based readout is the Address Event Representation readout. In this asynchronous readout type, once the data is ready, the pixel requests for the output bus. Once the acknowledgment takes place, a signal is



(a) PWM



(b) PFM

Figure 1.17: Pulse Modulation pixel schematics.

sent back by the chip-level AER control to start the transmission of data. When multiple requests occur, an arbiter decides the sequence of readout. This scheme provides an efficient use of the output bus. A PWM pixel with global reset is usually denominated “Time-To-First-Spike” (TTFS), where typically event-based readout is used.

Dynamic Range is no longer limited by the power-supply rails and therefore by the voltage swing of the pixel. Consequently, in these pulse-based (time-based) schemes, the dynamic range is limited by which range of integration time can be measured. The disadvantage of these methods is the high decrement of the fill factor.

1.5.5. Multiple Sampling Sensors

These sensors performs several frame captures but with different exposition times [60]. In every individual capture, the whole array of pixels will take the same integration time. This method produces

1. BACKGROUND

a set of images, which will be combined in order to obtain a final image with improved dynamic range.

The most straightforward method is to combine the frames linearly captured with exposition times that are half of the previous integration time ($T, \frac{T}{2}, \frac{T}{4}, \frac{T}{8}, \frac{T}{16}, \dots$). If a frame is taken with exposition time T_{exp} with a resolution of N bits, the half darkest pixels of a frame with exposition time $\frac{T_{exp}}{2}$ will contain the previous frame with resolution $\frac{N}{2}$. Additionally, the half brightest pixels will contain the frame with an extension of the dynamic range with resolution $\frac{N}{2}$. Therefore, one frame will be combined to the previous frame as the half brightest pixels of the combination. The image representation will be extended doubling the number of bits but the resolution toward the brighter pixels will be diminished. Therefore, this method will have a very extended bit representation that is not optimized. Therefore, there have been developed other methods, as for instance combination of several captures by closest value to saturation, multiple captures at high frequency, overlapping the captured frames, etc.

The drawback of these sensors is the high storage/computation resources usually required and the high time of per frame operation making difficult video applications. If the frame rate is low, due to the duration of captures and combination, there will be differences between frames appearing errors in the final combination, such as ghosting artifacts. As this method can be applied to any normal camera in post-processing of frames and even can be realized by software, it is by far the most extended method for general purpose applications.

1.5.6. Sensors with In-pixel Exposition Control

In this sensor, the exposition time is chosen by each pixel individually [61]. The exposition time is chosen depending on the level of incident illumination.

A typical approach is implemented by means of a conditional reset, which will be realized if the integration node crosses a reference at certain time points. The operation is similar to the PFM pixels but the point of the reset will be at a controlled time. The number of resets is stored in the memory unit and is updated every time the pixel is checked. At the end of the exposition, the analog value of the pixel is read as normal APS with an external Analog-to-Digital converter. The digital information of the number of resets is used to scale the digitized analog value to perform a “Mantissa-Exponent” representation.

The DR extension in such sensors is bounded by the minimal time of the memory update process, which occurs between the saturation checks. However, the final representation of the image is usually large, if no compression is applied, and the post-processing of these data will need high/moderate resources.

1.6. Conclusions

In this chapter a general view of the image sensor field, focusing on what can affect HDR applications, has been presented. Regarding the overview of HDR imagers, it must be noticed that an optimization light/representation is not a common scheme. In companding and multimode sensors, fixed compressions (such as logarithmic) are used. A fixed compression does not take care of the distribution

of the illuminations of the scene and cause a lack of details/contrast if a large population of pixels is in bright areas. The well capacity adjustment sensors usually provoke a limited extension of the dynamic range and similarly apply a non-optimized compression toward the brightest pixels. Pulse modulation sensors highly improve the dynamic range, however the representation of the images is not optimized for processing. Moreover, if the design uses event-based readouts a reconstruction of the frame must be performed in order to apply post-processing to the image. Regarding multiple sampling sensors, the generation of high quality images usually needs high resources or it is a slow approach. Sensors with in-pixel exposition control, such as “mantissa-exponent” approaches, generate a representation of the images, which is difficult to compute and contain a high amount of data.

1. BACKGROUND

Chapter 2

SCU: A Test Chip in CMOS Image Sensor Technology

2.1. Introduction

This chapter presents our first attempt to expand the dynamic range of the sensors. It is included in the works toward Focal Plane Processors that our group have been researching since more than 15 years. To that purpose, we explore here the possibility of increasing the Dynamic Range by reducing the noise floor by using a CIS technology. Obviously, our academic context does only allow us to use technologies that offer Multi Project Wafer (MPW) options, since our available budget cannot cover the costs of going through a dedicated fabrication of many wafers, nor even going for small engineering lots (fabrication of a reduced number of wafers [2-3]). The chapter starts with a description of the selected technology, it follows with a description of the design of a test chip, it continues with the obtained experimental results, and finally the conclusions and lessons learned are presented.

2.2. Selected CIS Technology

Due to its availability through Europractice [62] at the time of facing this design (year 2007) and the affordable fabrication cost (around 20.000 Euro for a 5x5mm² block), we selected the 0.18 μ m CIS technology offered by United Microelectronics Corporation (UMC) [63]. As depicted in figure 2.1 [63], the range of applications¹ of this technology is quite broad covering Mouse/Toys, Surveillance, Web Cams and DSC/DSLR/Camcoders.

The UMC CIS 0.18 μ m technology is based on a 2P4M generic mixed mode standard CMOS process. It offers two types of sensors to the users: (1) a conventional 3T-APS PN Photodiode with optimized implants and doping profiles, and (2) the so-called Ultra Photodiodes, which use a 4T-APS architecture

¹Though this is somewhat obsolete as it is a vision of the CIS scenario that is more than four years old.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

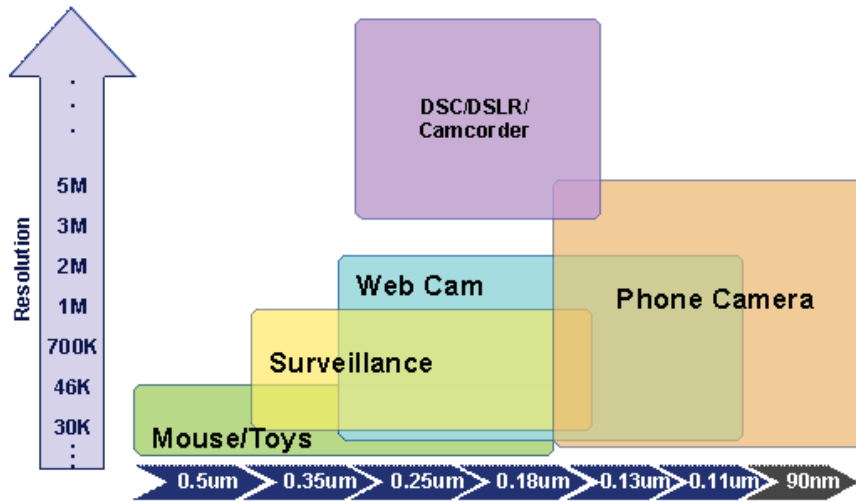


Figure 2.1: UMC CIS Technologies for Broad Applications.

and a pinned photodiode. The most distinctive characteristics of this CIS 0.18μm technology are [63]:

- Low resistivity epitaxy layer: 4μm or 7μm P-epi on P+ substrate
- Dedicated sensor implant
- Ultra-PD architecture → buried photodiode
- Sensor performance (dark current, etc.) depending on photodiode's architecture (conventional or ultra)
- Special Planarization (if microlenses are used)
- Shallow Trench Isolation
- Retrograded twin well (Deep N-well is optional) [64]
- Dual gate oxides
- CoSi₂ Poly gate and S/D
- Thinner Back-End-Of-Line (BEOL) stack (as compared to the standard CMOS process)
- 3-4 Al. metal layers with thin backend process (110nm):
 - 0.28μm pitch for M1 and mid Metals
 - 1.7μm film stack (from photodiode to color filter)
- In-house capability for incorporating Color Filter Arrays (CFA) and gapless Microlenses

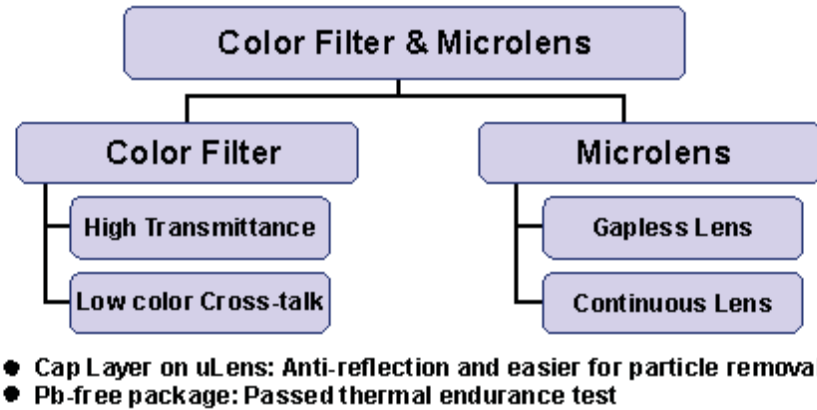


Figure 2.2: Color Filter and Microlenses characteristics.

The color filters are high transmittance and low color crosstalk with RGB color filter array. The microlenses are gapless (without a space between them) with a cap-layer on them [65], which improves anti-reflection and allows an easier particle removal [63]¹.

The devices available in the technology are:

- 1.8V Regular, Low and Zero Threshold Voltage Transistors
- 3.3V Regular, Low and Zero Threshold Voltage Transistors
- Passive devices: Resistor, Poly-Insulator-Poly (PIP) Capacitors and Metal-Insulator-Metal (MIM) Capacitors

Due to the modifications included in the fabrication process, only a restricted set of devices is allowed within the sensors area. Specially critical is the fact that PMOS are not permitted in the pixels, and they can only be used in the circuitry surrounding the array of sensors. A reason for this restriction is that, as explained in chapter 1, the incorporation of Nwells negatively affects enhanced photodetectors in Psubstrates, since these Nwells create large irregularities in the structure of the lattice that affect negatively to the noise performance of the sensor.

2.3. Evaluation Chip

As already mentioned, the main objective of this work was to explore the possibility of using CIS technologies to improve the sensing performance (basically dynamic range) of the Focal Plane Processor chips that are designed by our group. To that purpose, and due to the lack of information provided by the foundry regarding the performance of the sensors, which, on the other hand, are extremely dependent on sensor physical design (size, shape of the sensor, surrounding circuitry, structures on top of

¹More details about the technology cannot be disclosed here due to confidentiality agreements.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

it [vignetting], etc.), we decided to design a test chip in order to measure what this CIS 0.18 μ m technology could offer us. The objective was, then, to analyze the main characteristics of the photosensors and to evaluate possible strategies for dynamic range improvement methodologies [54], bearing in mind the desired compatibility with Focal Plane Processors typical circuitry. In our study, we have analyzed sensitivity, dark current and crosstalk, and how these factors are affected by photodiode's size, shape, use of microlenses, and use of Low-Voltage transistors in the 3T-APS circuitry.

UMC 0.18 μ m technology is offered through Europractice in two different choices: (1) the Conventional Photodiode, which is an enhanced PN photodiode, or (2) the Ultra Photodiode, which is a pinned device. Since neither a set of macros (predefined sensor layout) nor other information about the performance of the Ultra PD is provided by the foundry¹, and we cannot afford the cost of an iterative tuning (several runs) through design-fabrication-measurements of the pinned sensor, we opted for the conservative option of using the conventional photodiode, which is less complex as it contains a lower amount of modifications from an ordinary CMOS technology photodiode, making this option more secure for first silicon implementation success.

It is important to remark that N-Well and P-implant regions are only allowed outside the pixels area. Consequently, neither PMOS devices nor substrate contacts are permitted inside the pixels. The lack of substrate contacts is compensated by a very low resistivity substrate in pixels area (described low resistive epitaxial layer). However, NMOS only circuitry inside the pixels implies a serious constraint for focal plane processing architectures.

In order to characterize the performance of the conventional photodiodes in this CIS technology, the test chip (internally named as Sensors Chip from UMC or SCU) includes 12 arrays of pixels based in the 3T Active Pixels Sensors (3T-APS) structure, which is described in section 2.3.2. These arrays are addressed using an unified row and column bus (as parts of an unique sensor array whose dimension is 256x139). The chip incorporates a single output buffer to avoid column-wise fixed-pattern-noise (FPN) effects. The block diagram of the chip is shown in figure 2.3 and its main characteristics are summarized in table 2.1.

Technology	UMC 0.18 μ m 4M/2P CMOS Image Sensor (CIS) Technology, Conventional Photodiode
Package	84 pins Ceramic Leadless Chip Carrier (CLCC84)
Pixels Arrays	8 Arrays of 64x64 pixels 4 Arrays of 11x11 pixels
Pixel Pitch	3 types: 3.5, 5, 7 μ m
Fill factor	3.5 μ m pitch: 46 % 5 μ m pitch: 60 % 7 μ m pitch: 70 %
Number of transistors	139110
Power supply	dual supply: 1.8V and 3.3V
Die Size	2358x2358mm ²

Table 2.1: SCU Chip Main Characteristics

¹The Ultra PD is mainly devoted to company's own developments.

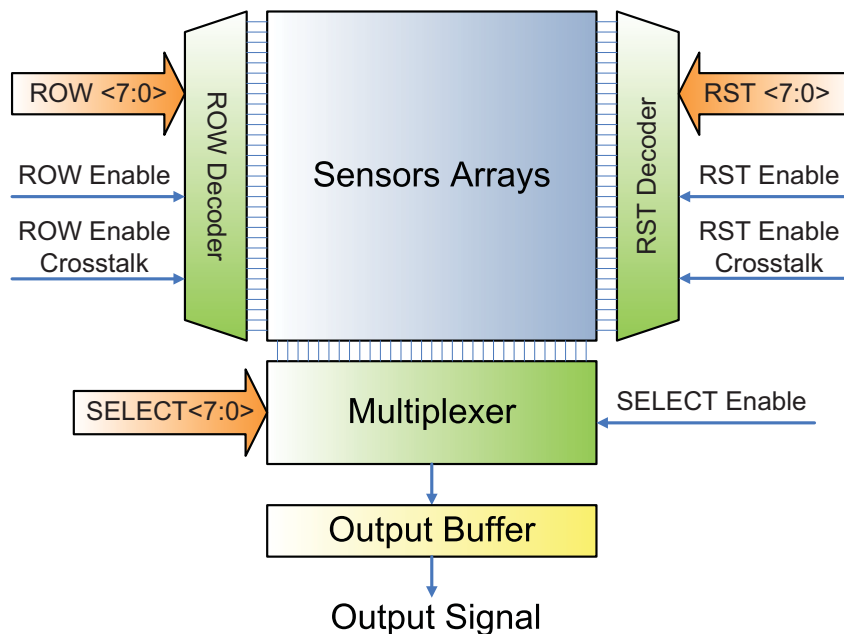


Figure 2.3: Block Diagram of SCU.

The chip employs a single row decoder with separated enable signals depending if the arrays are intended for either performance or crosstalk analysis (see section 2.3.1), whereas both types of arrays share the same addressing bus ROW<7:0>. Similarly, columns are addressed using a single multiplexer controlled by SELECT<7:0> and an enable signal. APS reset operation is executed row-wise (through a reset decoder) in order to allow for the Rolling Shutter operation mode (see section 2.4.2).

2.3.1. Sensors Array

The SCU test chip contains 12 arrays of 3-T APS pixels. These arrays can be categorized into two clusters according to the measurement that each array is intended for:

- Performance analysis arrays: 8 arrays of 64x64 pixels where shape, size, source follower threshold voltage, and usage of microlenses are modified in order to study how these variations affect the performance of the photosensors.
- Crosstalk analysis arrays: consisting on 4 arrays of 11x11 pixels, where only the central pixel [5,5] is supposed to receive light. To that purpose, all pixels but the central one are covered by metal structures, which serve for light blocking purposes. This arrangement allows studying how the light received by one pixel affects the signal in its near neighbors.

The layout of both types of arrays will be shown in the figures of subsection 2.3.3.2. Table 2.2 summarizes the variations on the different arrays in the chip (array number corresponds to that assigned in figure 2.4), where the layout of the chip is shown.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

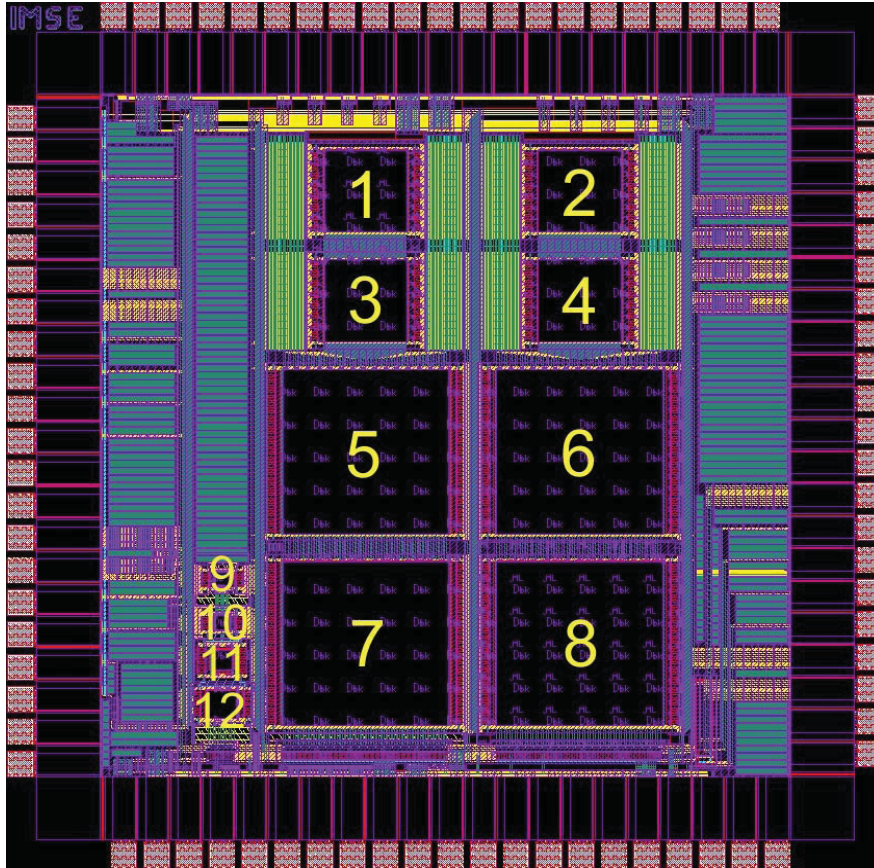


Figure 2.4: Layout of SCU Chip.

Array	Characteristics				
	Array Size	Pixel Pitch (μm)	Shape	Low V_{th}	Microlenses
1	64x64	3.5	Round-like	-	✓
2	64x64	3.5	Round-like	✓	-
3	64x64	3.5	Octagonal	-	-
4	64x64	3.5	Round-like	-	-
5	64x64	7	Octagonal	-	-
6	64x64	7	Round-like	-	-
7	64x64	7	Round-like	✓	-
8	64x64	7	Round-like	-	✓
9	11x11	3.5	Round-like	-	✓
10	11x11	3.5	Round-like	-	-
11	11x11	5	Round-like	-	-
12	11x11	7	Round-like	-	-

Table 2.2: Summary of the arrays included in the chip.

2.3.2. Pixels Circuitry

All the pixels in the chip are 3T-APS. This topology reduces read-out time and allows a non-destructive retrieval of pixel data. Just for illustration purposes, figure 2.5 shows the schematic of the pixels (using 3.3V transistors). Pixel voltage is defined as the voltage drop across the photodiode, which is read out by using a basic source follower. Indeed, only the active transistor of the source follower is included within the pixel (M_1 in figure 2.5). The required biasing current source (a single transistor usually) is either column-wise distributed or shared among all the pixels in the array (our case). The operation of the 3T-Active Pixel Sensor is very well known in the CIS literature and it has already been briefly described in chapter 1. Therefore, we will only provide here some highlights of how this configuration works.

The typical operation of the APS sensor involves three phases, as shown in figure 2.6:

1. Reset phase: During this phase, the photodiode's parasitic capacitor (more precisely, all capacitors connected to photodiode's sensing node, whose aggregation is usually denoted as C_{ph}) is initialized to the so-called reset voltage $V_{ph_{rst}}$ by switching on S_1 . Once this process settles down, the reset switch is turned off. It is straightforward to demonstrate that the starting voltage for the

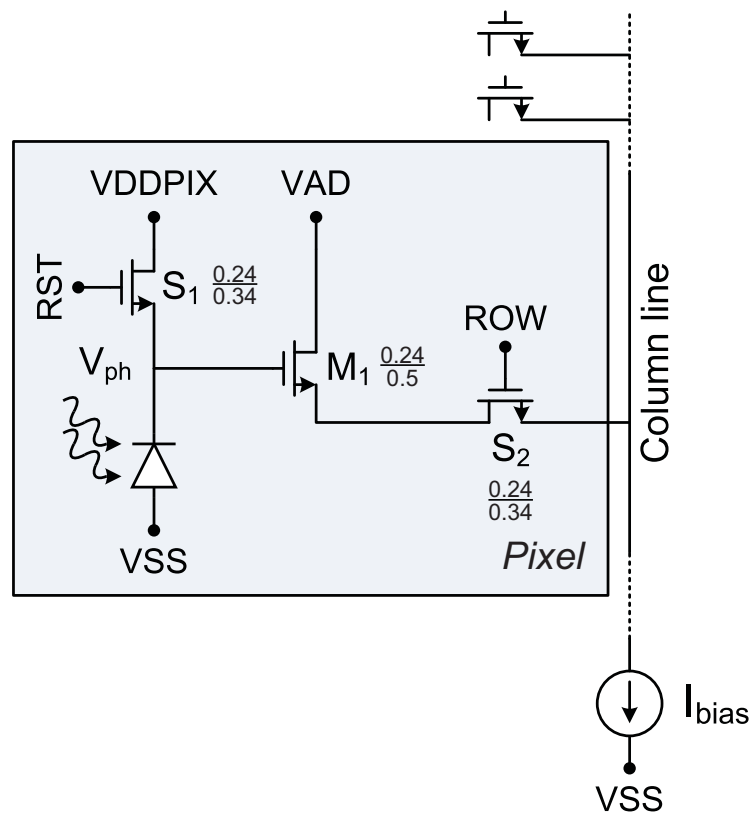


Figure 2.5: Pixel schematic.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

next phase is simply given by:

$$V_{ph_{rst}} = VDDPIX - V_{th} - V_{ft} \pm V_n \quad (2.1)$$

where V_{th} is the threshold voltage of the reset transistor S_1 ¹, V_{ft} is the feedthrough error introduced when switching off the reset switch [67], and V_n denotes the reset noise.

2. Exposition phase: When the reset signal is released, the photogenerated current discharges the integration node V_{ph} . This process, which is inherently non-linear due to the non-linearities of integrating capacitors, and also due to the dependency of the photogenerated current I_{ph} with the volume of the collection region, which is also a function of the voltage drop across the photodiode, is usually simplified to a linear process where the pixel voltage evolves as ²:

$$V_{ph} = V_{ph_{rst}} - \frac{I_{ph}}{C_{ph}} \Delta t \quad (2.2)$$

3. Readout phase: At the end of the exposition phase, whose duration is commonly referred as the exposition time or exposure, the row access transistor (S_2) is enabled and the pixel is connected to the readout circuitry, producing an output that can be expressed as:

$$V_{out} = G[V_{ph}(T_{exp})] \cdot V_{ph}(T_{exp}) - V_{shift}[V_{ph}(T_{exp})] \pm V_{rn} \quad (2.3)$$

where $V_{ph}(T_{exp})$ corresponds to equation 2.2 evaluated for $\Delta t = T_{exp}$, G is the gain of the source follower (usually simplified to one, but which is actually a function of the pixel voltage), V_{shift} is the voltage shift created by the source follower due to the threshold voltage of transistor M_1 (that is also a function of the pixel voltage), and V_{rn} is the so-called readout noise.

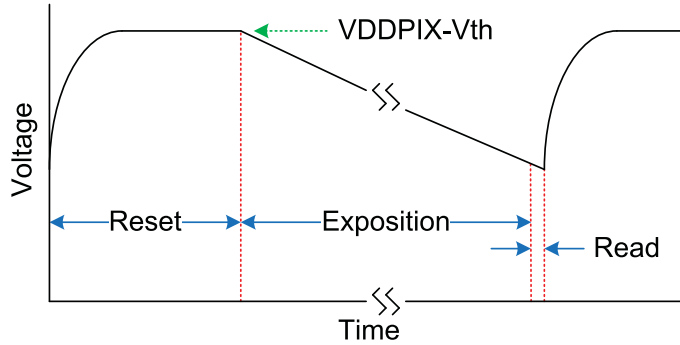


Figure 2.6: Pixel operation phases.

¹Due to the body effect this threshold voltage also depends on $V_{ph_{rst}}$ if the reset switch is an NMOS transistor. This can be eliminated if we use a PMOS switch as reset device [66], however this is not allowed in this CIS technology and, on the other hand, penalizes fill factor heavily in small pitch pixels.

²The effect of dark current and other non-idealities are not considered here for simplification.

2.3.3. Variations in Pixels Parameters

Using the previously described APS architecture, we have introduced variations in the physical parameters of the pixels in different arrays in order to evaluate how these parameters influence the performance of the sensors. Regarding the layout of the active diffusion, we have modified shape and size [68]. Additionally, to get information about the sensitivity gained by the use of microlenses, we have included versions (arrays) of the same pixels with a without microlenses, and the same has been done for the use of low threshold voltage transistors in the source follower. Briefly, the design parameters under variations are:

- Pixel Size (Pitch)
- Active Diffusion Layout
- Threshold Voltage of Source Follower Transistor
- Use of Microlenses

2.3.3.1. Pixel Size

Regarding the performance analysis arrays, two different pixel sizes are considered in the study: $3.5 \times 3.5 \mu\text{m}^2$ and $7 \times 7 \mu\text{m}^2$. Their layouts are shown in figure 2.7. The $3.5 \mu\text{m}$ pitch was selected because it contains the minimum photodiode area that preserves the correct angles and spacing in the connection between the active diffusion and the rest of the circuitry (as defined by UMC layout design rules). Therefore, this $3.5 \mu\text{m}$ option is the minimum achievable pixel pitch for this study and the $7 \mu\text{m}$ pitch pixels decrease the spatial resolution by a factor 4. The latter is expected to provide better fill-factor and sensitivity per area unit. On the other hand, the crosstalk arrays contain 11×11 pixels and selected pitches are $3.5 \mu\text{m}$, $5 \mu\text{m}$ and $7 \mu\text{m}$.

2.3.3.2. Layout of Active Diffusion

Studies recommend [69] to avoid abrupt angles in the layout of the active diffusion because they result in stress and malformation that lead to dark current and noise increment. Therefore, inner and outer angles are limited in the sensing diffusion. Two alternative shapes have been considered in our study:

- Octagonal: which results in active diffusions with 8 abrupt corners (135° degrees) (plus 4 corresponding to the region that connects to the reset transistor, see figure 2.7).
- Round-like: which results in active diffusions with several “softer” corners.

The shapes of the different active diffusions of the so-called “performance evaluation pixels” are shown in figure 2.7, whereas the shapes for the “crosstalk evaluation pixels” are shown in figure 2.8 which only adds the variation of having a $5 \mu\text{m}$ round-like pixel. Pixels having octagonal-shaped photo-sensors only contain 135° ($90^\circ + 45^\circ$) angles in the active diffusion and they are specially intended to

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

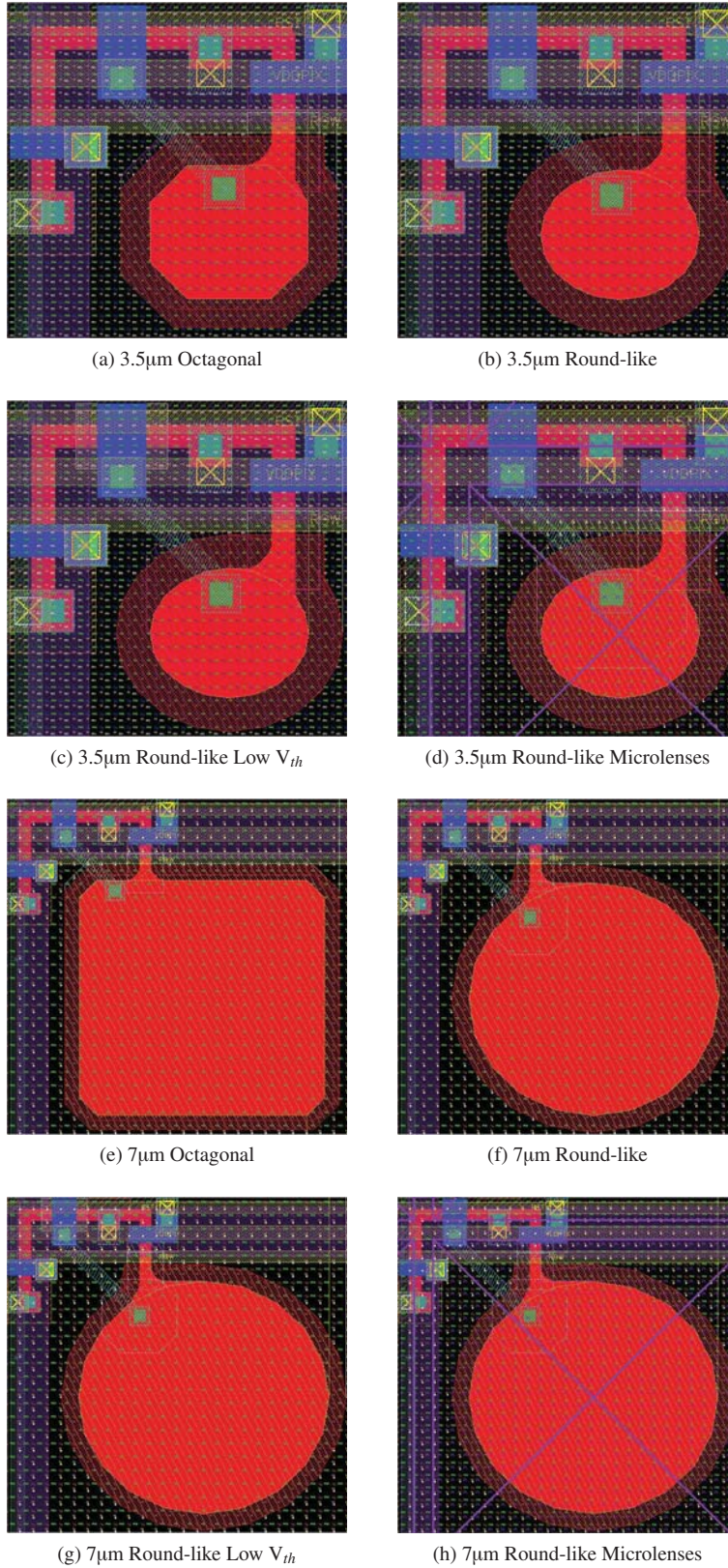


Figure 2.7: Pixels layout of performance analysis arrays.

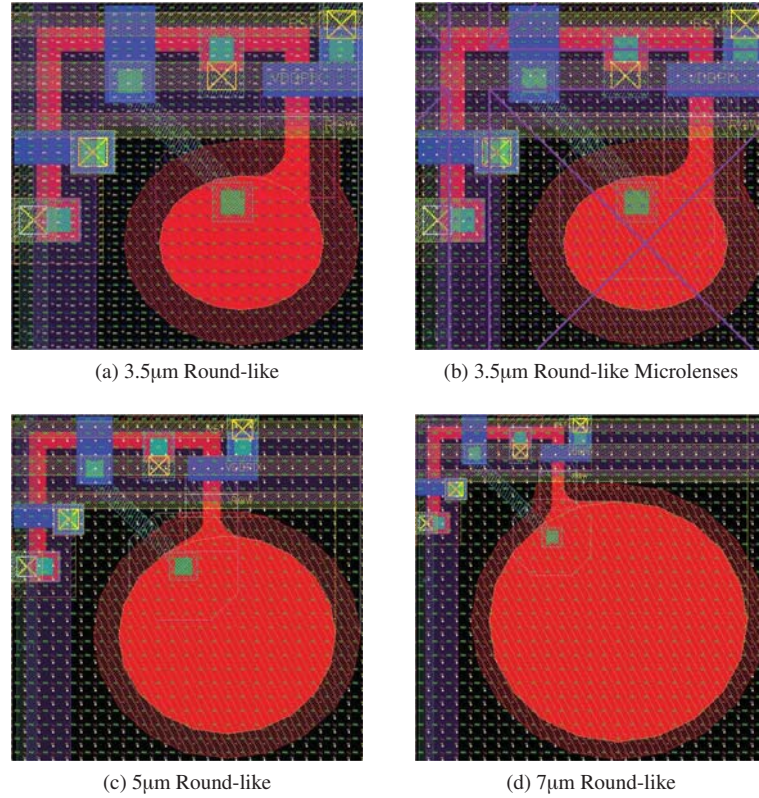


Figure 2.8: Pixels layout of crosstalk analysis arrays.

test the dark current of sensors having the minimum amount of corners (90° angles are not recommended by design rules).

2.3.3.3. Threshold Voltage of Source Follower Transistor

According to their threshold voltage, this technology offers three types of transistors, namely: Zero, Low and Regular V_{th} . The use of low V_{th} transistors in the source follower improves the dynamic range of linear APS, since it directly increases the attainable output swing. Pixels with identical layout have been included with low or regular V_{th} source follower transistors. This variation is implemented in order to determine whether there is any counterpart in the use of the low V_{th} transistors within the pixels. Zero threshold voltage transistors have not been used since they are forbidden within the sensor array (UMC design rule). The layouts in figure 2.7 (c) and (g), differ from (b) and (f) respectively, only in an added layer that defines the source follower transistor as a low threshold device.

2.3.3.4. Use of Microlenses

Microlenses increase sensitivity and minimize crosstalk [37]. In our design, microlenses are centered in the middle of the sensing diffusion and they focus the light to this point. It allows rays that otherwise would not hit the sensor area to arrive to the sensor surface and be detected. Clearly, this is especially important under no perpendicular incidence of light situations. The overall effect is that microlenses increase the amount of light power at the sensor surface (as compared to the same situation without microlenses) and, consequently, the number of photogenerated carriers.

In order to study the performance of microlenses in this technology, pixels with the same layout with and without microlenses are included in the chip. The main purpose is to characterize how they improve sensitivity and crosstalk. Figure 2.9 shows a simplified illustration of the effect of the microlenses over vertical and lateral incident rays of light. Figure 2.10 illustrates, in a 3D View, the position of the microlens over the metal aperture of the pixel. Dimensions in this figure illustrate microlenses effect in small pitch pixels, where the size of the aperture is comparable to the height of different metal layers and passivations.

The layouts of the pixels in figure 2.7 (d) and (h) are identical to (b) and (f), respectively, with the exception of the microlens layer definition, which appears as a purple crossed square over the photodiode diffusion. The microlens center has been located in the center of the photodiode diffusion, because the microlens will focus the light around this location. As the separation between microlenses is fixed, they are not completely enclosed by the pixel layout. In consequence, the pixel layout shows part of the microlenses of other pixels. The layout of the crosstalk pixel with microlens is shown in figure 2.8 (b).

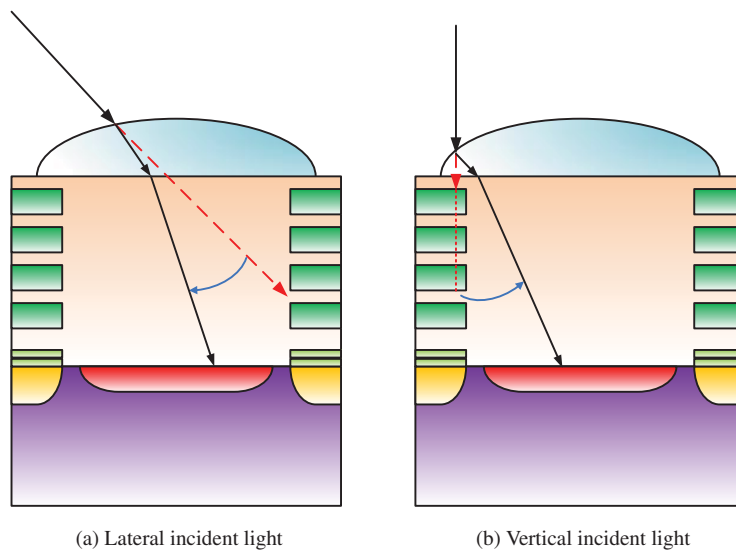


Figure 2.9: Microlenses effect on light.

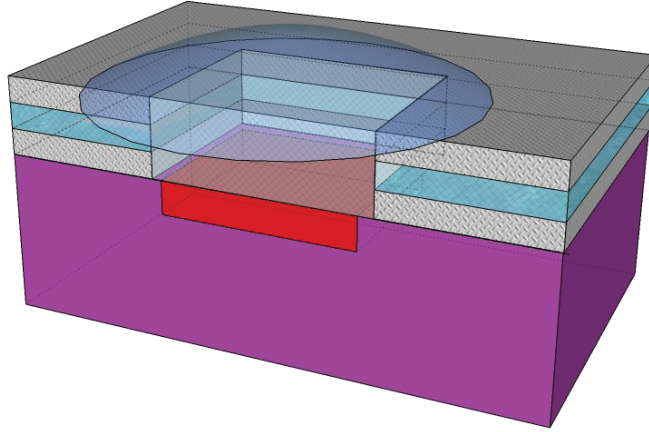


Figure 2.10: Microlenses 3D illustration.

2.3.4. Analog Output Buffer

This block includes the biasing circuitry of the pixel source follower plus an additional PMOS source follower. Figure 2.11 shows the schematic of the buffer. Transistor P_5 partially compensates the level shifting introduced by the pixel NMOS source follower and connects to the output pad. The NMOS transistors N_1 and N_2 complete the bias current circuitry of the pixel source follower connected through the input of the output buffer. The biasing is performed through an NMOS mirror with a ratio 93 to 1. An externally applied current of $111.34\mu\text{A}$ is scaled down approximately to $1.2\mu\text{A}$. The biasing of the PMOS section is accomplished by means of a PMOS current mirror, composed by transistors P_1 and P_2 . In this case, a cascode arrangement is added between the mirror and the source follower transistor in order to increase the output resistance of the mirror (transistors P_3 and P_4) [70]. The bias current of the PMOS section is $44.14\mu\text{A}$. These biasing currents can be approximately obtained by connecting a $10\text{k}\Omega$ resistor between 1.8V and $IBIASN$ through and a $50\text{k}\Omega$ resistor between $IBIASP$ and ground.

2.4. Control of Operation

The realization of any measurement on the SCU chip requires applying control signals in the proper order. Since the SCU chip does not include any internal control unit nor sequencer, everything is left to the user, which needs to provide the signals, timing, and sequences correctly. Table 2.3 summarizes the digital and analog control signals of the chip. Table 2.4 shows the addresses and enable signals of the different arrays.

2.4.1. Image Capture Operation

According to the typical 3T APS operation mode, the first operation is to reset the photodiode so that its associated capacitor is charged to a reference voltage V_{phrst} , as defined in equation 2.1. $VDDPIX$ is

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

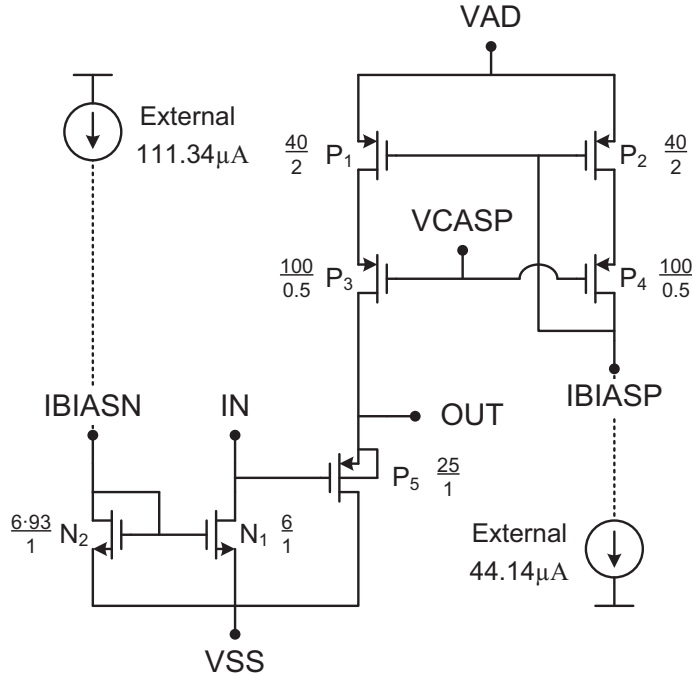


Figure 2.11: Schematic of the Analog Output Buffer.

Group	Pin Name	Range/Value	Type	Functionality
Address buses	RST<7:0>	Digital	Input	Selection of pixel row to be reset
	ROW<7:0>	Digital	Input	Selection of pixel row
	SELECT<7:0>	Digital	Input	Selection of pixel column
Control	RST_EN	Digital	Input	Reset enable in performance arrays
	RST_EN.CROSS	Digital	Input	Reset enable in crosstalk arrays
	ROW_EN	Digital	Input	Row enable in performance arrays
	ROW_EN.CROSS	Digital	Input	Row enable in crosstalk arrays
	SELECT_EN	Digital	Input	Column enable
Bias	VCASP	1.8V	Input	Output buffer cascode biasing
	IBIASP	44.14µm	Input	Output buffer bias current
	IBIASN	111.34µm	Input	Pixels source follower bias current
Supply and Ground	VDDPIX	3.3-0V	Power	Pixel reset voltage
	VAD	3.3V	Power	Power supply of the analog circuitry
	VDD33_DIG	3.3V	Power	Power supply of the 3.3V digital circuitry
	VDD18	1.8V	Power	Power supply of the 1.8V digital circuitry
	VSS	0V	Power	Ground supply
Data	OUT	1.019-2.714V	Output	All arrays output Used to read images data

Table 2.3: Analog and Digital Control SCU Pins.

Chosen Array	ROW<7:0> RST<7:0>	ROW_EN RST_EN	ROW_EN_CROSS RST_EN_CROSS	SELECT<7:0>	SELECT_EN
3.5 round-like μ lens	0:63	1	0	11:74	1
3.5 round-like lvt	0:63	1	0	75:138	1
3.5 octagonal	64:127	1	0	11:74	1
3.5 round-like	64:127	1	0	75:138	1
7 octagonal	128:191	1	0	11:74	1
7 round-like	128:191	1	0	75:138	1
7 round-like lvt	192:255	1	0	11:74	1
7 round-like μ lens	192:255	1	0	75:138	1
3.5 round-like crosstalk	64:74	0	1	0:10	1
3.5 round-like μ lens crosstalk	0:10	0	1	0:10	1
5 round-like crosstalk	128:138	0	1	0:10	1
7 round-like crosstalk	192:202	0	1	0:10	1

Table 2.4: Row and column addressing of the arrays.

configured via an external analog input pad. The reset action cannot be done pixel by pixel in this design; instead, the whole row selected by RST<7:0> is initialized when RST_EN is active. Once the reset signal is released (RST_EN is 0), this row starts the exposition phase where the photogenerated current will discharge the integration capacitance of these pixels. At the end of the exposition phase, pixel voltage must be read out. First, the row must be selected by ROW<7:0> and then activated by ROW_EN. Once the row switch is connected, the column switch will connect the pixel to the output buffer. Column selection is done through an address bus SELECT<7:0> and an enable signal SELECT_EN. Once the output node voltage has stabilized, the read out voltage is approximately¹ given by:

$$V_{pixel} = V_{ph_{rst}} - \frac{I_{pix}}{C_{pix}} T_{exp} - V_{th_n} = V_{ph_{rst}} - \frac{I_{ph} + I_{dark}}{C_{pix}} T_{exp} - V_{th_n} \quad (2.4)$$

$$V_{out} = V_{pixel} + V_{th_p} = V_{ph_{rst}} - \frac{I_{ph} + I_{dark}}{C_{pix}} T_{exp} - V_{th_n} + V_{th_p} \quad (2.5)$$

where $V_{ph_{rst}}$ is the voltage in the integration node just after reset, as expressed in equation 2.1, I_{ph} is the photogenerated current, I_{dark} is the dark current, C_{pix} is the overall capacitance of the integration node (due to the photodiode capacitance C_{ph} but also including the capacitances of the rest of circuitry connected to that node), T_{exp} is the exposition time, V_{th_n} is the threshold voltage of the source follower transistor of the pixel and V_{th_p} is the threshold voltage of the source follower of the output buffer.

Figure 2.12 shows the signals sequence for some sample pixels (72,132), (72,133), (72,134) and (72,135) (corresponding to the 3.5 μ m round-like pixel array) and figure 2.13 shows the same sequence for the pixels (0,0), (0,1), (0,2) and (0,3) (that are in the 3.5 μ m round-like μ lens crosstalk array).

An electrical simulation of the acquisition of the pixels (72,135) and (72,134), in this order, is

¹Assuming that the gain of the source followers is 1, and independent of the input voltage.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

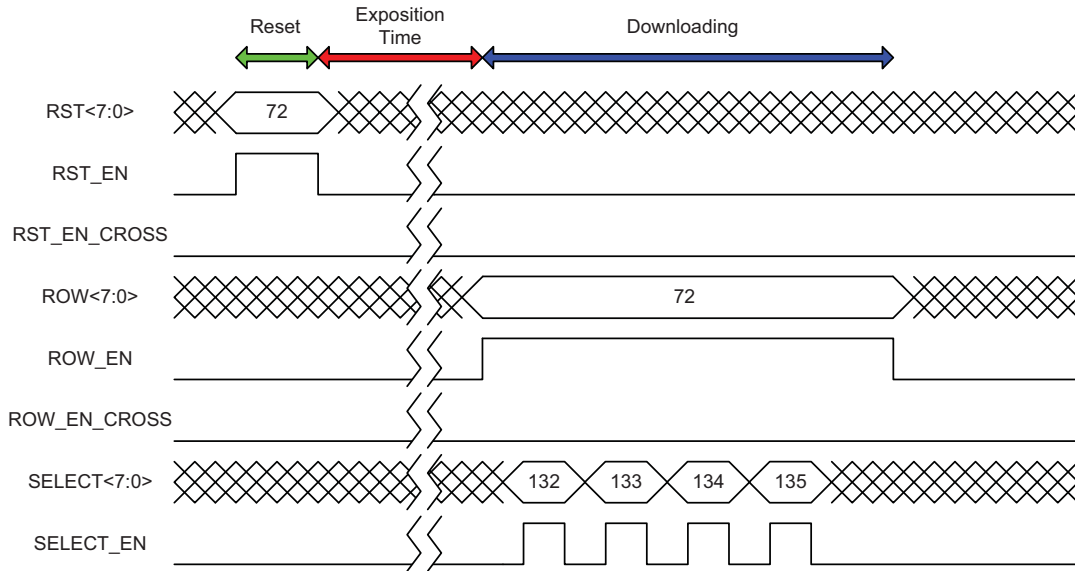


Figure 2.12: Control timing illustrating data capture in performance arrays.

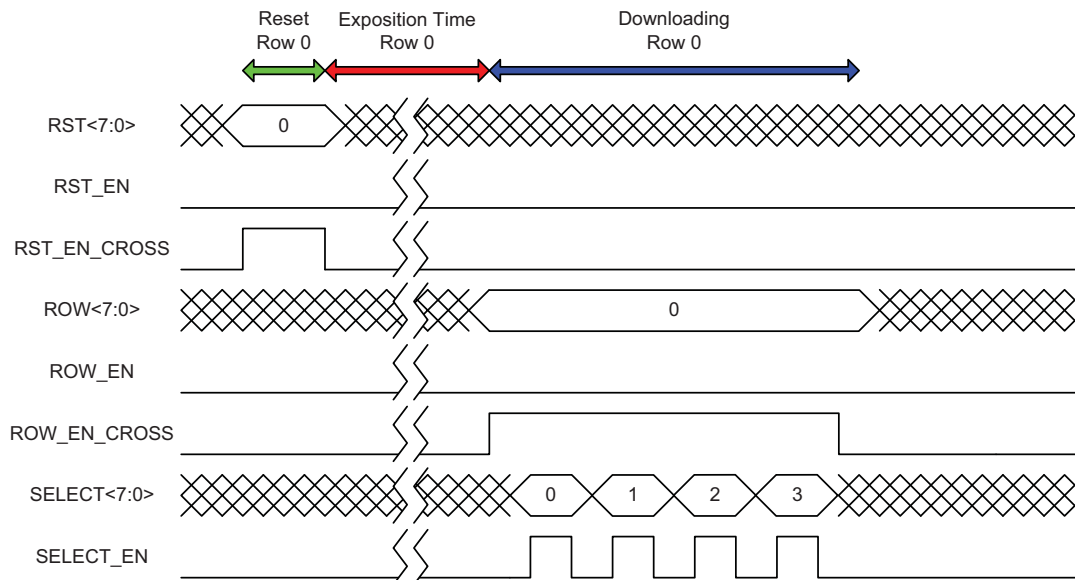


Figure 2.13: Control timing illustrating data capture in crosstalk arrays.

shown in figure 2.14, using 500 pA and 1 pA for the photogenerated currents¹ of first and second pixel, respectively.

2.4.2. Operation Modes

SCU chip includes a rolling shutter mode, which minimizes to a row the possible gradient due to sequential readout by reducing the exposition time difference to a maximum of the readout time of one row (between consecutive rows). The rolling shutter operation requires adding an additional row decoder for the RST signal. A typical sequencing for 3 rows would be:

1. Reset row 0.
2. Reset row 1 \Rightarrow while exposing row 0.
3. Reset row 2 \Rightarrow while exposing row 0 and 1.
4. Exposing row 0, 1 and 2.
5. Read row 0 \Rightarrow while exposing row 1 and 2.

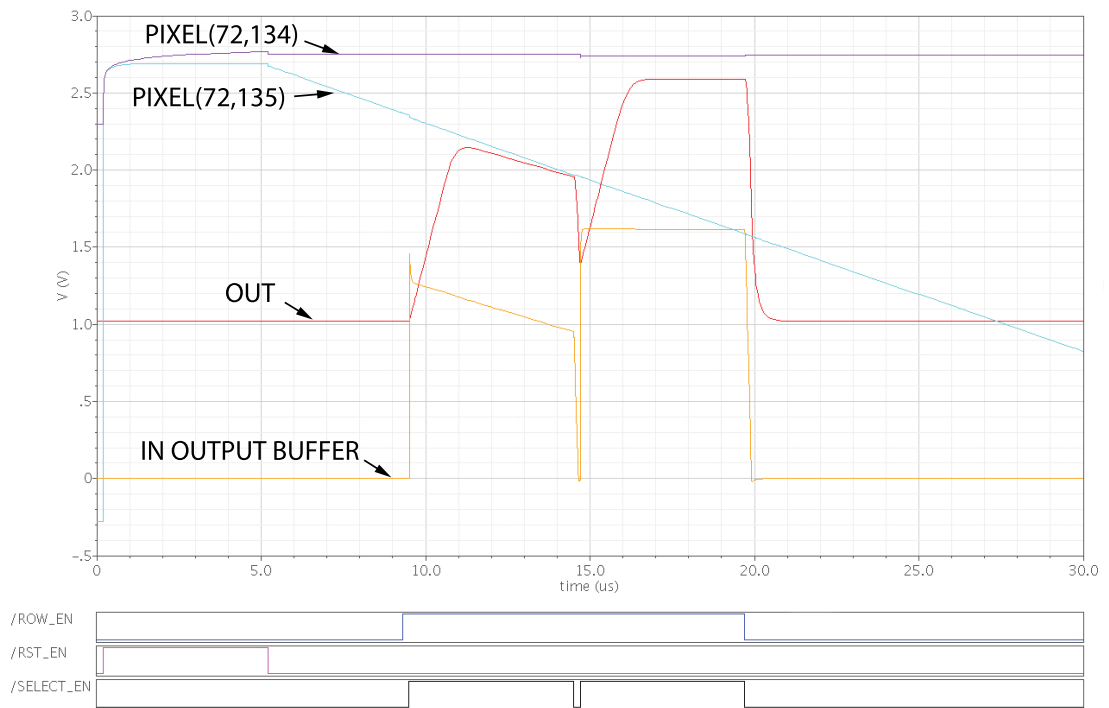


Figure 2.14: Electrical simulation of pixels data retrieval.

¹These photogenerated currents are too large for the photodiode sizes in this design. The expected photocurrents, under natural illumination conditions are between 1 fA and 500 fA. These large photocurrents are used in order to obtain short exposition times and ease the visualization of the Reset-Integration-Readout phases in a single plot.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

6. Read row 1 \Rightarrow while exposing row 2.
7. Read row 2.

Alternatively, for precise pixel evaluation, it is also possible to employ the so-called pixel-by-pixel mode, where we execute the Reset-Exposition-Readout sequence for every single pixel. Obviously, measuring the chip using this process takes longer. However, it provides the more accurate results as it is guaranteed that all pixels have been integrating the photogenerated current for the same time. Nevertheless, it possess strong requirements on the stability of the light source.

2.4.3. Timing Specifications

The user must provide the control signals for the SCU chip guaranteeing that the temporal specifications and sequencing in table 2.5 and figure 2.15, respectively, are met.

First, the row must be reset for a minimum time $t_{r_{min}}$. The row and column buses can be changed at any time while enable signals are off, however their activation through the enable signals should meet

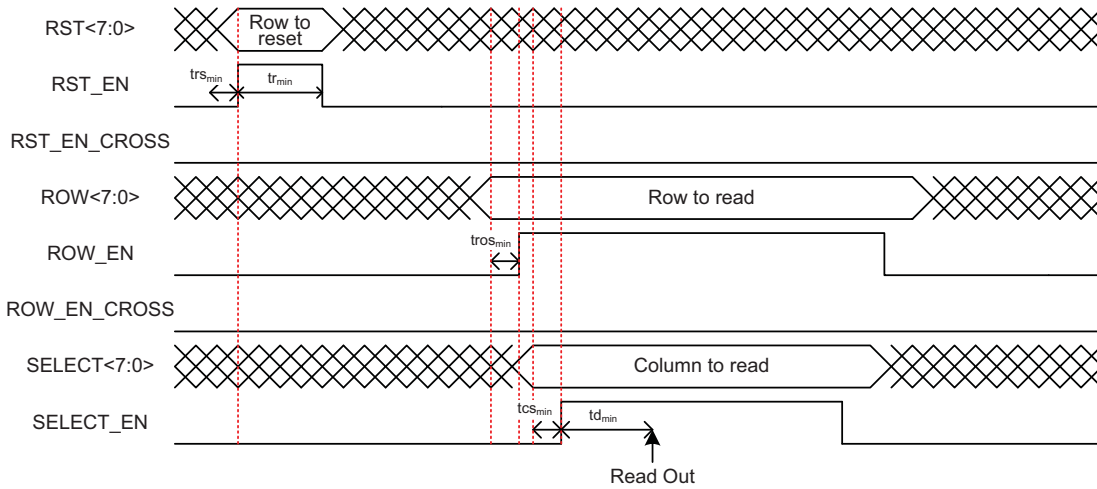


Figure 2.15: Timing illustration.

Time	Description	Value
$t_{rs_{min}}$	minimum reset bus establishment time	100 ns
$t_{r_{min}}$	minimum reset time (1% error)	1.22 μ s
$t_{ros_{min}}$	minimum row bus establishment time	100 ns
$t_{cs_{min}}$	minimum column bus establishment time	100 ns
$t_{d_{min}}$	minimum read-out waiting time ($C_{load} = 30$ pF)	3.75 μ s

Table 2.5: Timing description.

$tros_{min}$ and tcs_{min} requirements. After the read out path is enabled, the output signal would be available for read-out after td_{min} for a 30 pF capacitive load.

2.5. Measurements

The electro-optical characterization of the SCU chip has also required the design of a dedicated 4-layers Printed Circuit Board (PCB) including interfacing, biasing circuitry and an FPGA as controller. Basically, the purpose of this system is to generate (precisely) the required signal sequences to acquire and digitize pixel voltages, and to store them in a dedicated SRAM block. When the measurement is completed, the results are sent through an USB connection to a computer for further analysis.

The test board has been placed on an optical table for photometric characterization of the SCU chip. On top of this table, different devices have been included by means of docking hardware: a controlled light source, a beam splitter, a light power meter, bandpass optic filters, depending on the measurement to be performed.

2.5.1. Experimental Setup

The test board design is centered around an FPGA that generates the proper signals in order to retrieve the analog pixel data, digitize, store and finally send them to the computer. The block diagram of the board is shown in figure 2.16 whereas devices and their functions are listed in table 2.6. Figure 2.17 shows, for illustration purposes, the schematic of the board. Figure 2.18 shows a photograph of the board ready to be used.

A typical operation of the system begins with the reception of the configuration from the PC through either the USB or the RS232 link. This configuration determines, among many other parameters, the different timings, the mode of operation of the SCU (rolling shutter or pixel by pixel measurement) and configure the $VDDPIX$ voltage. SCU signals are generated by the FPGA for reset addressing, waiting for exposition and downloading data. Simultaneously, every time a new analog data is available, the FPGA produces the control signals for the ADC and store the digital result into a SRAM. Once the test is over, the FPGA sends the results to the computer to be analyzed. The whole process is integrated into a Matlab[®] environment including the control of electro-optical equipment (such as power unit for the light source), and retrieving measurements from the calibrated sensor (such that at the very end [after a hard coding work] everything consists on executing a single m-file with the right parameters). This is very important since it allows for automatically executing measurements just using loops around this file.

This board, in combination with the devices of the optical setup, has been mounted on top of an optical table RP Reliance[™] 46-8 Sealed Hole Table of Newport Corporation [71]. It offers broadband damping, static rigidity and thermal stability capabilities. This 8 inch thick optical table has 4 foot width, 6 foot length, and 1/4-20 holes on a 1 inch grid. Newport opto-mechanics have been used in order to insert devices on the table, and to isolate the light path to the chip using tubes. A photograph of the setup is shown in figure 2.19.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

The controlled stable light source is generated by a 250 Watt 24V DC Quartz Tungsten Halogen (QTH) Newport 6334 lamp. It is a visible and near infrared source with smooth spectral curve and stable output. Figure 2.20 shows its spectral irradiance. It produces an approximate flux of 10000 Lumens at 3400K. It contains a doped tungsten filament inside a quartz envelope filled with rare gas and a small amount of halogen. This lamp is mounted inside of a 66884 Research QTH Newport Lamp Housing. The housing includes a F/1 Two Element Fused Silica condenser for a 33 mm diameter collimated output beam. An included rear reflector collects the lamp's back radiation. A power regulated fan cools the lamp and housing. This lamp housing is supplied by a stable power supply, a 69931 Newport Radiometric Power Supply. This is a highly regulated source of constant current or constant power, which includes an RS-232 interface to be controlled by a PC.

Depending on the measurement to be performed, the white stable light generated by the source needs to be altered in a very accurate way in either wavelength or power. In order to produce light of a narrow wavelength, bandpass filters are applied in light path. This is used to measure the spectral response [72] of the SCU. The filters are Andover Corporation Standard Bandpass Filters. These filters have Central Wavelength (CWL) from 400 to 750nm (50nm steps) and Full Width at Half Maximum (FWHM) of 10nm in a 50mm diameter, (400FS1050 - 750FS1050).

The light power received in the SCU chip is measured by combined use of a calibrated photodetector and an optical power meter. The detector is a Newport 918-SL Low-Power Silicon Photodetector, which is calibrated in the range [400 - 1100]nm. Calibration setup is electronically stored inside the detector's head for use when connected to the used 1930-C Power Meter.

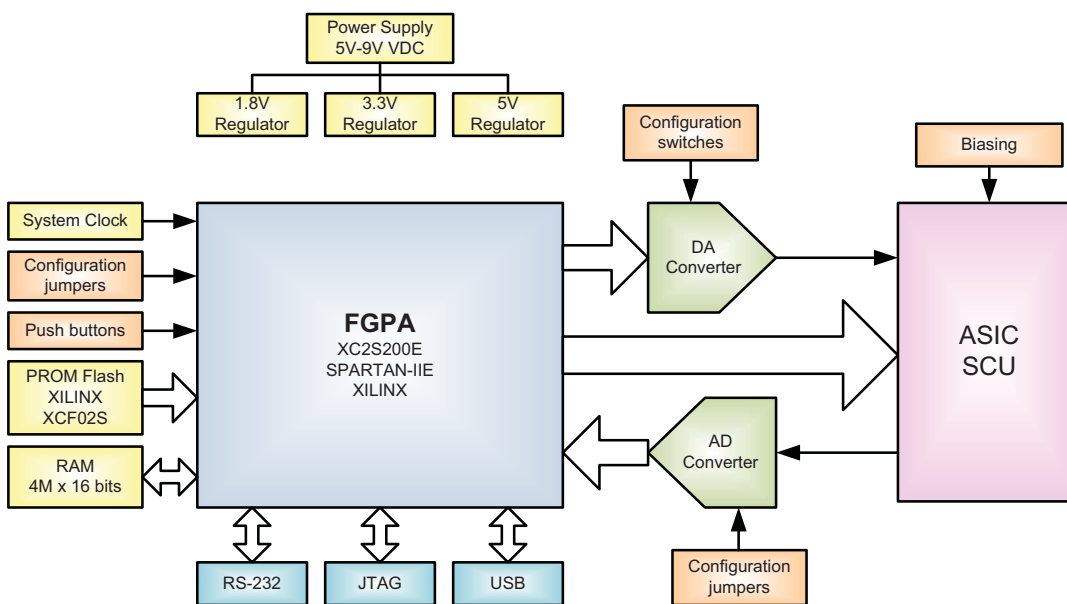


Figure 2.16: SCU PCB test board blocks.

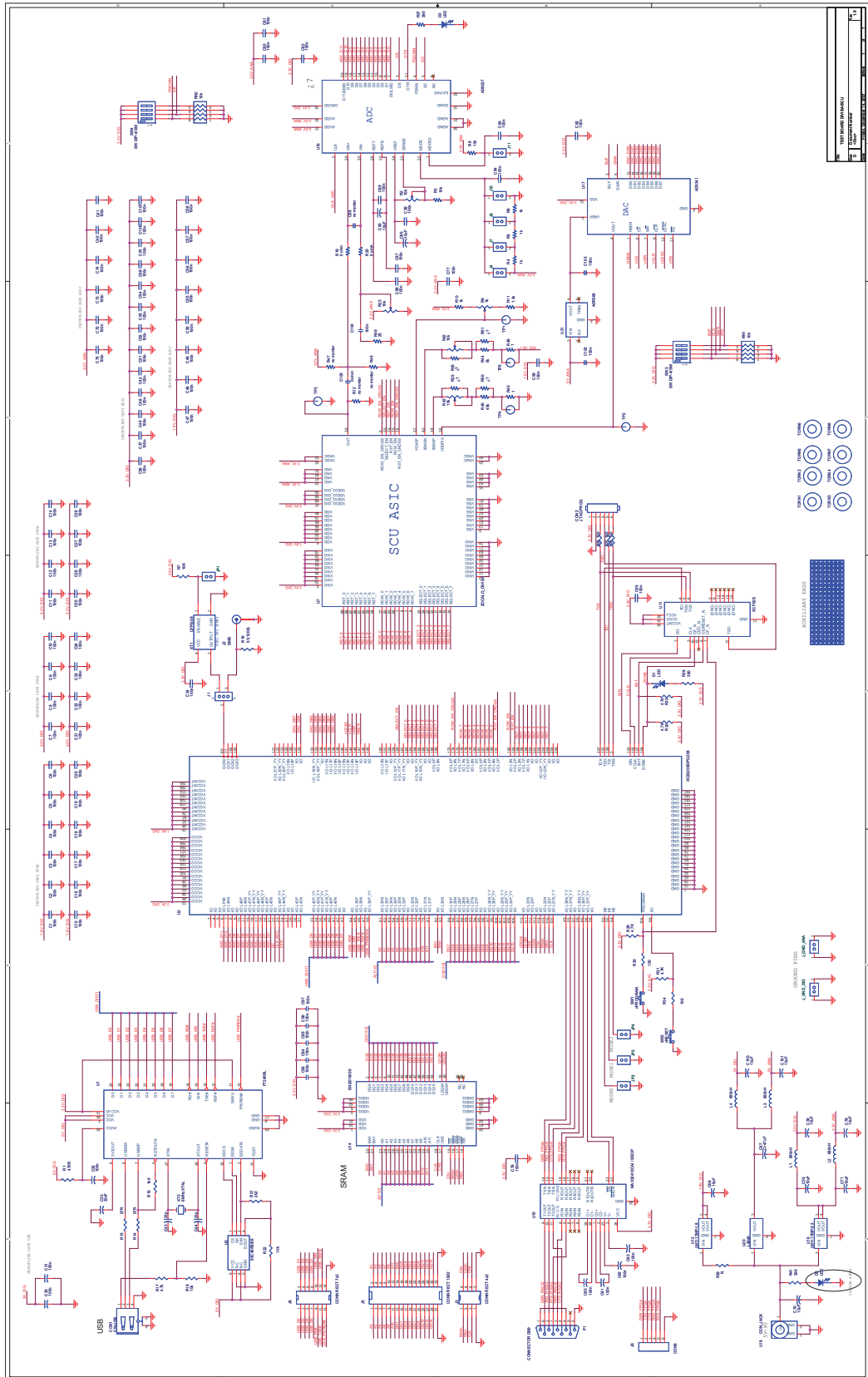


Figure 2.17: SCU PCB test board schematic.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

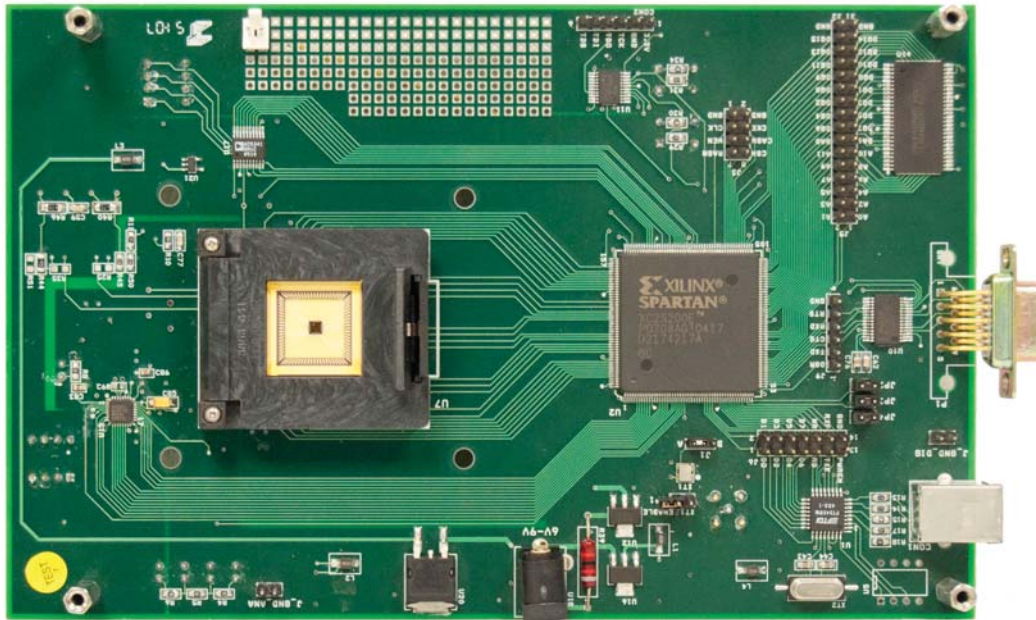


Figure 2.18: SCU PCB test board.

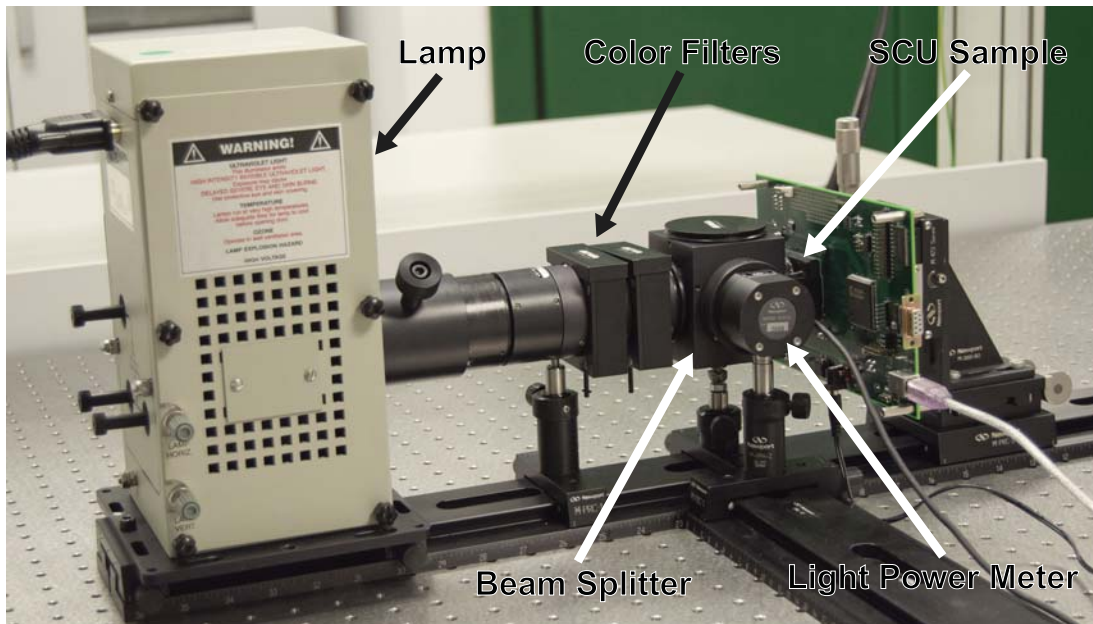


Figure 2.19: SCU measurement setup.

Block	Components	Function
CMOS Image Sensor Chip	SCU ASIC Chip in CIS 0.18 μ m 4M/2P technology	Sensor under test
FPGA	Xilinx Spartan II-E XC2S200E package PQ208	Central Control System
PROM Flash memory	Xilinx XCF02S 2 Mbit	Store FPGA program
RAM	Integrated Silicon Solution Synchronous Dynamic RAM IS42S16400D 1Mx16bits.x4 banks	Store images from SCU
Digital to Analog Converter	Analog Devices AD5341B	Generate $VDDPIX$ Voltage
Analog to Digital Converter	Analog Devices AD9327BCPZ-65	Digitize pixel data
JTAG port	6 pins JTAG connector	Port to program the FPGA from PC
RS232 port	MAXIM MAX3241	Serial communication with PC
USB port	FTDI FT245BL	Serial communication with PC
ASIC biasing circuit	Potentiometers and resistors	Generate $VCASP$ Voltage, $IBIASP$ and $IBIASN$ currents
Clock generator	IQD Frequency Products CFPS-39 crystal 50Mhz oscillator	System clock
Voltage supply generators	National Semiconductor LM1117MP-3.3 and LM1117MP-1.8, ST Microelectronics L4940D2T5 5V	Generate analog and digital supply voltages
Configuration jumpers	3 jumpers for FPGA, 5 jumpers for ADC, 2 jumpers for clock	Configure FPGA mode of operation, ADC and clock
Configuration switches	2 switches for ADC, 4 switches for DAC	Configure DAC and ADC operation
Manual user controls	2 push buttons	PROGRAM and RESET FPGA buttons
State leds	2 leds	Inform of active power supply and successful FPGA programming

Table 2.6: SCU PCB board devices.

2.5.2. Spectral Sensitivity

Spectral sensitivity has been measured for the range from 400nm to 900nm, in 50nm steps. Results are shown in figure 2.21 where the corresponding dark current contribution has been subtracted¹. Results are the average values of the array of pixels taking a central region of interest (ROI), which discards 5 rows and columns on each side, since those might be affected by the surrounding circuitry. The discharging rate is estimated from a least square linear polynomial fit of 20 measurement points equally distributed in time.

The first important highlight is that it is clearly noticeable that the round-like diffusion layout has better sensitivity for both 3.5 and 7 μ m pixel pitches. Although the aim of using round-like pixels was to diminish dark current due to the less stressing angles, we also find that it increases sensitivity. The reason for this seems to be based on the fact that round and octagonal pixels have the same aperture in metals but the round-like diffusions have less area than octagonal ones, therefore their associated parasitic capacitance at the sensing node is smaller. A smaller sensing area produces a smaller volume of the collection region underneath the diffusion. However, it seems that lateral collection and carrier diffusion into the collection region is significant. Consequently, in this case, the smaller capacitance,

¹The measurements are the average data of the interpolated slope of discharge minus the dark current contribution and divided by the light power applied at the different wavelengths.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

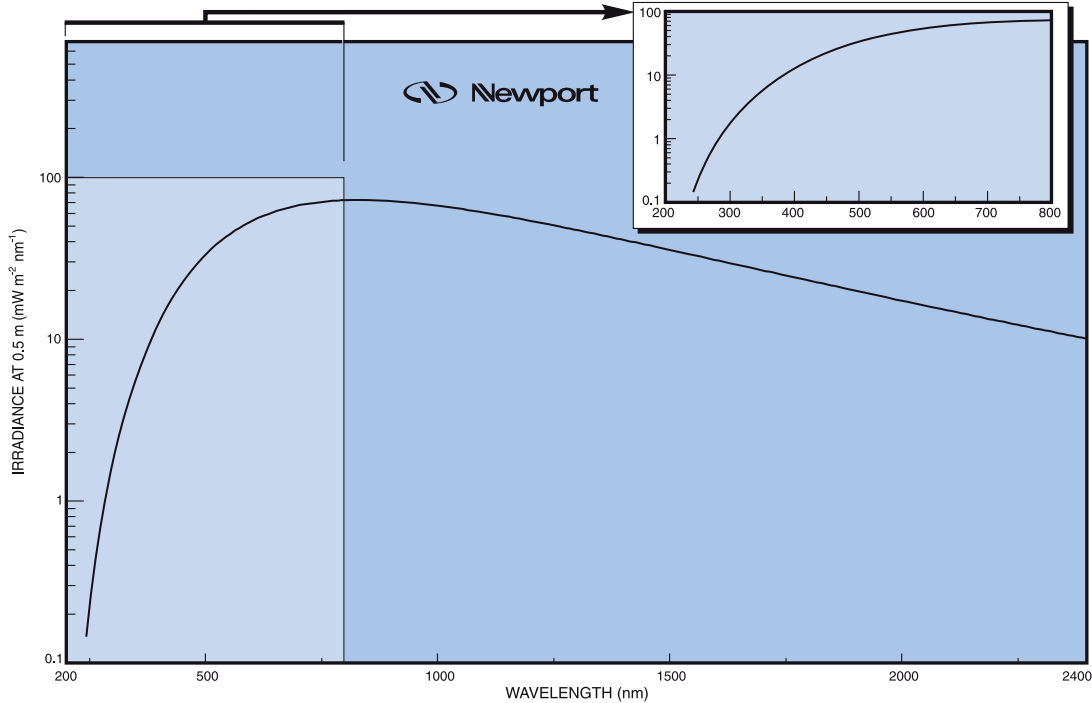


Figure 2.20: Newport 6334 lamp spectral irradiance.

which is direct consequence of the smaller area, produces a significant sensitivity improvement.

We can also observe that pixels with low threshold voltage transistors in the source follower have a slightly better sensitivity than the same pixels with regular transistors, something that adds to the positive effect of the expansion of the output voltage swing in about 0.4V. This effect is simply explained by the fact that low- V_{th} transistors have smaller gate capacitance than regular ones. Therefore, as it happened in the previous case, the overall capacitance at the integration node is smaller for the same collection region, thus producing a gain in the sensitivity of the sensor. The expansion of the output swing due to the low threshold voltage source follower can be seen in the static conversion gain of the signal path from $VDDPIX$ to the output pad, which is shown in figure 2.22.

Finally, we have also compared the results at 550nm, shown in table 2.7, with the results of a similar set of measurements over a test chip previously designed by our group, which employed the standard corresponding technology (UMC CMOS 0.18 μ m, but using 1.8V transistors[73]). Measurements in the standard technology resulted in a sensitivity of 0.1655 $\frac{V}{lux \cdot s}$ for 5 μ m pitch pixels and 0.0958 $\frac{V}{lux \cdot s}$ for 3 μ m pitch pixels. Hence, using the CIS technology increases sensitivity by about ten times at this wavelength.

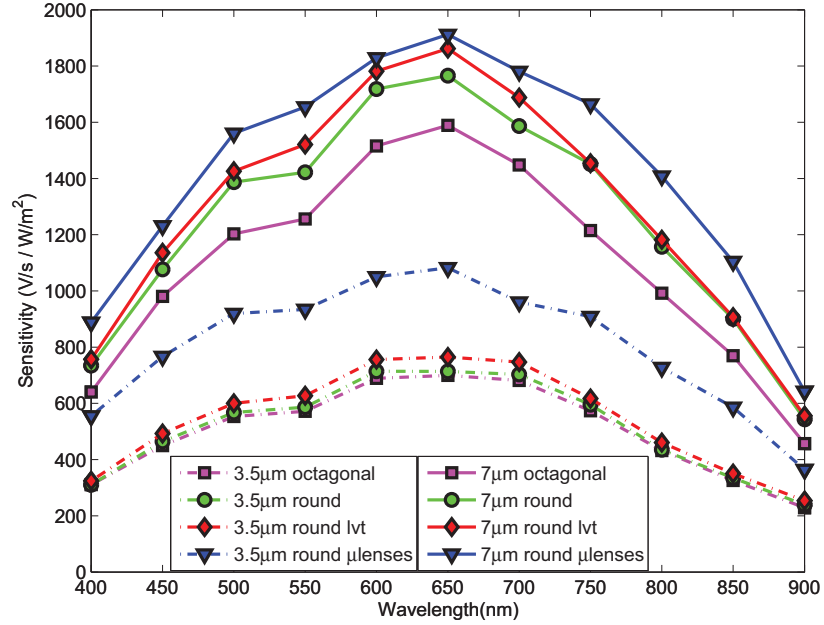


Figure 2.21: Spectral Response.

Array	Sensitivity($\frac{V}{\mu\text{lux}\cdot\text{s}}$)
3.5 round μ lens	1.37
3.5 round lvt	0.918
3.5 octagonal	0.837
3.5 round	0.858
7 round octagonal	1.84
7 round	2.08
7 round lvt	2.23
7 round μ lens	2.42

Table 2.7: Sensitivities at 550nm.

2.5.3. Dark Current

A very important topic in image sensors is the dark current effect, as this effect determines the minimum amount of detectable light power [74]. Dark current measurements have been taken in complete darkness at 20 different exposition times, ranging from 300ms to 5s. Repeated measurements are the average values of every array of pixels, using the same Region of Interest (ROI) as in the sensitivity measurements. Figure 2.23 shows the evolution of the pixel voltage (average) in total darkness, whereas table 2.8 summarizes the results for all the arrays in the chip.

First of all, it can be observed that, the average dark signal, which is mainly produced by the effect of dark current and integrating capacitor, expressed as $\frac{I_{\text{dark}}}{C_{\text{pix}}}$, is about $42.6 \frac{\text{mV}}{\text{s}}$ for the $3.5\mu\text{m}$ pitch pixels and around $16.9 \frac{\text{mV}}{\text{s}}$ for the $7\mu\text{m}$ pitch pixels. This certainly expresses a substantial enhancement considering the previous measurements made in the standard CMOS $0.18\mu\text{m}$ UMC technology, producing results of $340 \frac{\text{mV}}{\text{s}}$ for $5\mu\text{m}$ pitch pixels and $1.22 \frac{\text{V}}{\text{s}}$ for $3\mu\text{m}$ [73].

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

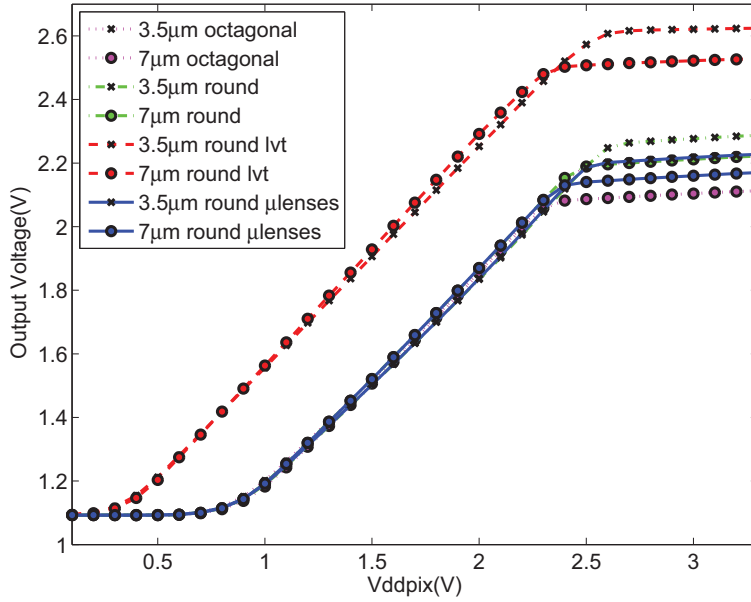


Figure 2.22: Conversion gain curves from $VDDPIX$ to out pad for all the arrays.

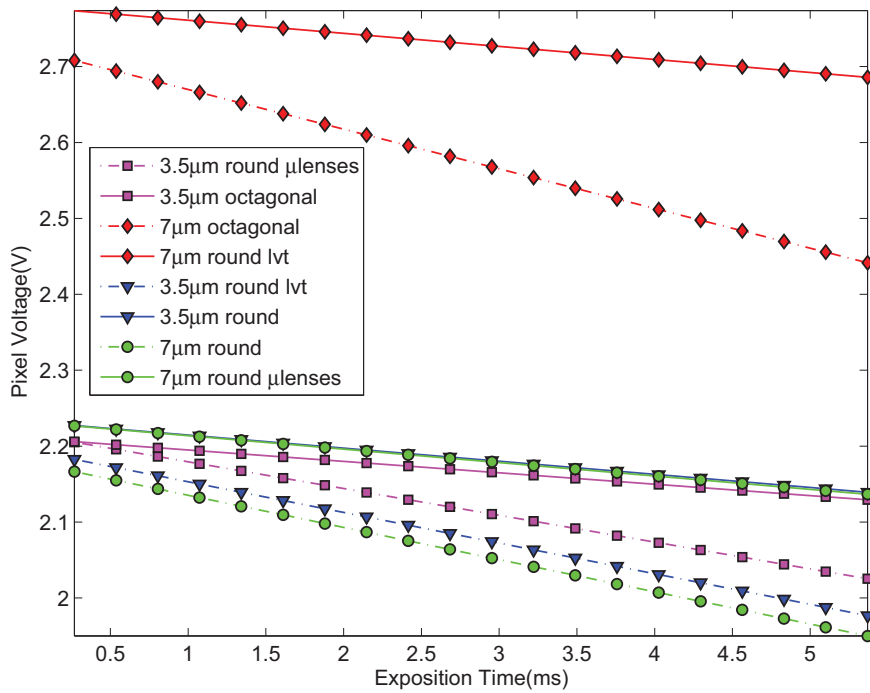


Figure 2.23: Discharge in darkness.

Array	Discharging rate in darkness($\frac{mV}{s}$)
3.5 round μ lens	40.358
3.5 round lvt	52.271
3.5 octagonal	35.28
3.5 round	42.408
7 round octagonal	15.018
7 round	17.654
7 round lvt	17.206
7 round μ lens	17.654

Table 2.8: Discharge slope in darkness.

Secondly, it is noticeable that the octagonal pixels, in both sizes, feature the lower dark signal. Although this seems to be contradictory to our assumptions about dark current and abrupt angles, we cannot conclude the opposite at this point, since we need to bear in mind that what we are measuring here is the combined effect of dark current and parasitic capacitance, and we need to check SNR measurements (combining sensitivity and dark signal) in order to extract a meaningful conclusion.

Thirdly, concerning the low threshold voltage in the source follower, we have to consider that these pixels exhibit a higher leakage due to the special doping profile in the channel and the thinner gate oxide [75]. Observing the result for the $3.5\mu\text{m}$ pitch pixels, we easily identify that it low- V_{th} pixels show the highest dark signal for this pitch. However, if we examine the $7\mu\text{m}$ pitch pixels, the relative difference is only 2%, a possible reason for that is the fact that in the $3.5\mu\text{m}$ pitch pixels, the smaller gate capacitance of the source follower reduces significantly the global capacitance at the integration node (as confirmed by sensitivity measurements) whereas this is not so important in the $7\mu\text{m}$ pitch pixels, where the photodiode's parasitic capacitance is largely dominant.

2.5.4. Dynamic Range

Figure 2.24 shows the Signal to Noise Ratio (here noise includes both readout noise and dark current effects). Obviously, dark current discharge will be the most remarkable contribution over the time (as readout noise power is not expected to increase with the exposition time). For the $7\mu\text{m}$ pitch pixels, it is observed that the low V_{th} exhibits the best SNR, as predictions forecast. The main reason for this is the 0.4V extended output swing due to the low- V_{th} source follower. Second position is occupied by the octagonal pixels, due to their lower dark signal, with the rest of options exhibiting very similar figures.

The $3.5\mu\text{m}$ pitch pixels set exhibits a slightly different behavior. In this case, the octagonal option shows the best performance, due to its lowest dark contribution with the low- V_{th} getting the second position in this figure of merit. In this case, the expansion of the output swing is not sufficient to counterpart the leakage of the low- V_{th} transistor. As in the $7\mu\text{m}$ pitch pixels, the rest of options exhibit very similar SNR.

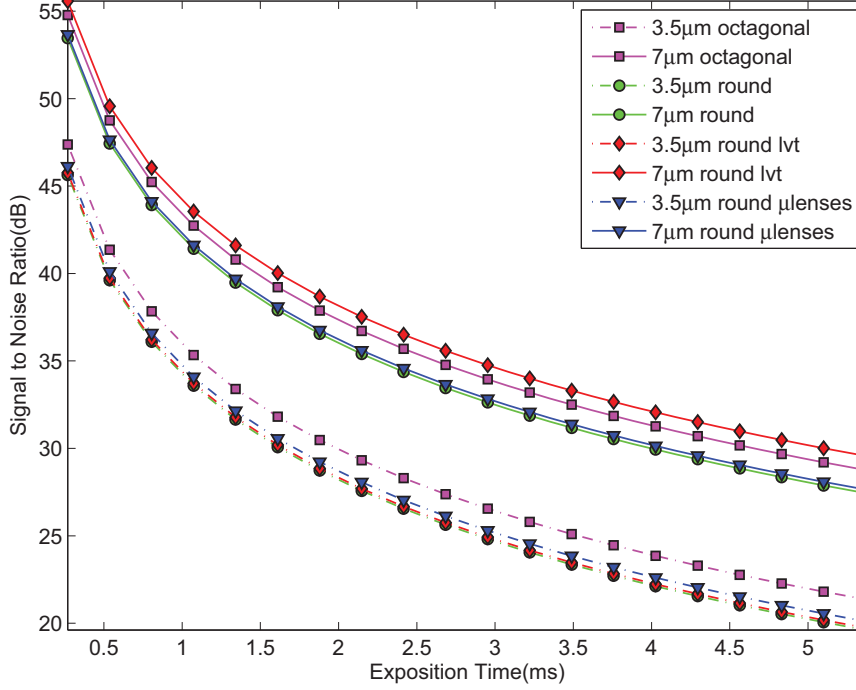


Figure 2.24: Signal to Noise Ratio.

2.5.5. Performance Comparison

We have discussed in the two previous sections how the octagonal pixels exhibit lower dark signal than the round-shaped pixels, and how this seems to be contradictory to the generally accepted assumption. Indeed our design does not allow us to measure dark current directly but its effect in the discharging rate of the photodiode’s associated capacitance in absolute darkness. Since the integration capacitance in round and octagonal pixels are different, measurements about dark signal are not absolutely representative of the real performance of the two shapes. On the other hand, we already described how the smaller associated capacitor of round-shaped pixels produced higher sensitivity than in octagonal pixels. Since our design does not allow us (this has certainly become a lack of the design) to extract the photodiode’s parasitic capacitances and simulation models available through Europractice are not precise at all, we have created a Figure Of Merit (FOM), which combines sensitivity and SNR. Indeed, if we consider that sensitivity increases as the parasitic capacitor shrinks down, whereas SNR increases as this capacitance grows, defining a FOM as:

$$FOM = SNR \times Sensitivity \tag{2.6}$$

may allow us to decide which of the many sensors in the SCU performs better, and to run comparisons considering pitch, shape, etc. Figure 2.25 shows the results for the FOM when the sensitivity is evaluated at 550nm. Results are clearer in this case, with round pixels with microlenses featuring the best FOM for the two pitches, followed by the round pixels including Low- V_{th} source follower transistor. Finally, when

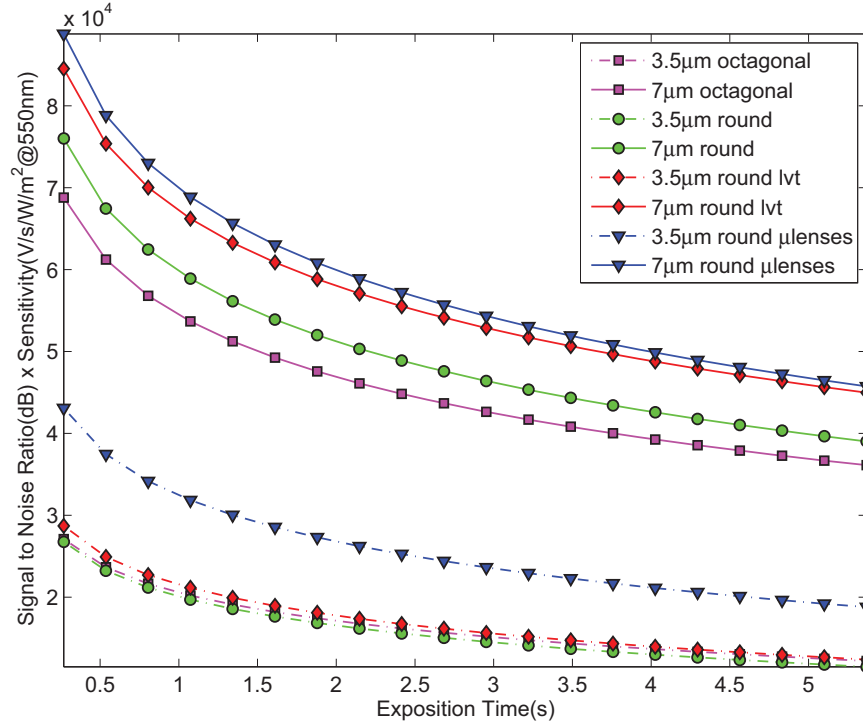


Figure 2.25: SNR x Sensitivity comparison

comparing the performance of the round and octagonal shapes without any additional extra feature, we see that for the 7µm pitch pixels, the round diffusions obtain a better FOM whereas for the 3.5µm pitch case, both shapes behave very similarly, as in this case, the gate capacitance of the source follower is an important component of the total integration capacitance.

2.5.6. Microlenses Effect

In order to observe the effect of the microlenses, we can simply consider the sensitivity results (figure 2.21) for the arrays including microlenses. For a simpler visual evaluation, figure 2.26 shows the ratios in the sensitivity of the same pixel with and without microlenses for the two pitches employed in this design. First of all, it is noticeable that the use of microlenses significantly improves the response of the 3.5µm pitch pixels, and it affects positively in the 7µm pitch pixels but at a lower ratio. Obviously, this improvement difference is caused by the relationship between pixel pitch, microlenses radius and thickness of the metals and passivations on top of the sensor. In the case of the 3.5µm pitch pixels, lateral dimensions are comparable to vertical dimensions and, therefore, the amount of light with large angle of incidence redirected into the sensor area is comparatively higher for this pixel than for the larger 7µm pitch pixels.

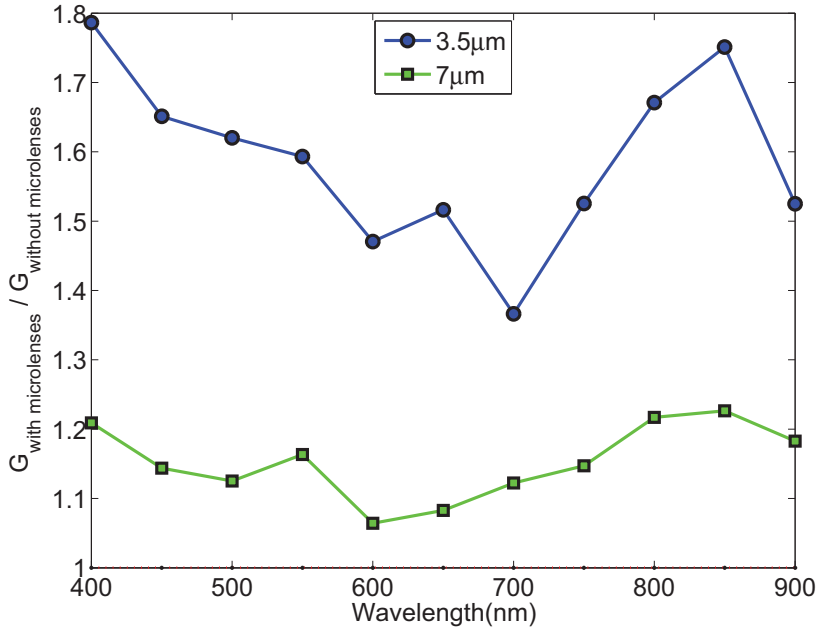


Figure 2.26: Relative sensitivity gain by microlenses.

2.5.7. Crosstalk

Crosstalk structures have been excited with 550nm light in order to study the signal acquired by supposedly blind pixels, which surround the central one. Exposure has been adjusted such that the central pixel is just at the limit of the saturation. Figure 2.27 shows a zoom around the neighbors of the central pixel. As expected, in the pixels without microlenses crosstalk diminishes as pitch increases. Regarding the 3.5µm pitch pixel with microlenses, it is observed that the crosstalk diminishes to an amount comparable to the results for the 5µm pitch pixel without microlenses. It is also remarkable that diagonal directions are much less affected than manhattan dimensions, as well as the fact that vertical asymmetry in the layout of the pixel (figure 2.8) produces a peak of crosstalk in the upper neighbor.

2.6. Conclusions

Through this chapter, we have presented the design and test of a chip whose purpose was to evaluate the possibility of using CIS technologies for obtaining high dynamic range operation in Focal Plane Processors. The chip has been designed in a 0.18µm CIS technology from UMC available through the Europractice consortium. Two basic APS structures have been included in the chip for performance evaluation, a 3.5µm pitch pixel and a 7µm pitch pixel. For this two pixel pitches, we have produced two versions of the layout of the active diffusion: round-shaped and octagonal. Additionally, two versions of the source follower: using/not using low- V_{th} transistor. And finally, two additional variations consisting in the use or not of microlenses. For different combinations of these features we have measured Sensitiv-

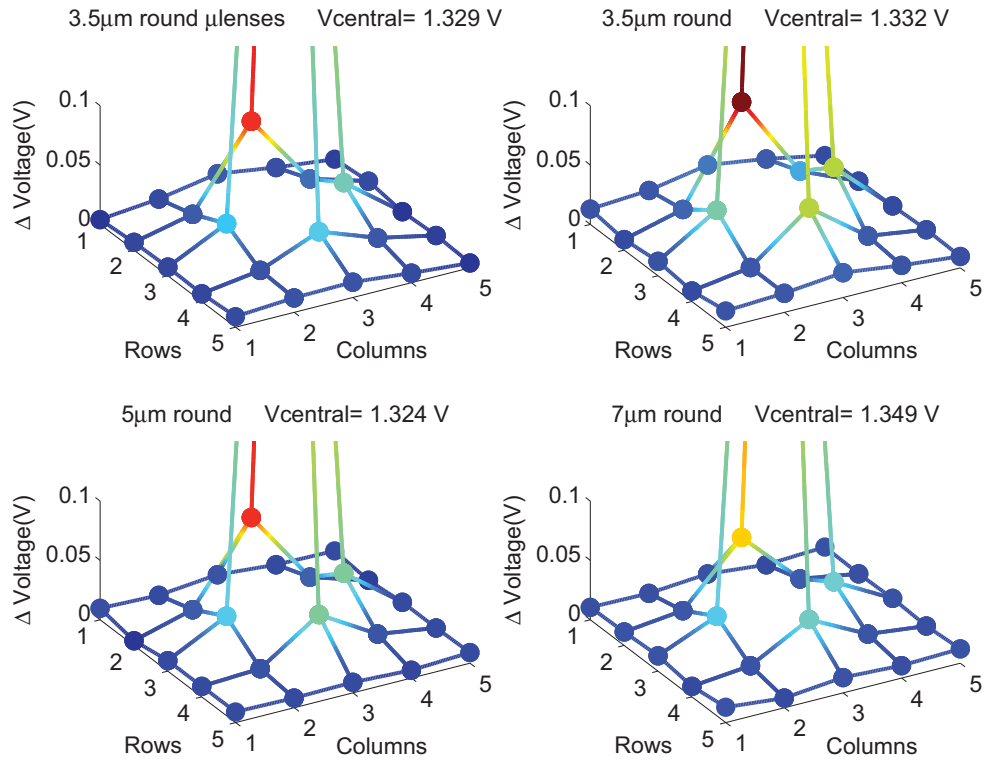


Figure 2.27: Crosstalk influence comparison.

ity, Dark Signal, Signal to Noise Ratio, Dynamic Range and Crosstalk, and we have also compared the resulting values with those obtained using the conventional CMOS 0.18 μ m technology from UMC. We have created a figure of merit to compare the overall performance of all kinds of sensors, obtaining that in general, the best performing sensor is a round sensor including microlenses, followed by the round sensor including Low- V_{th} transistor in the source follower. Regarding dynamic range, the improvement produced by the reduction of the dark signal is moderate, below 60dB. This, together with the fact that we cannot add additional processing circuitry at the pixel level in order to improve the dynamic range due to the strict design rules (specially the prohibition of using PMOS devices near the sensors), forces us to decide not to use CIS technologies for high dynamic range operation in Focal Plane Processor chips.

2. SCU: A TEST CHIP IN CMOS IMAGE SENSOR TECHNOLOGY

Chapter 3

Tone Mapping Algorithm

3.1. Introduction

During recent years, multiple HDR cameras have been developed in order to solve the dynamic range limitations reported by first CMOS cameras. However, they usually imply a large amount of data per image, such as computation over several frames, or a fixed non-optimized compression, as reviewed in chapter 1. Therefore, working with non-compressed HDR images requires employing long bit-words per pixel. These images are very difficult to handle either by processors or by visualization devices, which usually employ the typical 8-bit coding per color channel. Moreover, non-compressed HDR data can consume enormous amounts of storage space or transmission bandwidth.

In order to reduce the amount of data of the HDR imager presented in this work while keeping the details, it will be applied a Tone Mapping algorithm. Tone Mapping classically belongs to computer graphics and it is performed in software. This technique is used to compress HDR scenes in Low Dynamic Range (LDR) output representation while preserving the details contained in the scene, usually as a human would perceive it.

HDR tone mapping compression can reduce the size of visual data by more than several hundreds of times without introducing objectionable distortions. Therefore, this tone mapping compression has become an important part of many HDR image and video systems. HDR tone mapping techniques are also used to render digital 3D scenarios [76], these scenarios are simulated with a high dynamic range illumination and, therefore, rendering images have also the difficulty of how to represent these virtual HDR scenarios with the LDR representation methods that are commonly used nowadays. These tone mapping techniques are generally intended for still images and relatively very little effort has been put towards HDR video tone mapping.

Regarding tone mapping techniques, it is inevitable that the reader will make a visual evaluation of the goodness of the tone mapped image. In our case, the algorithm is not intended to maximize human perception (although it could be used for this purpose). Therefore, this characteristic can lead to mistakes towards the goodness of the tone mapped image quality. However, the human perception

3. TONE MAPPING ALGORITHM

will be a good evaluator, in a rough manner, of details lost in scenes. Consequently, it is necessary to clarify the perception of the representation of this thesis (or perceived HDR scenes) in combination with the effects of human perception. Therefore, this chapter will include background knowledge that will allow a critic evaluation of what it is perceived by the reader directly through real world or in representation methods. In this background section, human perception concepts that are typically used in Tone Mapping Algorithms are also included. These basic concepts will be used in the description of typical Tone Mapping techniques, which will be presented in this chapter.

After summarizing the general knowledge about human perception, image reproduction methods, and typical tone mapping techniques, the developed new hardware-aimed Tone Mapping Algorithm is described and a mathematical simulation over a HDR scene is presented. Finally, a comparison with other tone mapping algorithms is included.

3.2. Background

3.2.1. Human Vision

The human visual system deals with a similar problem as vision systems, a large amount of captured information to be directly processed. Human photoreceptors capture information, which need to be transmitted to the brain through the optic nerve. The amount of information that can pass through the optic nerve is limited, and therefore it constitutes a bottleneck. In fact, the number of photoreceptors in the retina is far larger than the number of nerve endings that connect the eye to the brain. The human eye solution is parallel processing near the photoreceptors. After light is absorbed by the photoreceptors, a significant amount of processing occurs in the next several layers of cells before the signal leaves the eye. The organization of the retina is shown in figure 3.1 [77], where we can observe that the signal is compressed, as there is more photoreceptors than connections to the optic nerve. This is performed due to the combination of bipolar, ganglion, amacrine and horizontal retina cells [52], which will perform a pre-processing of the signals.

In general, the human visual system processes HDR scenes and compresses them to a lower amount of data. Due to this mechanism, which is a sort of tone mapping system, a person can perceived a bad tone mapping algorithm, comparing its results with a similar memorized HDR scene which has been directly perceived. Humans can easily perceive the loss of information from real scenes to the display representation, such as overexposure, underexposure, halos over bright lights, or unnatural cartoon-like/painting-like images because our visual system does not perform these failures or features.

3.2.1.1. Dynamic Range Adaptation

The human eye is capable of adapting to lighting conditions that vary by nearly 10 orders of magnitude [52] inter-frame and it can perceive a range of about 5 orders of magnitude intra-frame. Therefore, inter-frame adaptation makes our visual system less sensitive in daylight and more sensitive at night.

Visual adaptation to varying conditions of illumination is performed by the coordinated action of the

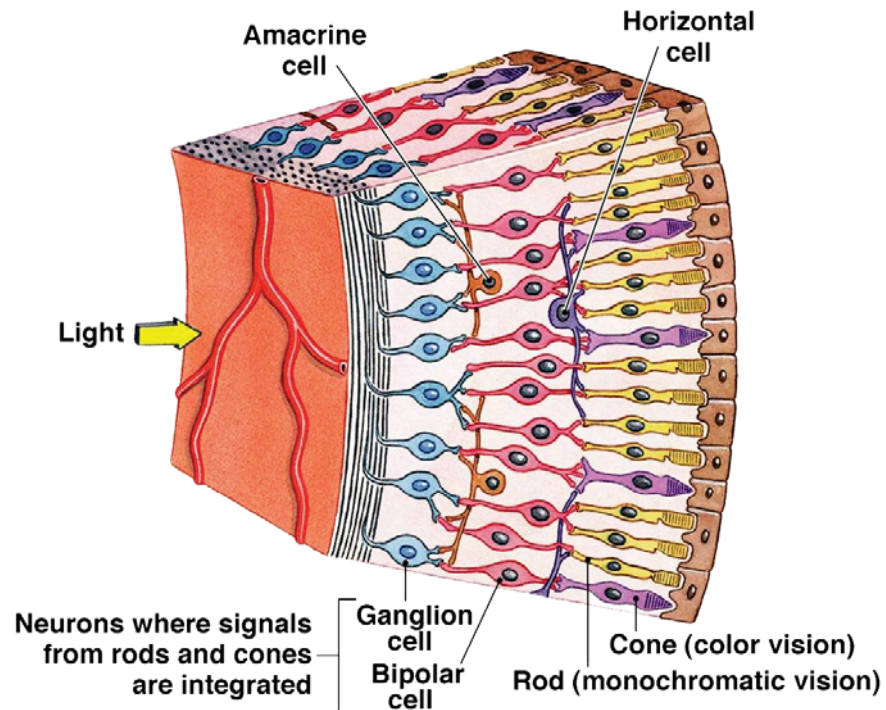


Figure 3.1: Organization of the retina.

pupil and the rod-cone system. The pupil is the opening within the iris through which light passes before reaching the lens and being focused onto the retina. Rod and cones are the two types of photoreceptors present in the human retina.

The pupil changes its size in response to the background light level. Its diameter changes from a minimum of about 2mm to a maximum of about 8mm. This change accounts for a reduction in light intensity entering the eye by only a factor of 16. Hence, the pupil does not play the key role in visual adaptation. The rod and cone combination are the main responsible of the effective dynamic range of the human visual system.

The retina contains approximately 5 million cones and 100 million rods in each eye. The rods are more sensitive than the cones. Therefore, rods perform the vision under low illumination levels, which is called scotopic vision. Cones are intended for high illumination levels, which is named photopic vision. There is a range of illuminations where both rod and cones are active, which is defined as mesopic vision. The mesopic range takes place between moonlight to indoor light. The boundaries of this vision region are depicted in figure 3.2 [78]. In this figure, the illumination conditions are at the top while threshold and saturation or damaged point are at the bottom, where it is indicated color perception and visual acuity, which is the ability to perceive details. Under scotopic illumination, rods are more sensitive than cones because they have a much lower threshold (*Absolute rod threshold*). In the increment of illumination, cones start to sense (*Cone threshold*) and become increasingly sensitive. At higher illuminations, rods begin to saturate (*Rod saturation begins*) and eventually the rod system

3. TONE MAPPING ALGORITHM

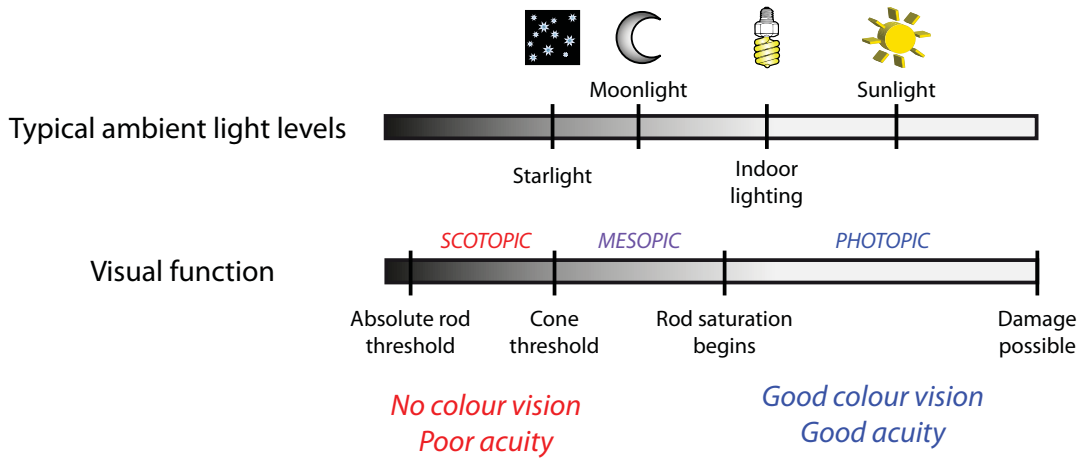


Figure 3.2: Light levels effect on photoreceptors visual function.

becomes incapable of discriminating details. Finally, the superior boundary for cone system implies a physical damage (*Damage possible*) to the photoreceptors, such as stare directly to the sun.

Regarding light wavelength, rods acquire illumination in a monochromatic way while cones perceive colors. In order to perform color perception, there are three types of cones: L, M and S as they perceive mainly red, green and blue, respectively. Combining these mechanisms, the human eye is sensitive to light between approximately 380 to 830 nanometers.

3.2.1.2. Tone Mapping Techniques: Basic Concepts

The perception of the illumination in terms of brightness is usually used in tone mapping algorithms. The perception of brightness is expressed as Luminance ($\frac{cd}{m^2}$). It constitutes an approximate measure of how bright a surface appears to a human person. Luminance is in fact the most relevant photometric unit to HDR imaging [79]. It is generated weighting the measured radiance to the human perception characteristic in wavelength.

The perception of brightness is not linear in humans but it is compressed toward higher values. Moreover, it approximately follows a logarithmic behavior over a broad range of illuminations. This is proposed by the Webber-Fechner law. Weber's law was first described in 1834 by the German physiologist Weber [80], and was later formulated quantitatively in 1860 by the experimental psychologist Fechner [81], founder of the modern psychophysics. In psychophysical studies, human visual adaptation is evaluated by measuring the minimum amount of incremental light by which an observer distinguishes a test object from the background light. This minimum increment is called a visual threshold or Just-Noticeable Difference (JND). Weber stated that the change in stimulus intensity that produces a just noticeable difference in a human sensation is a constant fraction of the starting intensity of a stimulus. In other words, the minimum amount by which a stimulus magnitude must be changed in order to produce a noticeable variation in sensory experience is a constant fraction of the level of stimulus intensity. So, Weber's Law, which has wide generality across different sensory magnitudes and modalities, is ex-

pressed in equation 3.1 where c is a constant and I is the level of stimulus intensity. The constant nature of this fraction suggests that visual adaptation acts as a normalizer, scaling scene intensities to preserve our ability to sense contrasts within scenes.

$$\frac{\Delta I}{I} = c \quad (3.1)$$

Fechner based his work on Weber's Law. He proposed that as the stimulus intensity increases, it takes greater and greater changes in intensity to change the perceived magnitude by some constant amount. Therefore, the perceived magnitude is a logarithmic function of stimulus intensity multiplied by a specific constant. Fechner's Law relationship is expressed in equation 3.2, where S is the perceived magnitude, k is a constant and I is the stimulus intensity.

$$S = k \cdot \log(I) \quad (3.2)$$

These fundamental laws are usually valid for general sensory phenomena and can account for many properties of sensory neurons. However, the human visual perception is not necessarily logarithmic in the entirety of the dynamic range and the logarithmic law is not the only psychophysical law to explain brightness perception. The Stevens' Power Law was introduced in 1957 by Stanley Smith Stevens [82]. It proposed that perception follows a power law as expressed in equation 3.3,

$$S = k \cdot I^\gamma \quad (3.3)$$

where S is the perceived magnitude, k is a constant, I is the stimulus intensity and γ is an exponent constant that depends on the type of stimulation. In practice, the Weber law is a good approximation for bright illumination conditions and for lower adaptation. Luminance levels power functions fit better to the experimental data. However, it is indubitable that, apart from the accuracy of the perception function itself, the theories propose the increasing compression toward higher stimulus intensities.

Summarizing, some concepts that must be considered in the general description of tonemapping algorithms have been presented: Luminance, Webber-Fechner Law, Just-Noticeable Difference (JND) and Stevens' Power Law.

3.2.2. Dynamic Range of Reproduction Devices

Image visualization devices fall into two major categories: hardcopy and softcopy devices. Hardcopy devices are static and passive. Some hardcopy devices produce transparencies filtering the light, but most produce the reflection of ambient illumination. In general, they are many methods for permanent reproduction, such as:

- traditional ink presses
- dye-sublimation
- thermal, laser and ink-jet printers

3. TONE MAPPING ALGORITHM

Regarding DR, hardcopy media will never become HDR because this would entail the invention of methods such as light-emitting paper. Hardcopy are usually limited to a contrast ratio of 200:1, and so it is usually insufficient even for 8-bit LDR representations.

Softcopy devices are dynamic and active. Most of these display devices include an integrated light source. Displays are commonly not able to generate pure black emission because there is leakage of light from other pixels or reflections from ambient light. The limitations of the most typical softcopy devices are:

- Cathode-ray tubes (CRTs): The fundamental constraint for dynamic range in CRTs is their maximum brightness, which is limited by the amount of energy we can safely deposit on a phosphorescent pixel without damaging it or generating unsafe quantities of X-ray radiation. The CRT displays have a very high dynamic range, but it is not useful because the range is mainly at the low end, where we cannot see it under normal viewing conditions. A good CRT monitor will be close to displaying a pure black, and therefore offering a contrast that exceeds 1000:1.
- LCD displays: There is no fundamental limit to the amount of light one can pass through an LCD screen, as it only imply changing the backlight source. Regarding the backlight, LCD displays can be divided in two main categories: Light Emitting Diode (LED) and Cold Cathode Fluorescent Lamp (CCFL) backlight. LED backlight can achieve better brightness and an individual control over pixel backlight (white or RGB) improving the dynamic range. LCD display operation is not able to completely turning on and off the backlight, and therefore they achieve a typical static contrast ratio of only 400-600:1.
- Plasma displays: Despite not being able to display the purest black, Plasma Panel exhibit the best contrast, offering a 1000-3000:1 ratio, due to higher brightness capabilities.

Moreover, LCD panel suffers from angle view limitation in comparison with CRT and Plasma displays. If the viewer is situated in a certain angle regarding the center of the display, the person will experience a loss of brightness, contrast, and color. In certain cases, angled view will imply the perception of image details in the darkest or brightest areas, which are lost in direct viewing due to contrast limitations.

CRT displays were in previous years the common monitor for computer. However, they have been surpassed by LCD panels because they are thinner and lighter for same display size. Regarding softcopy DR, HDR displays are in development and will become available in the future. However, typical monitor displays nowadays have a contrast ratio capability of only 100-1,000:1.

Summarizing, typical hardcopy methods of representation are limited in DR even for the representation of 8-bits photography. Regarding softcopy, the reader must consider the depicted dynamic range and angle of view limitations of the used display.

3.3. Tone Mapping Techniques

Tone mapping techniques are intended to adapt HDR images to LDR representation. However, since the early development of these techniques, the final objective was the presentation of images to human users in conventional displays. Since early stages in the research of tone mapping, it is noticeable that, direct compression of the intensity level and the contrast range (to fit them into the display limits) is not sufficient to reproduce the accurate visual appearance of the scene. In the beginning, researchers studied the problem and suggested the use of models of the human visual system. Therefore, tone-mapping algorithms usually incorporate visual models in their operation. Due to the complexity of this task, tone mapping have been classically performed in the field of the computer graphics. Therefore, image sensor architectures toward tone mapping are not common.

Tone mapping techniques are usually divided in four categories:

- Global operators: The same function is applied to all the array of pixels. The function depends on global variables, such as mean luminance.
- Local operators: They apply a different function depending on the pixel. The function usually depends on local variables, such as the value of the pixel and its neighbors.
- Frequency Domain operators: The function depends on the spatial frequency of the image.
- Gradient Domain operators: The applied function depends on the derivative of the image; it is to say, the direction and value of the change between pixels. Usually, they attenuate the image gradients with large magnitude while magnifying the small ones.

The most used schemes are global and local operators. However, given the nature of the problem (infinite kinds of scenarios), it is not possible to have an all-in-one method. Depending on the user or application requirements and the HDR scene in particular, a method or even a combination of them will be better suited.

3.3.1. Global Operators

Global tone mapping algorithms apply the same function to all pixels of the image. The applied function can be a power function, a logarithm, a sigmoid, or a function that depends on a global descriptor, such as the histogram.

Global operators have the advantage of being computationally efficient. In general, global operators are the fastest to execute because normally only two or three passes over the image are required. In each pass, very simple computations are performed. Many of them may be executed in “real time” in computers. Moreover, the performance of global operators usually depends only on the size of the image. Global operators are usually much faster than the other operators. Consequently, applications that require speed or limited resources should consider global operators over all others.

Global tone mapping operators are the simplest to implement, but they tend to lose details. Global operators functions usually increase monotonically. Otherwise, visually unpleasant artifacts will be

3. TONE MAPPING ALGORITHM

introduced because the produced image seems rare or artificial to the human visual system. In fact, the higher the dynamic range of a scene the more values must be mapped to the usual 256 levels of display by a monotonically increasing function. Hence, their disadvantage is their lack of visibility and contrast preservation in very large dynamic range scenes. Additionally, some operators are not able to handle all kinds of HDR scenes.

In order to have a general idea of the algorithms that can be performed globally to the HDR scenes, some typical global tone mapping operators are briefly summarized in the following subsections.

3.3.1.1. Miller's Operator

Miller et al. developed in 1984 a tone-reproduction operator that tries to preserve the sensation of brightness of the image before and after dynamic range reduction [83]. Brightness is a complex function of both luminance and spatial configuration. Miller et al. proposed a model that keeps the brightness ratios constant before and after dynamic range compression. Stevens' psychophysical data is used for luminance calculation.

3.3.1.2. Tumblin-Rushmeier's Operator

The Tumblin-Rushmeier operator is based on the same psychophysical data as Miller's operator (Stevens' studies), but the brightness function is stated slightly differently [84]. Tumblin-Rushmeier attempts to preserve the brightness values themselves, in opposition to Miller et al. who proposed to preserve brightness ratios.

3.3.1.3. Ward's Scale Factor

This technique tries to preserve contrast instead of brightness [85]. It proposes that the smallest perceptible difference (Just Noticeable Difference [JND]) in a real scene corresponds to the smallest perceptible difference in the image. The psychophysical used data belongs to contrast sensitivity model of Blackwell's studies [86].

3.3.1.4. Ferwerda Visual Adaptation Operator

The concept of matching JNDs of real world and displays was extended by Ferwerda et al [87]. However, they extended the complexity of the psychophysical data. Ferwerda et al. added a scotopic vision component while Ward's contrast-based scale factor only takes care of photopic vision. The perception data also modeled the loss of visual acuity (the ability to perceive details) under scotopic lighting and the process of illumination adaptation.

3.3.1.5. Logarithmic and Exponential Operators

These techniques apply a simple global tone mapping curve following a logarithmic or exponential behavior [79]. These compressions are consistent with the logarithmic like behavior of brightness

perception stated by Fechner's law. Logarithmic and exponential operators are the most uncomplicated tone mapping operators. Therefore, they provide a baseline for comparison of resources effort. This last usually determined by execution time in personal computers. For medium dynamic range scene, these simple techniques provide a final image quality comparable with operators that are more complex.

3.3.1.6. Drago Logarithmic Operator

Drago et al. also propose a logarithmic function application, but it is extended to manage a wider dynamic range than the simple logarithmic operators [88]. This operator implies a logarithmic relationship of the tone mapping technique following Stockham's work [89]. He recommended that the displayed luminance L_d is derived from the ratio of world luminance L_w and maximum luminance in the scene L_{max} . This tone mapping operator realizes the adaptive adjustment of the logarithmic base depending on pixel radiance. The base is varied between 2 and 10, allowing contrast and detail preservation in dark-medium regions and large compression in bright regions.

3.3.1.7. Reinhard-Devlin Photoreceptor Operator

Reinhard and Devlin [90] proposed in this operator a modified version of the photoreceptor model which was developed by Hoods et al. [91]. This operator is applied to each of the red, green, and blue color channels separately. An important contribution of their work is the control of color saturation in the way of von Kries model of 1902 [92], which enunciated that chromatic adaptation occurs independently in the LMS cones.

3.3.1.8. Ward Histogram Adjustment

This operator proposed by Ward Larson et al. is inspired in the typical histogram enhancement techniques, which are applied to LDR images to improve contrast [93]. This operator computes a histogram of a subsampled density image to obtain the distribution of pixels in luminance. The density image will be a logarithmic representation of the pixels in an approximation of brightness from luminance values. The result of the cumulative histogram can be directly used to map luminance values to display values.

This technique tries to keep the contrast rather than maximizing it. Therefore, post processing steps are applied to the model to simulate glare, color sensitivity and spatial acuity to maintain subjective correspondence and visibility.

3.3.1.9. Schlick's Rational Operator

This technique was introduced by Christophe Schlick [94] in 1994. It is intended to improve computational efficiency and simplify parameters rather than the subjective correspondence achieved by previous methods. This operator applied a first-degree rational polynomial mapping function as an alternative to previous linear, exponential or logarithmic mappings. The application of rational functions is efficient to compute and easy to tune due to only two user parameters. This operator works prop-

3. TONE MAPPING ALGORITHM

erly when applied to high-contrast scenes but it is especially well suited for scenes containing strong highlights.

3.3.2. Local Operators

Local operators generally perform a local adaptation depending on the value of the pixel of interest and a set of pixel in its neighborhood. Therefore, a set of the pixel neighborhood determines how this pixel is compressed. Hence, a bright pixel in a dark neighborhood is different from a bright pixel in a bright neighborhood and the same for dark pixels. Local operators do not provide a monotonically increasing function. Therefore, two pixels with equal luminance values can be mapped into two different LDR values.

Regarding the computation of local operators, several parameters are widely implied:

- amount of neighbor pixels to be included in the computations
- weight of each neighbor pixel in relation to the pixel of interest
- once the adaptation reference levels are determined, how are they finally applied

In general, local operators provide much better perception results than global ones. This is explained by the fact that the human vision system is more sensitive to local contrast than to global contrast. This means that the human vision does not adapt to the scene as a whole but to smaller regions instead. Hence, local operators are frequently inspired by biological features of the human visual system. However, this good performance has the disadvantage of an increased complexity, as these operators are much more computationally intensive and harder to tune than global operators. The tuning difficulty is due to several parameters in the algorithms that have to be set empirically. Moreover, local operators may have the disadvantage of halo or ring artifacts. These artifacts are usually present in the surrounding of sources of light due to the presence of high gradients in illumination.

The local operators to be applied in tone mapping techniques are not as straightforward as global ones. Following sections will present the main characteristics of the most typical local operators.

3.3.2.1. Chiu's Operator

The first local operator was developed by Chiu et al. in 1993 [95]. This operator was inspired by artistic techniques that frequently use spatially varying techniques to trick the eye into perceiving a wider dynamic range than the real one. Particularly, the areas around bright features may be dimmed somewhat to accentuate them. They stated that the eye is more sensitive to reflectance than luminance, and hence slow spatial variation in luminance may not be greatly perceptible.

This operator multiplies the pixel luminance by a local average producing a low-pass filtered version of the image. Consequently, the pixels in this image are a weighted local average. Finally, these values are inverted and then multiplied by the original pixel luminance scaled by a constant factor.

This operation suffers of one of the typical problems of local operator: halos around bright features. In order to solve this issue, Chiu's operator first approach included a smoothing stage, which iterates at

least 1,000 times over the image with a small filter kernel. This would reduce the effect of halos but it makes this operator to have an unpractical amount of computation.

3.3.2.2. Rahman Retinex Operator

Rahman et al. [96] developed their operator over a reinterpretation of the Land's retinex theory [97] of 1971.

There are two version of this operator: single-scale and multiscale. The multiscale version was introduced to reduce halo effects. In the single-scale version, this operator applied a low-pass Gaussian filter over the image to find local multiplying factors. They divide the pixel value by its low-pass filtered value. Finally, the logarithm of the result composes a reduced contrast "single-scale" version. In the multiscale version, the filter is applied to the image with kernel sizes doubling every step. This produces a group of images with increasing blurring. Finally, the final image is realized by the weighted sum of this stack of images with scaling and offset constants.

It is remarkable that, RGB channels are computed independently rather than unified on a single luminance channel. Therefore, it implies that the Gaussian-blurring convolution must to be repeated at least three times per image. Even more computation will be required in the multiscale version, where it would make three Gaussian filters for every image of the stack.

3.3.2.3. Pattanaik Multiscale Observer Model

An observer viewing a real scene would perform a chromatic adaptation. If the same scene were represented in a display, the observer would adapt differently to the image in the display device. Consequently, the displayed image is likely to be perceived differently than the real scene. Chromatic adaptation should be considered in HDR imaging and specially in tone mapping. However, tone mapping operators usually ignore this characteristic for simplicity.

Pattanaik et al. proposed a multiscale observer model, which includes one of the most complete color appearance models and consists of several steps of operation [98]. The appearance of color to humans are described by a group of properties named "color appearance correlates" [99]. These variables can be computed from the color's tristimulus (RGB) values as well as the description of the environment. Typically useful appearance correlates from a visual stimulus are:

- Brightness: Visual stimulus appears to emit more or less light.
- Lightness: The area where visual stimulus takes place appears to emit more or less light in proportion to a similarly illuminated area that is perceived as a white stimulus.
- Hue: attribute of color. Perfect grey, where all color (RGB) components are the same, is an achromatic color.
- Chroma: difference between a visual stimulus and an achromatic stimulus of the same brightness.

3. TONE MAPPING ALGORITHM

- Saturation: difference between a visual stimulus and an achromatic stimulus judge regardless of their brightnesses.

Appearance correlates are not computed directly but it made necessary the use of color spaces [100].

This multiscale color appearance model simulates luminance, pattern and color processing of the human visual system to accurately predict the color appearance attributes of spectral stimuli in complex surroundings under a wide range of illumination and viewing conditions. Therefore, the computation of this model is complex due to a large amount of steps in the process. Therefore, this model must only be used for images with an extreme dynamic range. In lower dynamic range scene, simpler models will be best suited.

3.3.2.4. Ashikhmin's Operator

The multiscale observer model try to include in the model all the steps of human visual processing that are currently well understood. However, for the limited present purpose, this complexity is not necessary. Ashikhmin's operator attempts to model human visual perception in useful characteristics toward dynamic range compression [101]. Therefore, this operator tries to obtain results similar to Pattanaik's operator but in a significantly simpler computational model.

3.3.2.5. Reinhard et al. Photographic Tone Reproduction

The Reinhard et al. photographic tone reproduction operator [102] is inspired by the Zone System described by the photographer Ansel Adams in 1940s [103]. A zone is defined as a Roman numeral associated with an approximate luminance range in a scene, as well as an approximate reflectance of a print. There are eleven print zones, ranging from pure black (zone 0) to pure white (zone X), each doubling in intensity, and a potentially much larger number of scene zones. The photographer uses measured information in the field to improve the chances of producing a good final print.

3.3.3. Frequency Domain Operators

Frequency domain operators rely on the fact that low frequency content in an image tends to be high dynamic range and high frequency content tends to be low dynamic range. Attenuating the low frequencies in the Fourier domain, HDR scenes could be compressed while the high frequencies (the low dynamic range details) are preserved.

Frequency domain operators are computationally based on the Fast Fourier Transform (FFT) to obtain a frequency domain representation. In these operators the performance of the FFT transform determine the computation time. The execution time of an FFT depends on the size of the image. However, this dependency is not linear. Images that are powers of two will be faster computed, and images of prime numbers size will be the slowest to compute.

Three typical frequency domain operators are: Oppenheim Frequency-based Operator [104], Durand Bilateral Filtering[105] and Choudhury Trilateral Filtering [106].

3.3.4. Gradient Domain Operators

Considering the gradients in the image, it is possible to partially distinguish between illuminance (incident light, slow gradients with arbitrary HDR) and reflectance (reflected light, high frequency, and LDR). The differentiation transforms images from the luminance to the contrast domain. The final integration of the thresholded derivatives leads to the reconstruction of image reflectance. Working with diffusely reflecting scenes, the separation of illuminance and reflectance may be reasonably successful.

Contrast domain offers several advantages over luminance domain:

- Contrast can be modified depending on its magnitude, hence taking advantage of contrast perception characteristics.
- Extremely sharp contrast can be achieved without introducing halo artifacts.
- The pyramid of contrast values, which approximates the localized contrast perception, can take into account both local and global contrast relations. Without this, undesired effects can take place in the areas where the input luminance is the same.

Although it has advantages, the gradient domain HDR compression consumes significant computational time. Two well-known gradient domain operators are: Horn Lightness Computation [107] and Fattal Gradient Domain Compression [108].

3.4. Proposed Algorithm

The present tone mapping algorithm is intended to capture images from HDR real world scenes and compress them into a low dynamic range representation. At the same time, this algorithm must preserve the details of the objects in the scene minimizing the computational complexity taking advantage of the parallelism of the focal plane circuitry. Under these assumptions, our algorithm performs a global tone mapping operator in the HDR scene. Once the compression variables are set, it will adapt the compression curve to the image to be processed. This adaptation will take a previously captured subsampled non-linear histogram as an indicator of the probability of illuminations. This method is capable of producing visually satisfactory results for a large set of examples. Moreover, as the computation of the tone mapping curve can be performed with little processing effort, this algorithm is suitable not only for tone-mapping over still images but also for HDR videos. However, for still images, it implies two image captures for every image due to the necessity of a previous image as an estimator of illumination probabilities.

Our algorithm performs two kind of compressions derived from the type of circuitry used in the approach: analog and digital compression. The analog compression is performed due to the mathematical intersection of the discharge signal in the integration pixel with a variable analog voltage reference. The discharge signal of a pixel is linearly dependent on the illumination of the scene. The digital compression is performed by assigning a digital code of the tone mapping curve of compression to each

3. TONE MAPPING ALGORITHM

pixel. The assigned code will depend on the time when the analog intersection described in the analog compression takes place. Both mechanisms of compression are explained in the following sections.

3.4.1. Analog Compression

The analog compression is a consequence of the method used for sampling the photogenerated current. In order to capture the illumination distribution of a scene, the time of intersection between the discharge signal of a pixel and an analog reference is measured. The analog reference will be fixed during most of the exposition time, till a certain time called T_{fixed} , as illustrated in figure 3.3, when it will ramp up to allow for poorly illuminated pixels to intersect. If we analyze the distribution of these two signals, we will notice an increasing compression toward higher photocurrent. In figure 3.3, the discharge signals of equally distributed photocurrents and an example reference are represented. The pixels discharge signals for the different illumination levels are depicted by blue lines, the voltage reference by a green line and their intersection by red dots. The start point of the ramp, T_{fixed} , and its slope are variable. The maximum exposition time T_{max} is given by the time needed by the ramp to reach the starting point of discharge of the pixels, the reset voltage V_{rst} , plus the time T_{fixed} .

Consequently, two analog compressions take place:

1. The compression applied to pixels whose discharged signal crosses during the fixed reference voltage.
2. The compression applied to pixels whose discharged signal crosses during the ramp-up reference voltage.

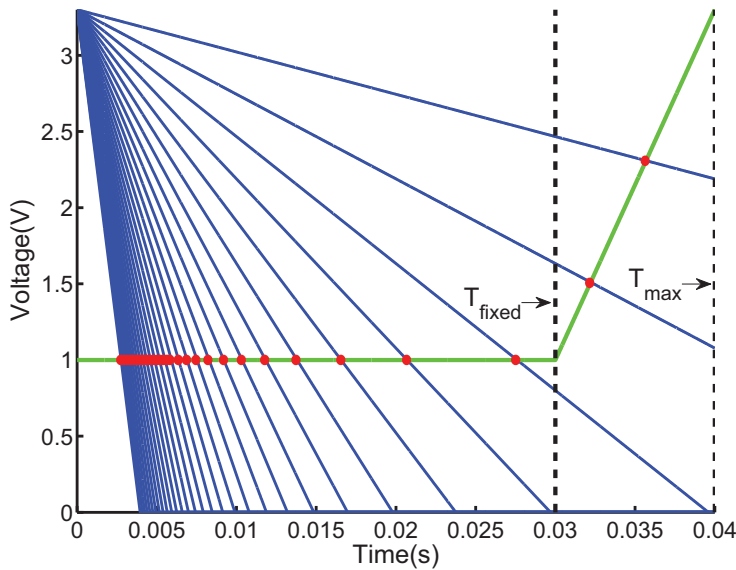


Figure 3.3: Intersection of discharge signals with ideal analog reference.

The discharge signal is directly proportional to the photogenerated current. Its behavior is depicted in equation 3.4, where V_{ph} is the discharged signal in the integration node, V_{rst} is the reset voltage (starting point of the discharge), I_{ph} is the photogenerated current, C_{ph} is the capacitance of the integration node and Δt is the time since the start of the discharge¹.

$$V_{ph} = V_{rst} - \frac{I_{ph}}{C_{ph}} \Delta t \quad (3.4)$$

Analyzing the first compression, the time that takes the discharged signals to reach a fixed voltage value is indicated in equation 3.5. Where T_{cross} is the time of intersection and V_{fixed} is the fixed reference voltage. The first compression due to this analog behavior will be an inverse function of the photocurrent.

$$T_{cross} = \frac{C_{ph}}{I_{ph}} (V_{rst} - V_{fixed}) \quad (3.5)$$

Regarding the second compression, the ramp voltage behavior is defined in equation 3.6. Where m is the slope of the ramp and T_{fixed} is the time when the fixed voltage ends and starts the ramp.

$$V_{ramp} = V_{fixed} + m (\Delta t - T_{fixed}) \quad (3.6)$$

The behavior of the crossing time of the discharged signal and a ramp-up voltage is indicated in equation 3.7.

$$T_{cross} = \frac{V_{rst} - V_{fixed} + m \cdot T_{fixed}}{m + \frac{I_{ph}}{C_{ph}}} = \frac{C_{ph}}{m \cdot C_{ph} + I_{ph}} (V_{rst} - V_{fixed} + m \cdot T_{fixed}) \quad (3.7)$$

Once other variables (V_{rst} , V_{fixed} , T_{fixed} , C_{ph}) are chosen, the slope m will determine the type of compression. The extreme values of the compression are expressed by means of the limits of T_{cross} for the slope 0 and ∞ , which are expressed in equations 3.8 and 3.9.

$$\lim_{m \rightarrow 0} T_{cross} = \frac{C_{ph}}{I_{ph}} (V_{rst} - V_{fixed}) \quad (3.8)$$

$$\lim_{m \rightarrow \infty} T_{cross} = \lim_{m \rightarrow \infty} \frac{C_{ph} \cdot \cancel{m} \cdot T_{fixed}}{\cancel{m} \cdot C_{ph}} = T_{fixed} \quad (3.9)$$

Therefore, if the ramp-up has a very high slope, the point of intersection will be determined as the intersection of the discharged signal with a vertical line at a fixed time, and it will determine a behavior directly proportional to I_{ph} ; this behaviour is given by the following equation:

$$V_{ph} = V_{rst} - \frac{I_{ph}}{C_{ph}} T_{fixed} \quad (3.10)$$

¹Here, the ideal behaviour is considered, where only the photogenerated current I_{ph} discharges the integration node whose capacitance is only due to the photodiode capacitance (C_{ph}).

3. TONE MAPPING ALGORITHM

The slope in this algorithm must be chosen as high as possible in order to take high advantage of the digital compression, which will be explained in the following section.

Summarizing, the analog compression will be inversely proportional to the photocurrent in the pixels crossing during the fixed voltage and almost directly proportional to the photocurrent in the pixels crossing during the ramp if the slope is high.

3.4.2. Digital Compression

The analog compression does not depend on the distribution of illuminations in the HDR scene. Therefore, high compression can be applied in very populated regions (of illumination) of the image. In order to avoid this behavior and considering that we have not yet digitize the data, we can modified the analog compression controlling the digital codes assigned at the intersection time, and applying different compressions to different zones of illumination of the HDR scene.

Two digital codes will be assigned to each pixel at the time of intersection of the discharged signal and the analog reference depending on their values: Time Stamp Code (TSC) and Tone Mapping Code (TMC). Both codes are given by the digital generated TMC and TSC curves. TSC is stored in order to have information about the time of crossing. TMC is the final tone mapped representation of the pixel and therefore the output value of the algorithm for that pixel. How TSC and TMC values are obtained for a low and high illuminated pixel is depicted in figure 3.4, where in this example the TSC curve has 2 bits, TMC has 3 bits and the slope of the ramp is moderate.

The applied TMC curve, $TMC_{<2:0>}$ in the example of figure 3.4, is the tone-mapping curve and determines the compression applied to the pixel values. This tone-mapping curve will be calculated from the cumulative histogram of non-linearly distributed crossing time zones measured by the capture of TSC digital reference (usually from the previous image).

The complete process to obtain the final pixel values includes different steps:

1. The photocurrents behavior from a subsampled image is analyzed by obtaining the histogram representation of the image composed by the TSC data, what we call Time Stamp Image (TSI). This histogram will indicate the distribution of the intersections in time.
2. The TMC digital reference curve is generated depending on the histogram of TSI.
3. The pixels acquire the TMC values depending on the crossing time with the analog reference obtaining the Tone-Mapped Image (TMI), which is composed by all the captured TMC values.

How TSC and TMC curves are generated is explained in the following subsections.

3.4.2.1. Time Stamp Code

The TSC curve is used to obtain the histogram of an image employed to identify those intersection times where the discharge signal commonly crosses with the analog reference. In order to perform this, the exposition time is divided in different zones depending on the value of the analog reference voltage: (1) intersections with the fixed voltage (before T_{fixed}) and (2) intersections with the ramp (after T_{fixed}).

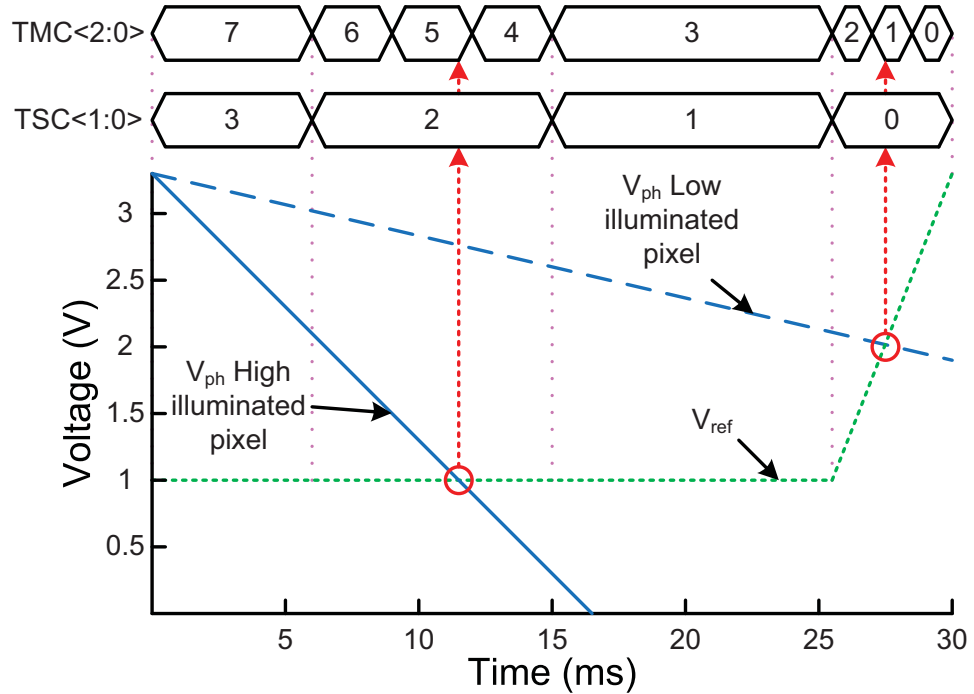


Figure 3.4: Assignment of digital codes at intersection time.

These zones are time divisions, which are named bins, as they are the possible values that take place in the histogram. The obtained histogram will be non-linear, as it is an accumulation of the intersections occurrence, which are non-linear. The bins are distributed in: (1) one bin for intersections during the voltage ramp reference, and (2) several bins for the intersections during the fixed reference voltage.

The time limits of the single ramp bin (T_{fixed} , T_{max}) are clear because it will only depend on the characteristics of the voltage reference behavior, which must have the highest slope permitted by the hardware. However, the distribution of the bins during the fixed voltage reference must be tuned by the user, depending on the target application or the illuminations required to be enhanced. As a first and general approach, we will choose a distribution of progressively increasing durations of bins, as the analog compression increases for lower intersection times during the fixed voltage reference.

The number of bins will depend in the number of bits chosen for the representation of TSC. If number of bits for TSC is 4, we will have 15 during fixed voltage reference and only 1 bin during the ramp. Generally, a scheme with duration of bins that increases with time, except the bin during ramp, will be more suitable. An example of 16 bins distribution with a base clock of $1\mu s$ is shown in table 3.1, where the TMC curve will have 7-bits and the clock cycle is $1\mu s$, resulting in a maximum exposition time of 15.488ms.

The bin with $TSC<3:0>=15$ corresponds to intersection times occurring between 0 and $128\mu s$, $TSC<3:0>=14$ between $128\mu s$ and $384\mu s$, $TSC<3:0>=13$ between $384\mu s$ and $768\mu s$, $TSC<3:0>=12$ between $768\mu s$ and $1.28ms$, $TSC<3:0>=11$ between $1.28ms$ and $1.92ms$, $TSC<3:0>=10$ between $1.92ms$

3. TONE MAPPING ALGORITHM

TSC<3:0>	Clocks per bin	Time per bin (μ s)	TSC<3:0>	Clocks per bin	Time per bin (μ s)
15	1	128	7	9	1152
14	2	256	6	10	1280
13	3	384	5	11	1408
12	4	512	4	12	1536
11	5	640	3	13	1664
10	6	768	2	14	1792
9	7	896	1	15	1920
8	8	1024	0	1	128

Table 3.1: Example of bins distribution.

and 2.688ms, TSC<3:0>=9 between 2.688ms and 3.584ms, TSC<3:0>=8 between 3.584ms and 4.608ms, TSC<3:0>=7 between 4.608ms and 5.76ms, TSC<3:0>=6 between 5.76ms and 7.04ms, TSC<3:0>=5 between 7.04ms and 8.448ms, TSC<3:0>=4 between 8.448ms and 9.984ms, TSC<3:0>=3 between 9.984ms and 11.648ms, TSC<3:0>=2 between 11.648ms and 13.44ms, TSC<3:0>=1 between 13.44ms and 15.36ms and TSC<3:0>=0 between 15.36ms and 15.488ms. Therefore, the lower TSC (bin index), the nearest it is to the ramp bin, which will be always coded as TSC=0. Hence, the bins are ordered in the table in temporal occurrence.

The TSC data is not stored for every pixel, but in a subsampled set of pixels, in order to reduce storing and computation efforts over this data.

3.4.2.2. Tone Mapping Code

Once the TSC data is captured, it must be accumulated to form a non-linear histogram of the illumination distribution. This histogram is the distribution of number of pixels intersecting the different bins given by the TSC curve. In order to determine the distribution of the digital codes of the TMC curve, a number of levels (TMC codes), which will be the final representation of the image, must be assigned to the bins. A first logical approach will be a distribution depending on the percentage of populations in the bins. Therefore, the more pixels intersect during a bin, the more codes of TMC (levels of output signal) are assigned to that bin.

The calculus of the number of levels assigned to the bins will be done dividing the number of pixels in each bin of the histogram of TSI (image composed by the TSC values) by a constant. This constant will be the quotient of dividing the total number of pixels in the TSI histogram by the total amount of levels, which are available for the final image representation. The quantity of total levels for the final image will depend on the final target representation. Therefore, $levels = 2^{bits\ of\ TMC}$, which are the number of possible different codes of TMC curve. However, the calculation by simply dividing by a constant can lead to non-integer assignments of levels. If we perform a normal rounding of the calculated levels, it can lead to assigning more levels than we have. Hence, first approach is the floor

rounding¹. However, it will lead to losing levels, as the sum of all dismissed decimal parts is the number of unassigned levels. In order to solve this issue, the lost levels are distributed in the bins that are closer to have another level, which are those with the highest remainders of the division by the constant = $\frac{\text{Pixels in TSI histogram}}{\text{Levels to be assigned}} = \frac{\text{Number of pixels in TSI}}{2^{\text{bits of TMC}}}$.

Table 3.2 shows an example of a simple case. The image dimensions is 40x40 resulting in 1600 pixels to be compressed represented with a 3-bits TMC, which implies an amount of $2^3 = 8$ levels. TSC is represented with 2 bits resulting in 4 bins, which are stored only one every two pixels in row and columns, which implies a subsample by 4 of the total amount of pixels in TMI (image composed by TMC values). The total amount of pixels in the TSC histogram is $\frac{1600}{4} = 400$. Therefore, the constant for the division is $\frac{\text{Pixels in TMI/subsample}}{2^{\text{bits of TMC}}} = \frac{400}{8} = 50$. In the floor round, $1+1+4+1=7$, one level is lost. The lost level is assigned to the maximum remainder of the division, which belongs to bin with $\text{TSC}<1:0>=2$.

Once the levels per bin have been obtained, they must be equally distributed in time inside the bins in order to generate the TMC reference over time. Additionally, the TMC reference must decrease over time so the pixels receiving highest illumination are represented by the highest final digital code (and lowest code are for lowest illumination), as it is the usual standard in image representations. It must be noticed that the pixel receiving highest illumination will be the fastest to cross the reference analog voltage.

The time per bin in the example of table 3.2 are 8, 16, 32 and 8 μs for bins 3, 2, 1 and 0, respectively. Therefore, if the levels must be equally distributed inside the bins, the time when a TMC change inside a bin can take place must be a multiple of $\frac{\text{Time per bin}}{\text{Number of total levels}}$, which in this case is $\frac{\text{Time per bin}}{8}$, resulting in 1, 2, 4, and 1 μs , respectively. These times divide the total time in subdivisions where the TMC digital reference is constant. These subdivisions are named subdivisions for evaluation since only between their boundaries the crossing of the pixel signal with the reference voltage can be evaluated. Figure 3.5 shows the TMC digital reference which must be generated for this example. Figure 3.5(a) shows the reference over time, where maximum exposition time is 64 μs . Figure 3.5(c) shows also the corresponding TSC reference represented as digital values. Red dotted vertical lines indicate the boundaries of the bins and green dotted vertical lines indicate the boundaries of subdivisions for evaluation. In more complex examples, some bins durations can be very reduced in comparison to the others, which will not allow a clear visual comprehension of the TMC curve applied. Therefore, figure 3.5(b) shows the TMC reference over the subdivisions for evaluations, in this case the number of evaluations are 4 bins \times 8 TMC possible codes = 32 subdivisions for evaluations. These subdivisions are calculated by dividing

TSC<1:0>	Pixels in bin	$\frac{\text{Pixels in bin}}{50}$	Floor rounded levels	Remainder	Levels per bin
3	50	1	1	0	1
2	90	1.8	1	40	2
1	200	4	4	0	4
0	60	1.2	1	10	1

Table 3.2: Levels per bin calculus.

¹Only the integer part of the number remains.

3. TONE MAPPING ALGORITHM

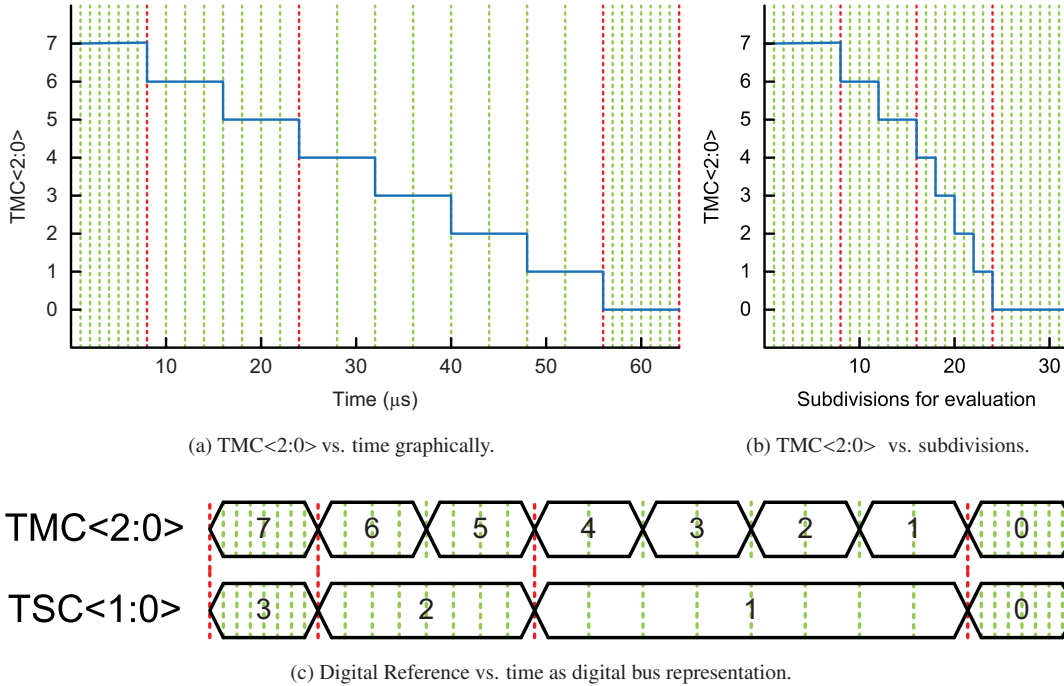


Figure 3.5: Example TMC curve.

the duration of every bin by the number of possible levels to be assigned ($2^{\text{bits of TMC}}$). This calculation comes from the possibility of assigning all the levels to a single bin, possibility which is shown in figure 3.6 that occurs when the whole population of the image intersects in a single bin.

In this case, the example of table 3.2 is simple; the number of levels per bin for the different bins are a submultiple of the number of subdivisions for evaluation. Generation of the TMC reference can simply be performed by decreasing one code every $\frac{\text{Number of subdivisions per bin}}{\text{Number of levels per bin}}$ subdivisions. In this case, the number of subdivisions is 8, so the change will take place every $\frac{8}{\text{Number of levels per bin}}$ on a bin depending on its assigned levels per bin, resulting in a change of code every 8, 4, 2 and 8 subdivisions inside the bins for bins 3, 2, 1 and 0, respectively.

However, when the number of levels per bin is not a submultiple of the number of subdivisions for evaluation (3, 5, 6 and 7 levels per bin), the generation of TMC reference is more complex. Figure 3.7 shows the example for 3 and 5 levels per bin in bins 2 and 1 respectively. The assignment of

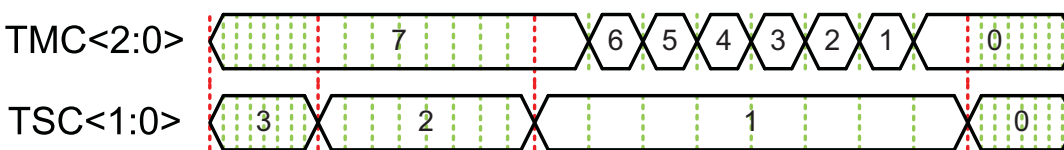


Figure 3.6: Example of the distribution of TMC codes all in one bin.

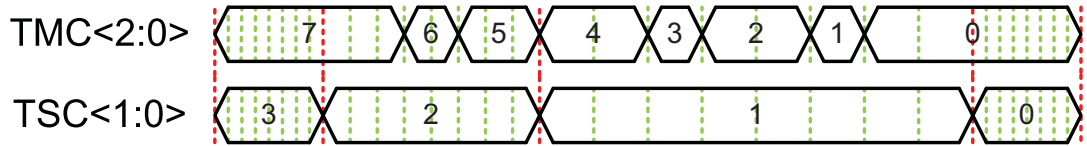


Figure 3.7: Example of TMC codes when levels per bins are not submultiple of total subdivisions.

codes cannot be exactly linear inside the bins. In order to solve this issue, the ideal rate must contain decimals. Therefore, the boundaries of subdivision where the code decrement must take place is found accumulating $\frac{8}{\text{Number of levels per bin}}$ with no rounding and round in the final value. The case for 7 bins per levels is shown in table 3.3 where the ratio of change is $\frac{8}{7}=1.1429$. Now, the problem is how many decimal are needed in order to avoid wrong results. This is important because it will not be possible to have infinite accuracy in the physical implementation of this algorithm. As it can be noticed in table 3.3, if only 1 decimal is used in the ratio, it provokes a mistake which is indicated with bold fonts. In the case of the calculation with all decimals, the subdivision where no decrement of TMC takes place is 4 but using only 1 decimal it is subdivision 5.

In this case, the number of bits used for TMC is low. If the number of bits used is higher, as an example 7 bits, the worst ratio for accuracy will be $\frac{128}{127 \text{ levels per bin}} = 1.0079$ where all decimal must be used in order to avoid an error. If this number is represented in binary, it will be necessary 14 bits ($\log_2(10079) = 13.2991$) plus the 7-bits necessary for the integer part of the accumulation, resulting in 21-bits. If this method is used, it will be also necessary to calculate the ratio in every level per bin with an accuracy of 14 bits, or having a memory of 128 words of 14 bits that contains the previously calculated values. Both options could imply a delay in the calculations between bins. In order to avoid this, another option have been developed, the subdivision with a decrement is already stored in a Look Up Table (LUT).

This LUT contains the positions of TMC one code decrement in the subdivision boundaries of a bin depending on the assigned levels per bin. The row address of the LUT corresponds with number of assigned levels, which have been obtained in the previous calculations. An example of the LUT for TMC of 3 bits is shown in figure 3.8, where 1 indicates the position of one code decrement and 0

Level	Level · 1.1429	Round (Level · 1.1429)	Level · 1.1	Round (Level · 1.1)
1	1.1429	1	1.1	1
2	2.2858	2	2.2	2
3	3.4287	3	3.3	3
4	4.5716	5	4.4	4
5	5.7145	6	5.5	6
6	6.8574	7	6.6	7
7	8.0003	8	7.7	8

Table 3.3: Subdivisions with decrement in the case of 7 levels per bin in 8 levels.

3. TONE MAPPING ALGORITHM

0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0
2	1	0	0	0	1	0	0	0
3	1	0	0	1	0	1	0	0
4	1	0	1	0	1	0	1	0
5	1	0	1	1	0	1	1	0
6	1	1	0	1	1	1	0	1
7	1	1	1	1	0	1	1	1
8	1	1	1	1	1	1	1	1
	1	2	3	4	5	6	7	8

Figure 3.8: Look Up Table of code decrements for TSC representation of 3-bits.

indicates no change in TMC code. Every change of bin must imply a decrement, in order to not share levels between bins (except in the 0 levels per bin case¹). Hence, the first bit of a row of this LUT always starts with a jump (1) but in the first row, belonging to 0 levels per bin data. Having the levels per bin and this fixed LUT for single code decrements, the tone-mapping curve can be generated in capture time just retrieving the corresponding 1-bit word for a subdivision for evaluation.

Figure 3.9 shows the composition of the LUT for a TMC of 7-bits, where green pixels indicate the position of one code decrement and blue pixel indicate no change. The rows indicate the number of levels assigned to a bin.

The calculation of this LUT has been performed distributing geometrically the code decrements. The positions of bits 1 will take place in positions which are rounded multiple of $\frac{128}{levels\ per\ bin}$. Then for n levels per bin the positions of bits 1 are rounded $1 \cdot \frac{128}{n}$, $2 \cdot \frac{128}{n}$, ... $n \cdot \frac{128}{n}$, which implies always a bit 1 at

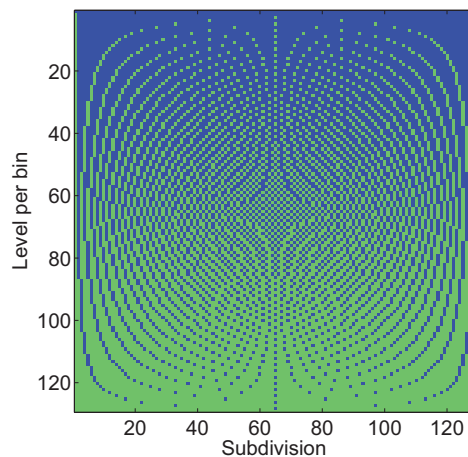


Figure 3.9: Look Up Table of code decrements for TSC representation of 7-bits.

¹If a change takes place the LUT row will be equivalent to the 1 levels per bin case.

position 128 ($\lfloor \frac{128}{n} \rfloor = 128$ independently of n). In order to differentiate code between bins, it is preferred to perform a code decrement in boundaries between bins. Therefore, it is required always a bit 1 (TMC one code decrement) in position 1 of every word; the solution followed in this work is to perform a flip left to right of the LUT. For example, the positions of bits 1 of the look up table for 3 levels per bin will be 3. Unflipped positions 3 will be $\frac{128}{3} = 42.6667$, $\frac{128 \cdot 2}{3} = 85.3334$, $\frac{128 \cdot 3}{3} = 128$. These numbers are rounded to be 43, 85 and 128. Finally, the results of flipping left to right will be $(129-128)=1$, $(129-85)=44$ and $(129-43)=86$ ¹.

Additionally to the LUT information, every first code change is dismissed. In the example of figure 3.5(c), bin with $TSC \langle 1:0 \rangle = 3$ is assigned 1 level per bin. Then, we go to the second row of the LUT, which corresponds to 1 level per bin, it indicates that the row is "10000000" that corresponds to activation or not of a decrement for the subdivision indicated by its position in the row, subdivision, as indicated in the x-axis of figure 3.8. This row of the LUT indicates a single decrement at the beginning (subdivision 1), and no decrement in the next evaluation subdivisions of that bin. As in figure 3.5(c), the code starts already at maximum code, this first decrement is not necessary. However, in the figure 3.10, bin $TSC \langle 1:0 \rangle = 0$ is assigned with 2 levels per bin, the LUT row will be the third row of the LUT "10001000" but here the first decrement of a bin is necessary to change from TMC code 2 to 1. Therefore, the first code indicated by the corresponding row of the LUT for a bin will be dismissed only if it is the first decrement that will take place in the TMC reference. In the case of a bin with 0 levels per bin assigned, no code change takes place, it is shown in figure 3.10 (0 levels in bin 3) and figure 3.11 (0 levels in bin 2).

Summarizing, the necessary steps to generate the TMC curve are:

1. Accumulate TSC data to obtain a non-linear subsampled histogram.
2. Divide the number of pixels in the bins by the quotient of dividing the total number of pixels in the subsampled histogram by the total amount of levels at disposal, $2^{\text{bits of TMC}}$.
3. Floor round the result of step 2.
4. Distribute lost levels in the bins with highest remainders in step 2.
5. Compose TMC signal in time with the positions of the decrements in the subdivisions indicated in the LUT, depending on the values of levels per bin obtained in step 4.

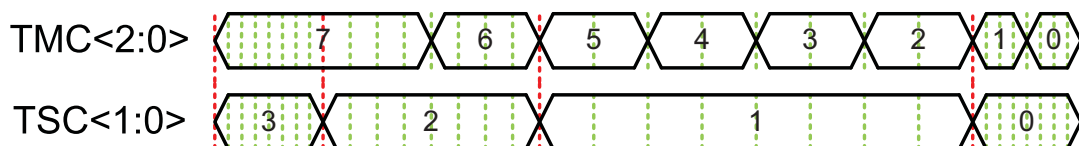


Figure 3.10: TMC code distribution with zero levels assigned to first bin.

¹It must be noticed that the LUT of TMC with 3-bits in figure 3.8 does not correspond to the calculations in table 3.3 because of the flip left to right.

3. TONE MAPPING ALGORITHM

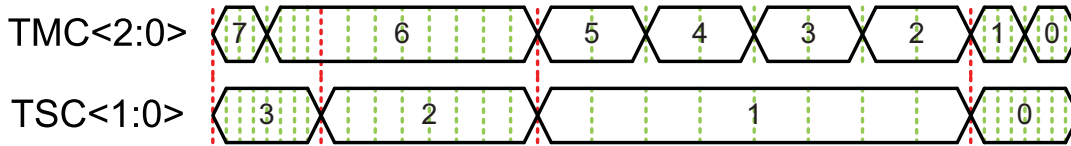


Figure 3.11: TMC code distribution with zero levels assigned to a middle bin.

3.4.3. Some Alternatives for Level per Bin Assignment

The method used in the distribution of levels per bin depending on the histogram of the previous TSI will determine the tonemapping operation mode. The distribution of levels per bin can be applied by a simple statistical calculation; i.e. depending on the weight of the bins in the histogram. However, variations over this method can be useful in some circumstances. Some possible modes of operations by simply changing the mechanism in the calculus of the levels per bin are:

1. Equal distribution: The same number of levels per bin is assigned in all active bins. It is used as the mode for the first shot, when we do not have TSI information.
2. Weighted: The levels are distributed depending on the weight of the bins in the histogram of TSI. The number of levels per bins is obtained by dividing the values of the histogram by $\frac{\text{number of pixels in TSI}}{\text{number of levels}}$ and floor rounding. The non-assigned levels are distributed between the bins with the higher remainders of the division. This option has been already described in previous sections.
3. Bin threshold: A bin is active when the histogram bin value is higher than a threshold. Levels are equally distributed between the active bins.
4. Avoid concentration in one bin: This mode operates as the weighted mode but it avoids a high concentration of too many levels in a single bin. If the number of levels in a bin exceeds a configurable limit, saturation is applied and the rest of levels are distributed between the bins with higher assigned levels.
5. Weighted with bright priority: If the most populated bin is a bright one, then this bin and the near lower neighbors are amplified by a fixed number of levels if they exceed a threshold. The rest of levels are assigned proportional to its weight in the TSI histogram.
6. Low populated priority: Low populated non-empty bins are amplified by assigning a fixed number of levels. The remaining levels are distributed as in the weighted mode among all the bins.
7. Weighted with minimum threshold: All non-empty bins are pre-assigned with a fixed number of levels. The rest of levels are distributed as in the weighted mode.
8. Non-linear levels adjustment: A non-linear function is applied to the TSI histogram, then the levels are weighted distributed. Here, the applied function was $y = x^{1.75}$.

-
9. Weighted with minimum low light priority: A minimum amount of levels is applied to non-empty bins. Then, if the most populated bin corresponds to one bin receiving high light (short crossing time), the levels per bin of bins receiving low light (long crossing time) are amplified by assigning an addition of several levels. Finally, the remaining levels are distributed as in the weighted mode.

3.5. Simulation Results

In the present section, an example over a real HDR scene will be illustrated. First, it is explained how to obtain an HDR scene mathematical representation. Second, the obtained TSI is presented as well as the performed calculation, resulting in the levels per bin. Third, the tone mapping curve (TMC codes) is obtain from the levels per bin. Finally, the tone mapping curve is applied to the original HDR image obtaining the result image of the algorithm.

3.5.1. Composition of HDR Image using Multiple LDR frames

In other to have a mathematical representation of a HDR scene, the combination of several Low Dynamic Range (LDR) frames is performed. This technique obtains a HDR image combining multiple images of the same scene that differ only in exposition time. These images are supposed to be linearly captured, and therefore they are not compressed or tone mapped. In our case, the photographs have been taken with different exposition times by the technique of exposure bracketing, where several shots of the same subject are captured using different camera settings. This has been performed varying the camera options manually: the aperture has been fixed to $f/16$ (low aperture to have all the objects in focus), the focal length is 55mm and the shutter speed has been varied. There have been obtained 18 frames with shutter speeds of $\frac{1}{4000}$, $\frac{1}{2000}$, $\frac{1}{1000}$, $\frac{1}{500}$, $\frac{1}{250}$, $\frac{1}{125}$, $\frac{1}{60}$, $\frac{1}{30}$, $\frac{1}{15}$, $\frac{1}{8}$, $\frac{1}{4}$, $\frac{1}{2}$, 1, 2, 4, 8, 15 and 30 seconds. Exposure Value (EV) [109] in photography indicates all combinations of a camera's shutter speed and relative aperture that give the same light exposure. The EV definition is shown in equation 3.11, where f is the f-number of the relative aperture and t is the shutter speed in seconds.

$$EV = \log_2 \frac{f^2}{t} \quad (3.11)$$

In the exposition compensation, which take as a base a calculated optimum EV value, $EV=0$ is the considered correct general exposure, positive values are overexposed and negative values are underexposed. An increment of 1 EV of compensation denotes a double exposure from the previous step. The EV calculated by Adobe® Photoshop® CS5 in the 18 frames with the variations are -8.09, -7.06, -6.06, -5.06, -4.06, -3.06, -2, -1, 0, +0.91, +1.91, 2.91, 3.91, +4.91, +5.91, +6.91, +7.81 and +8.81 respectively. The images are shown in figure 3.12. These pictures or photographs have been acquired with a commercial camera Nikon D90 [110][111], which includes a CMOS sensor [112]. The photographs have been taken using a tripod and an infrared remote control in order to avoid vibrations, which will cause differences in the scene between different frames. In order to reduce noise, as it will be greatly amplified in this kind of compositions, the frames have been captured with the lowest possible ISO

3. TONE MAPPING ALGORITHM

sensitivity: ISO 100 equivalent (named by Nikon L1.0). Moreover, a dark frame has been subtracted in the frames with exposition times 15" and 30" by the large exposition noise reduction option of the camera. The images has been taken with 12-bits AD conversion in RAW format (NEF in Nikon) that does not alter the information taken by the camera by post-processing (tonal adjustments, compression, etc.), which occurs in formats such as JPG.

Proper HDR composition of a set of images is not an easy task as, in commercial cameras, the exposition time is difficult to tune. Therefore, the obtained images have been introduced in the Adobe® Photoshop® CS5 HDR merging engine [113]. This engine uses the metadata in the image files to determine the exposure values of the images and the tonal response curve of the camera that took the images. It will perform an auto blend of parts of the scenario captured by each photograph. The engine also performs automatic alignment of the LDR frames and ghost removal in the final HDR image (caused by scene changes between photographs). The engine combines the images in an HDR representation of 32 bits known as the Radiance picture format (.hdr) [79], which is composed by three color channels of 8-bits (Red, Green, Blue) plus a common exponent of 8-bits.

3.5.2. Mathematical Simulations

The original HDR image is cropped and subsampled to the size we are aiming to implement in the hardware version, which will be a low resolution frame, a QCIF image (144x176 pixels). The final radiance picture format (.hdr) is then imported by the Matlab® Image Processing Toolbox engine in double data type. Then, it is converted to monochrome resulting in pixels values between $391 \cdot 10^{-6}$ and 29.36. This data imply a dynamic range of about $20 \cdot \log_{10}(\frac{29.36}{391 \cdot 10^{-6}}) = 97.5 \text{dB}$. In order to study the algorithm in higher dynamic range, the data are linearly ranged to fit in a set of data between 10^{-9} (for 29.36) and 10^{-15} (for $391 \cdot 10^{-6}$). Therefore, it will result in the same relative distribution between photocurrents but extended to the DR of choice (120dB). These artificial photocurrents values, which are scaled pixel data, can be considered as a representation of the pixels photocurrents because the multiexposure LDR frames has been linearly captured.

Representing images using a gray colormap is usual for monochrome images. However, this will lead to loss of details in the several possible representation of this work combined with the vision system of the reader. Therefore, figure 3.13 shows the photocurrents image in gray colormap and in jet colormap for the improvement of details perception. The jet colormap performs a rainbow set, which goes from blue assigned to lower values to red in the higher values passing through the colors green, yellow, and orange. From now on, the images will be shown in both colormaps with a right side bar indicating the range of the colors (colorbar). Figure 3.14(a) shows the histogram of the photocurrent images, however little can be distinguished in the linear representation of the image and its histogram due to the high differences between values. In order to achieve a proper illustration of the distribution of illumination zones, figure 3.14(b) shows the histogram of the logarithm of base 2 of the photocurrents, which is a typical representation in HDR histograms as it corresponds with the distribution of EV. It is noticeable that there exists 3 zones, the left zone corresponds to the background, the middle zone is the lamp holder and surroundings that reflects high amounts of light, and the right zone is the light bulb that emits the

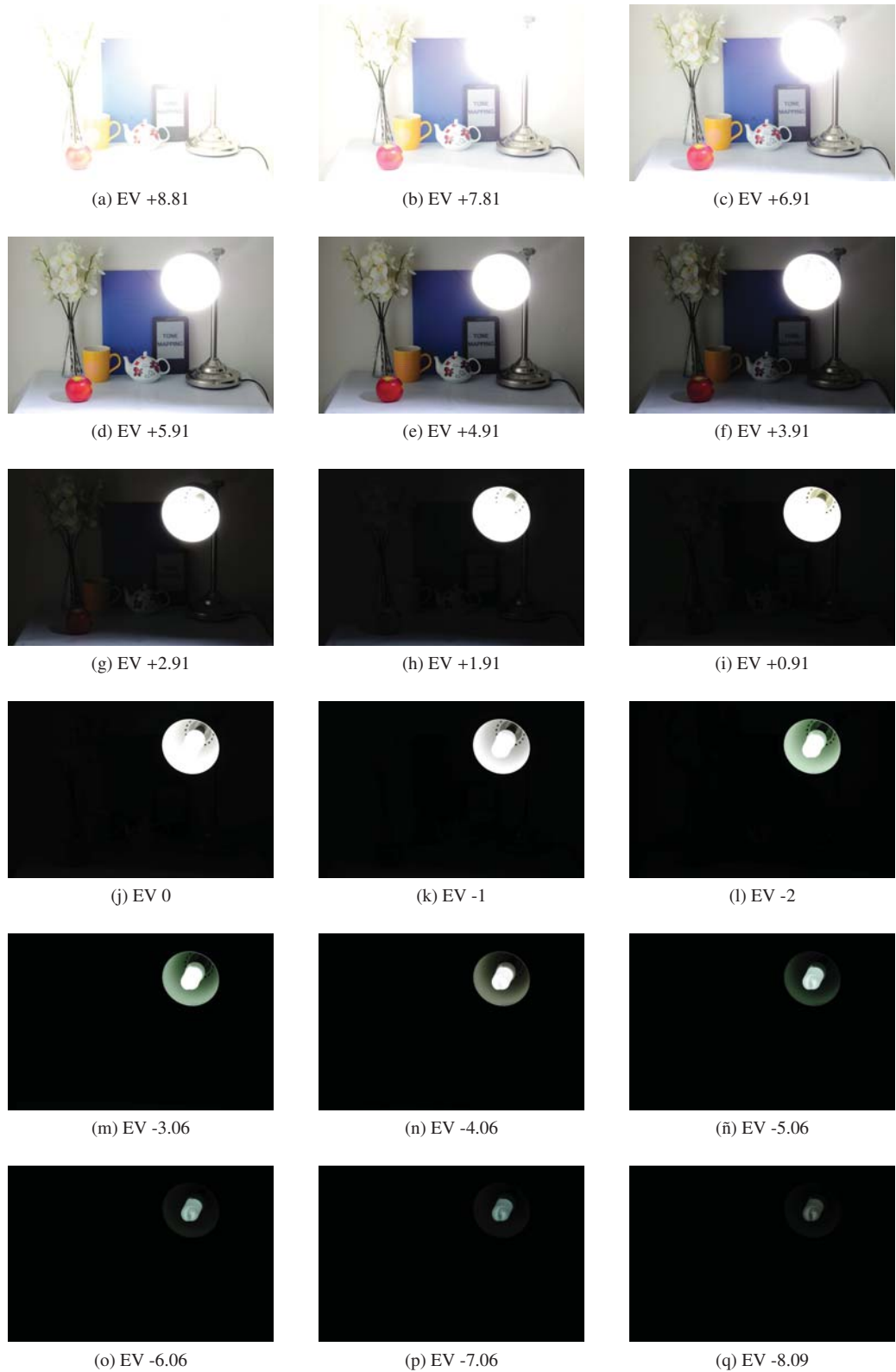


Figure 3.12: LDR frames used for HDR composition.

3. TONE MAPPING ALGORITHM

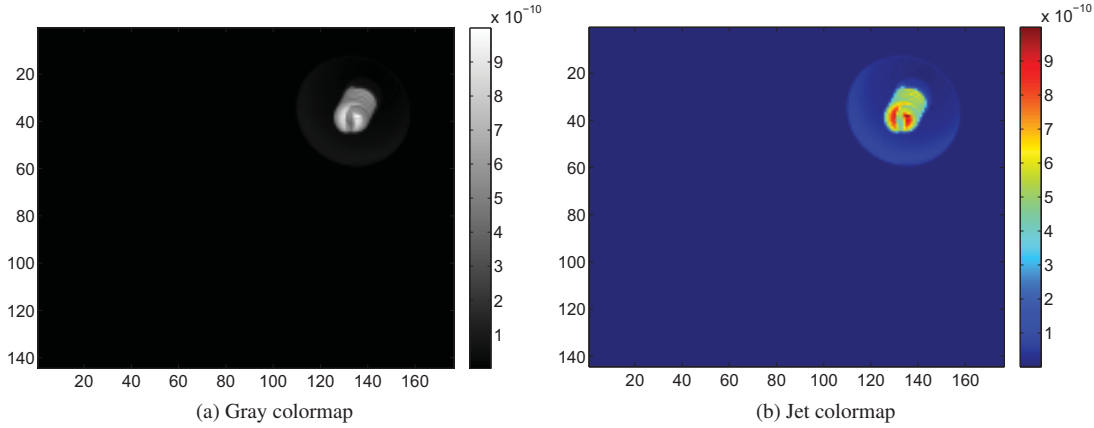


Figure 3.13: Photocurrent image.

light.

In order to apply the algorithm to this image, the discharge signals of the pixels are simulated taking the values of the HDR image (with values from 10^{-9} to 10^{-15}) as photocurrents. The discharge is evaluated via equation 3.4. The capacitor in the photodiode (C_{ph}) node is assumed to be 30fF and V_{rst} is 3.3V. The fixed reference voltage during the time bins is 1V. Therefore, during the last bin a voltage ramp from 1V to 3.3V has been simulated. The TSC curve has 4-bits for representation, achieving 16 bins, and TMC curve has 8-bits resulting in a final tonemapped image of 256 gray levels.

Once, the image containing the 4-bits TSC data is obtained, the histogram of TSI is calculated to obtain how many pixels cross in every bin. The TSI is shown in figure 3.15 and its histogram in figure 3.16, where the commented 3 zones are noticeable.

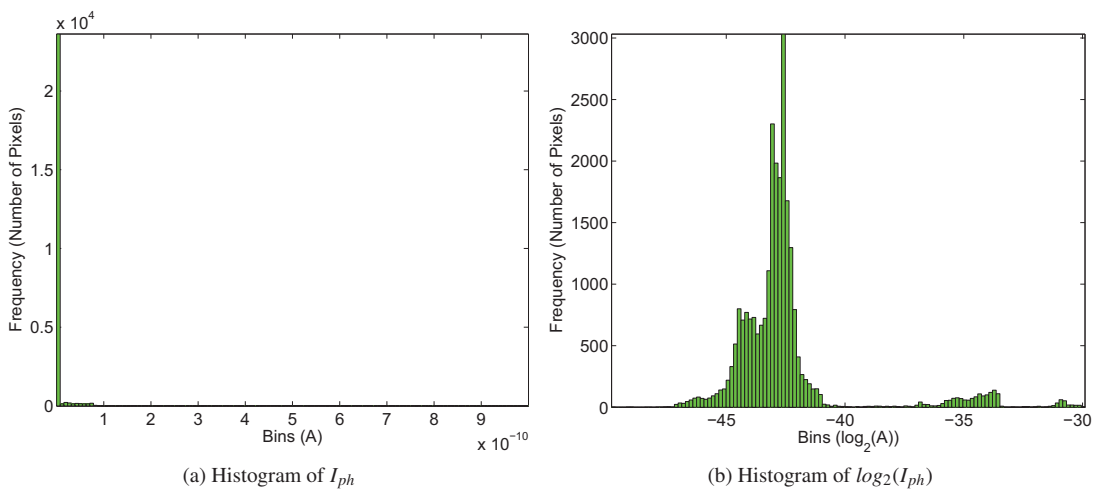


Figure 3.14: Photocurrent image histogram.

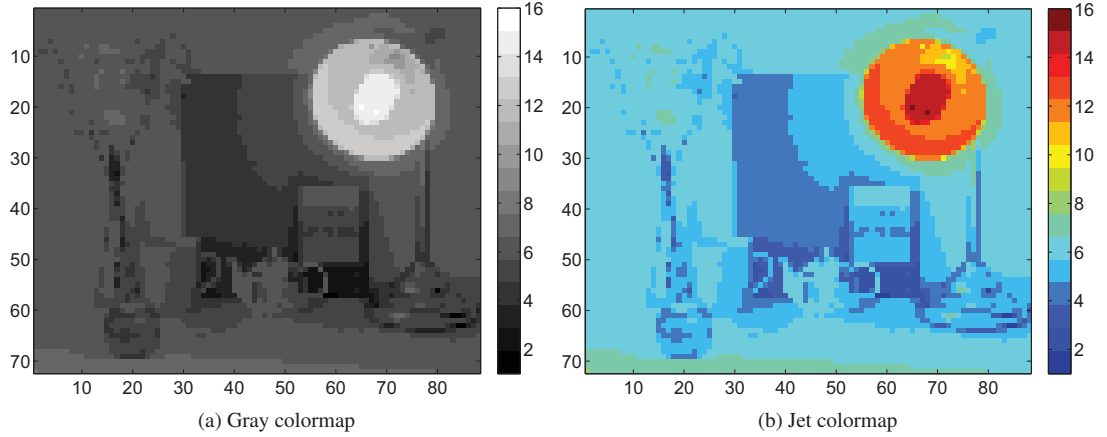


Figure 3.15: Time Stamp Image.

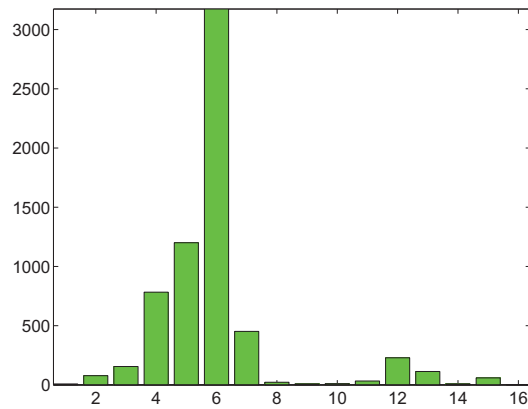


Figure 3.16: TSI Histogram of photocurrent image.

Next, the levels per bin must be calculated. This calculation is shown in table 3.4. We have divided the numbers of pixels in the bin by the constant $= \frac{\text{Number of pixels in TSI}}{2^{\text{bits of TMC}}}$. As the histogram image is subsampled by four in a QCIF image, the amount of pixels in the TSI is $\frac{144 \cdot 176}{4} = 6336$. Then, this quantity is divided by the number of total levels considering a final representation of 8 bits, $\frac{6336}{256} = 24.75$. Now, the result of the division of the number of pixels in the bin by 24.75 is floor rounded. However, the remainder of that division is kept to be sorted. The sum of the assigned levels will lead to only 250 levels. The lost levels will be $256 - 250 = 6$. These levels will be equally distributed in the bins with 6 highest remainders. The increment of levels in table 3.4 is highlighted by the use of bold letters. Now, we have a total of 256 levels assigned.

Tone Mapping Code composition over time is now achieved by the retrieval of the LUT data, which contains the evaluation times when the TMC decrements occurs. The LUT, in this case, have 257 (256 plus 0) positions of 256 bits. The information of the LUT is retrieved in every bin depending on the

3. TONE MAPPING ALGORITHM

Bin Index (TSC)	Pixels in bin	$\frac{\text{Number of pixels}}{24.75}$	Floor rounded levels	Remainder	Levels per bin
0	8	0.32	0	8	0
1	78	3.15	3	3.75	3
2	155	6.26	6	6.5	6
3	783	31.63	31	15.75	32
4	1200	48.48	48	12	49
5	3174	128.24	128	6	128
6	452	18.26	18	6.5	18
7	22	0.89	0	22	1
8	9	0.36	0	9	0
9	10	0.4	0	10	1
10	33	1.33	1	8.25	1
11	229	9.25	9	6.25	9
12	113	4.57	4	14	5
13	9	0.36	0	9	0
14	59	2.39	2	9.5	3
15	2	0.8	0	2	0

Table 3.4: Levels per bin calculus for photocurrent image.

assigned levels per bin. The final curve is shown in figure 3.17 where the x-axis is not the time but the subdivisions where the evaluation of the intersection takes place ($16 \text{ bins} \cdot 256 \text{ evaluations} = 4096 \text{ total evaluations}$) and the red vertical lines indicate the boundaries between bins.

Finally, the simulation of the discharge of the pixels with the values of the photocurrent image is repeated. If an intersection takes place in a subdivision of a bin, the TMC code present at that time is assigned to the pixel obtaining the final image.

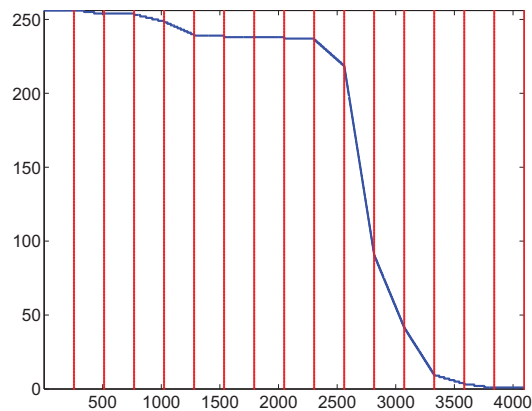


Figure 3.17: Tone Mapping Curve vs. evaluation subdivisions for photocurrent image.

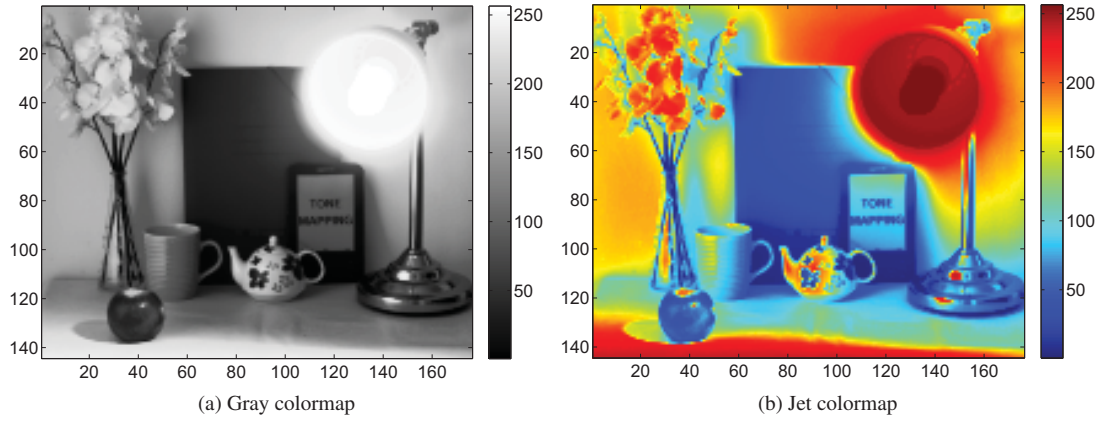


Figure 3.18: Final Tone Mapped Image.

The final Tone Mapped Image (TMI) is shown in figure 3.18. It is noticeable that, the objects present in the scene can be distinguished. The objects can be distinguished in low light (teapot, cup, flowers, etc.) and in medium light (apple) illumination environments. Moreover, if the representation in jet colormap is observed, the objects in very high light zones have been captured, which are the shape of the light bulb as well as its surrounding area in the reading lamp. Additionally, the final image will tend to have an equalized histogram. However, this algorithm assumes that the population of pixels in a bin is equally distributed inside this bin provoking an imperfect equalization. The histogram of the final image is shown in figure 3.19.

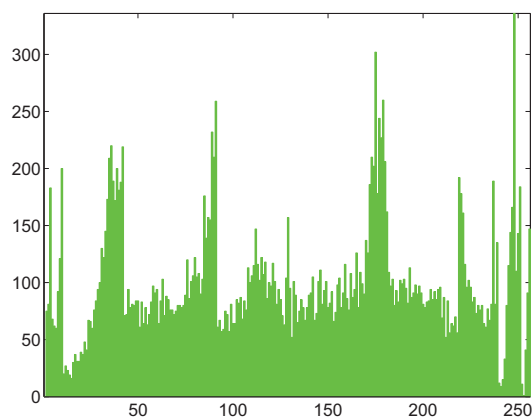


Figure 3.19: Histogram of the Tone Mapped Image.

3.6. Algorithm Comparison

To the knowledge of the author, it has not been reported another tone-mapping hardware-aimed algorithm for focal-plane implementation. Hence, the comparison with an algorithm of the same characteristic is not possible at this moment. Instead, a comparison with computer graphics tonemapping techniques is presented.

Different tonemapping techniques have been applied to the HDR composed image using the open source program Luminance HDR [114]. This program performs the tone mapping of HDR images in radiance format (.hdr) applying some tonemapping algorithms [115]. The references for these algorithms are:

- Mantiuk'06: A perceptual framework for contrast processing of high dynamic range images [116].
- Mantiuk'08: Display adaptive tone mapping [117].
- Fattal: Gradient domain high dynamic range compression [108].
- Drago: Adaptive Logarithmic Mapping For Displaying High Contrast Scenes [88].
- Durand: Fast bilateral filtering for the display of high-dynamic-range images [105].
- Reinhard'02: Photographic tone reproduction for digital images [102].
- Reinhard'05: Dynamic range reduction inspired by photoreceptor physiology [90].
- Ashikmin: A tone mapping algorithm for high contrast images [101].
- Pattanaik: Time-dependent visual adaptation for fast realistic image display [118].

The final tone mapped images of these algorithms are shown in figure 3.20. The result images are presented in color as they are intended for human perception in common reproduction systems, which is a fair representation of their results. However, it is difficult to perceive some details, as the background has been very compressed in order to have more details in the light bulb area. Moreover, it is not easy to compare with our tonemapping algorithm, which results in a monochrome image. Therefore, these images have been converted to a monochrome representation. Then, in order to enhance the contained details, Contrast-Limited Adaptive Histogram Equalization (CLAHE) [119] of the Image Processing Toolbox of Matlab[®] has been applied to the images and also to the result of the new algorithm presented in this work. The parameters of the CLAHE are 8x8 Tiles, Clip Limit of 0.01, 256 bins, full range and uniform distribution. This images are shown in figure 3.21.

It is noticeable that some algorithms create artifacts in the surroundings of the reading lamp: Mantiuk'06, Reinhard'02, Ashikmin and, at a lower level, Pattanaik. Parallely, the new algorithm presents a halo artifact in the surroundings of the reading lamp due to the high compression applied to this zone. In general, all the computer graphics tonemapping algorithms have applied less compression in the lamp and therefore light bulb has more details than our algorithm. Regarding the background, our algorithm has more details than the rest in the objects in the low-middle illuminated zones. This fact is due to



Figure 3.20: Results from Luminance HDR.

the high compression in the high illuminated zones, which gives more levels to represent the rest of the scene. In our algorithm, the tonemapping curve is monotonic, as darkest areas correspond to lower pixel values and brighter areas correspond to higher pixel values in the entire image. It is to say, two pixels receiving the same photocurrent cannot have different pixel values, which can happen in local tonemapping algorithm. Therefore, there is a tradeoff between the compressions of the different zones as the number of levels for representation are limited, because it is a global tone mapping operator.

In comparison with general tone mapping algorithm, this new algorithm will not generate an a priori better result. These algorithms are intended for computer graphics. Consequently, the hardware limitations are inexistent or not affected when compared with our approach. The new algorithm is computationally simple and requires neither intensive memory accesses nor memorization of many images.

3. TONE MAPPING ALGORITHM

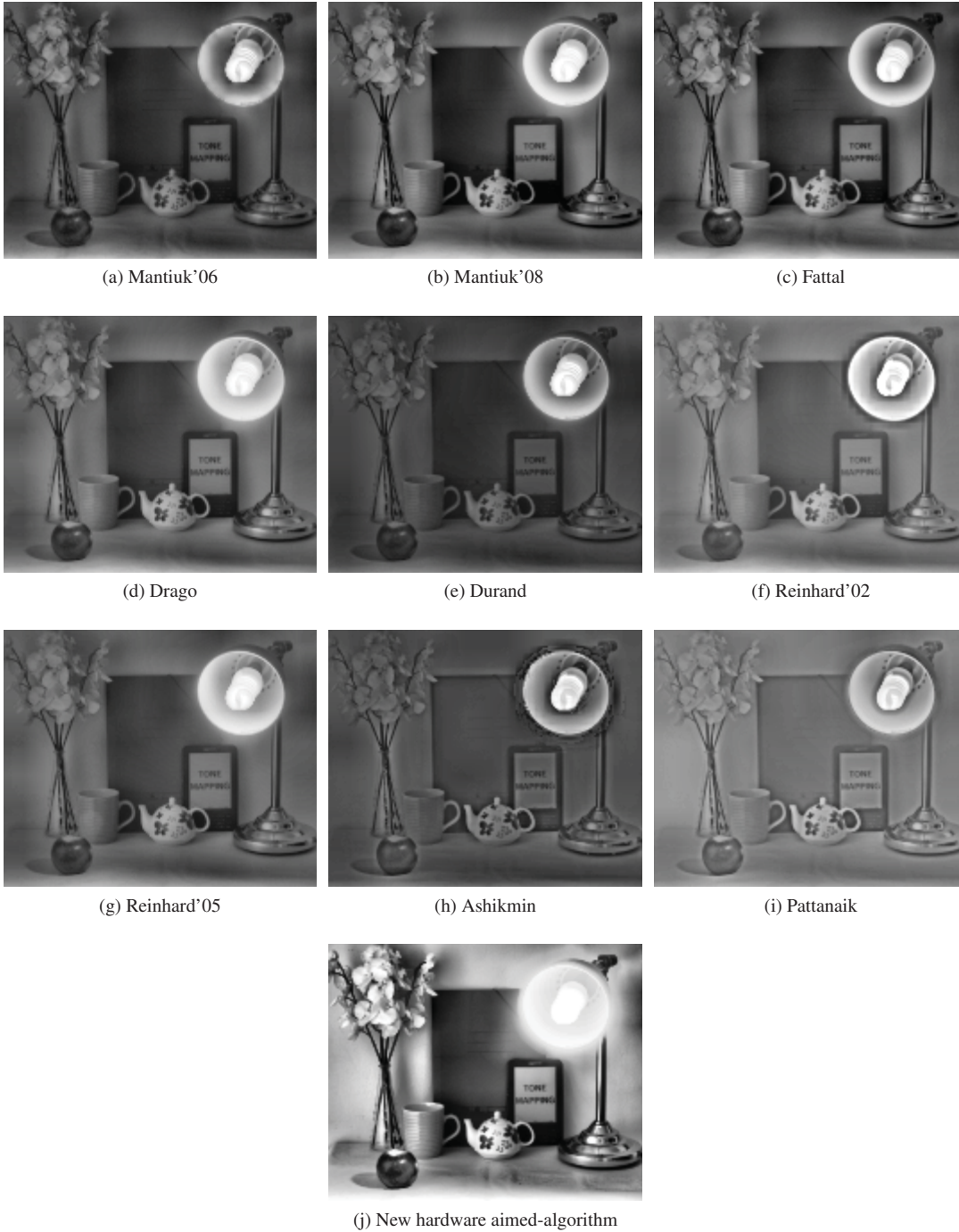


Figure 3.21: Results after the application of CLAHE.

Tone mapping algorithm usually realize very complex computation to be implemented at a system on chip (such as filter kernel, FFT or complex photoreceptor model functions, etc.). On the contrary, our algorithm requires in-pixel analog comparators, in-pixel memory, comparative accumulation for the histogram, a very limited amount of divisions by a constant value and a look up table. These requirements are not computationally expensive and can be implemented on a single chip. Summarizing, the new algorithm is intended for further processing, not for representation, and therefore the result presented here is valid for our purposes. Moreover, the algorithm can be modified to enhance the details in selected areas.

In other to enhanced different details, the different modes of assigning levels per bin have been applied to the photocurrent image. The results are shown in figure 3.22, where the different tone mapping curves that have been applied (x-axis are evaluations subdivision) are also included. It can be observed that most of the modes enhance the high light areas compared with previously used mode 2. However, the number of levels for image representation is limited. It is a tradeoff, which caused the rest of areas to lose some contrast in low-medium light areas in comparison to mode 2. Mode 6 seems to have the best enhancement in the high light area. In order to illustrate the improvement, the CLAHE enhanced version of mode 6 is presented in figure 3.23 again in comparison with the previously presented CLAHE enhanced computer graphics tone-mapping results.

3.7. Conclusions

A tone mapping algorithm has been successfully developed. The algorithm is capable of representing an HDR scene with a reduced amount of data by applying different compressions to different zones. Moreover, this compression of the zones will adapt depending on the distribution of illuminations in the scene. This makes the algorithm not only applicable to still image but also to video as it adapts to changes of the HDR scene. Despite considerable progress in HDR image tone mapping algorithms for still image in the past decade, little work has been done for HDR video. The present algorithm presents a real-time video tone mapping capability since our algorithm adapts to extreme changes of illumination distribution in only one frame.

As already mentioned, the proposed algorithm is not intended for human perception of details in displays, which is the general objective of computer graphics tone mapping operators. If this were our objective, the display output characteristic and the human perception should be considered. Nevertheless, our algorithm proposes an a priori compression toward higher illumination levels, which is consistent with the natural behavior of the human eye. This fact makes our algorithm suitable for human perception in displays but with limitations because it is not optimized for such objective.

The algorithm have been developed adapting the most basic ideas of tone-mapping techniques to the possibilities of focal-plane circuitry. Although tone-mapping techniques have not inspired the functionality, the hardware limitations have led us to some parallelism with the Ward Histogram Adjustment Operator; a histogram is performed over a subsample density image. In our case, the compression applied to obtain the density image is not logarithmic (Ward's Operator) but a combination of an inverse

3. TONE MAPPING ALGORITHM

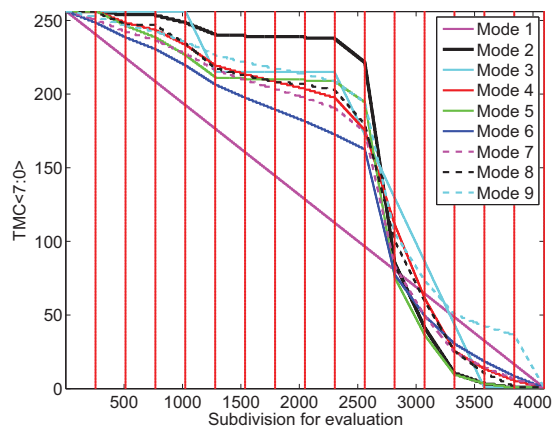
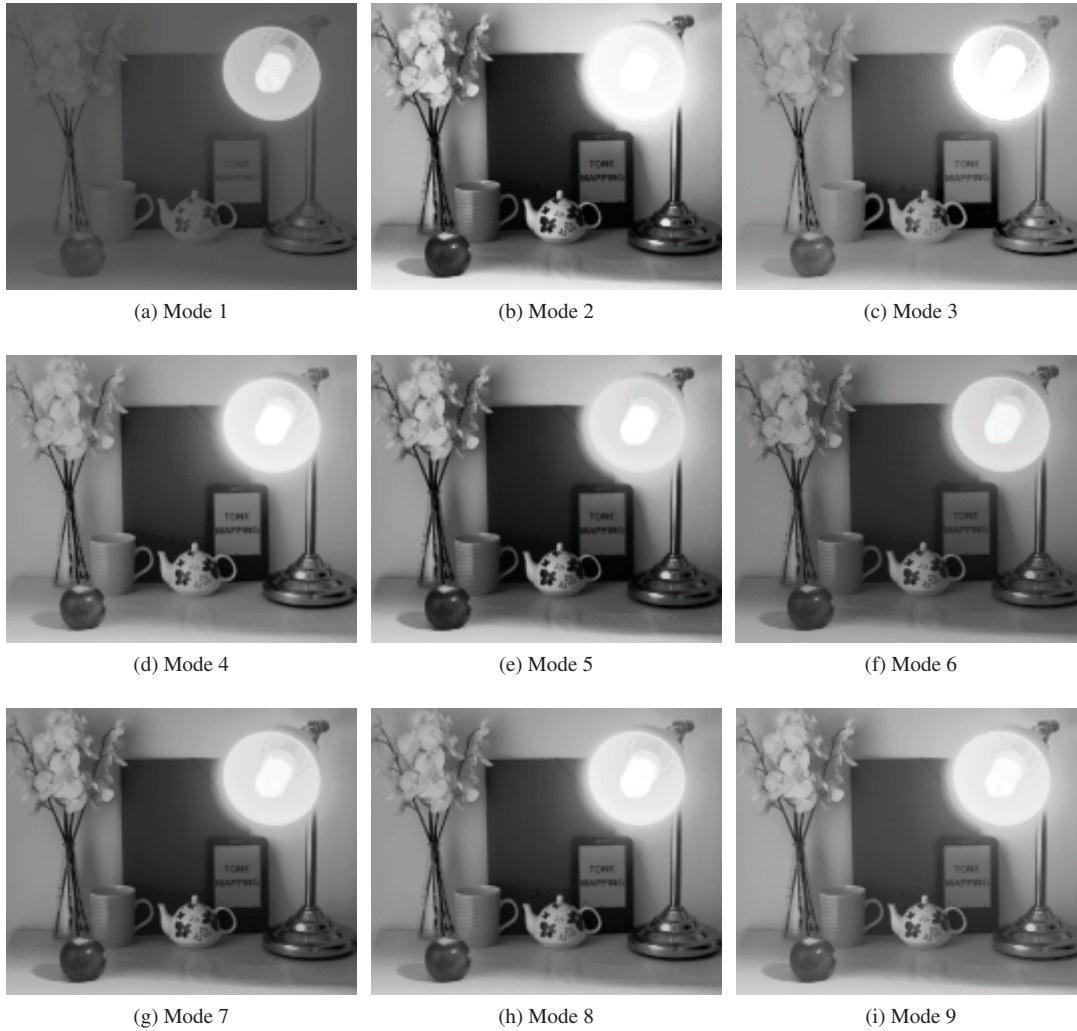


Figure 3.22: Results of applying the different levels per bin assignment modes (a-i), and (j) their corresponding tone-mapping curves.

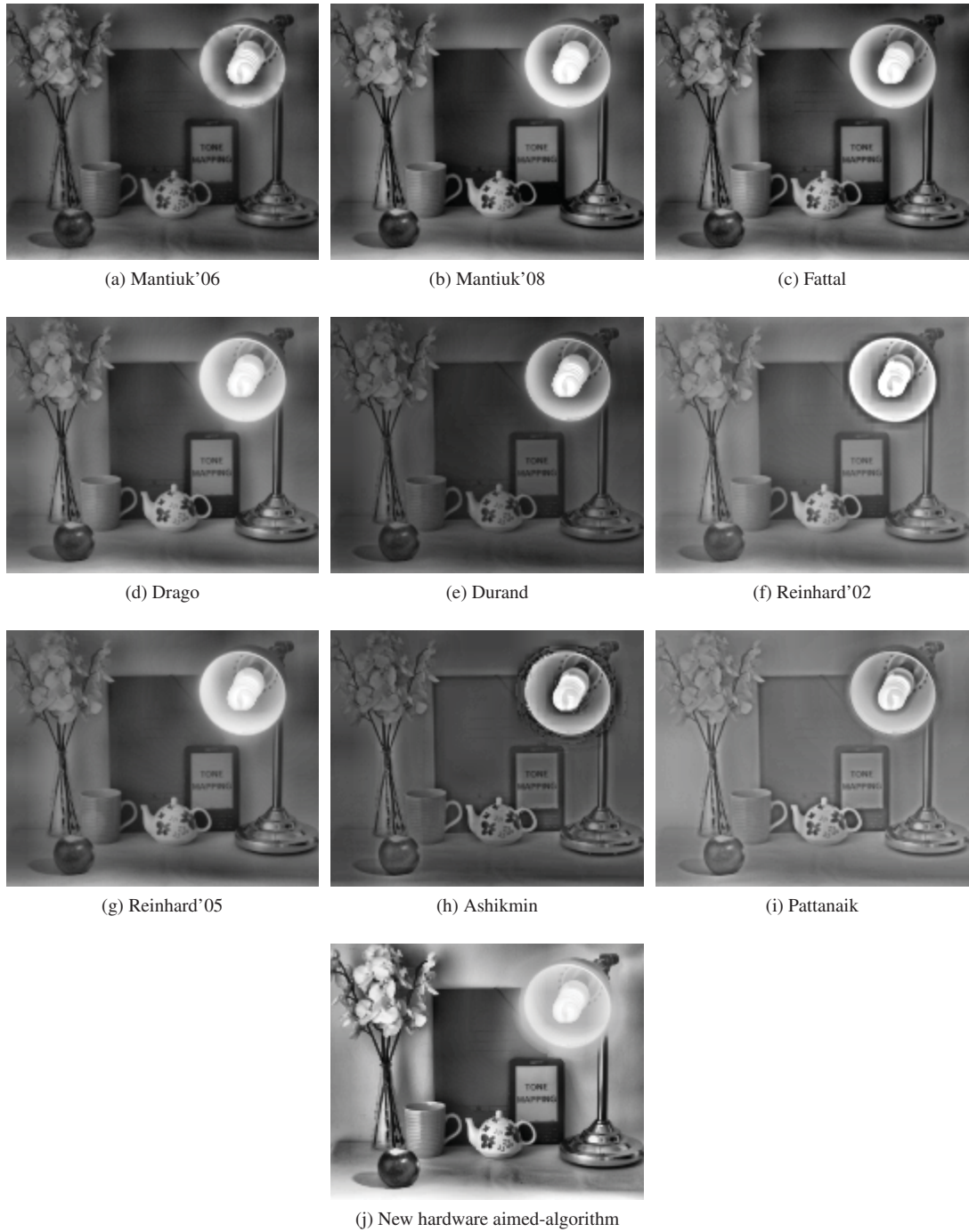


Figure 3.23: Results of CLAHE in comparison with mode 6.

3. TONE MAPPING ALGORITHM

function and a pseudo-linear function. Moreover, the result of the cumulative histogram is also used to calculate the global tone mapping curve of compression.

This tone mapping algorithm is suitable for on-chip implementation taking advantage of the reduced amount of computation that it is required. Since only comparison with an analog signal and storing capabilities are required inside the pixel, it is possible to implement part of it as focal plane circuitry. The parallelism of operation, which is allowed by focal-plane circuitry, makes this algorithm especially well-suited for focal plane processors or 3D integration technologies in future approaches. Therefore, the algorithm minimizes the amount of data in the representation of HDR scenes for later use in post-processing, and therefore cutting down the necessary resources for these future evolutions.

The algorithm itself does not have a limited dynamic range to handle or an established number of bits for pixel representation. However, a low compression will take place in low dynamic range scenes. Similarly, low compression will occur if a high number of bits is chosen for image representation. The amount of bits in the image representation will be defined by a trade-off. This will depend on the hardware implementation, which implies physical limitations, and the application, which determine the limits of the target representation. The physical limitations are derived from characteristics such as, noise, area of the pixels, transistors mismatch, etc.

Compared with computer graphics tonemapping, our algorithm compresses the steps necessary in a normal tone mapping operation, as the pixel data is captured already tone mapped. Whereas in computer graphics tonemapping, the necessary steps will be capturing several images with a camera, storing, downloading to PC or system, and the application of the algorithm in the computer. In our case, we do not need to move unnecessary data, as the information is already captured and stored at pixel level.

Chapter 4

TVHC: A HDR Tone Mapping Imager in Standard CMOS Technology

4.1. Introduction

This chapter describes the so-called Time Voltage Histogram Camera chip, or simply TVHC, a QCIF resolution HDR imager, which implements the tone-mapping algorithm for DR improvement, which has been described in chapter 3. Clearly, the differences between a simulation environment like Matlab[®] and the mixed-signal on-chip implementation of our algorithm, will force some adaptations in the real hardware implementation. Up to now, the algorithm has been designed and simulated by the infrastructure provided by Matlab[®], with floating point (nearly unlimited to our purposes) accuracy, huge memorization resources, and not paying too much attention to the attainable frame rate (as we basically dealt with still images). Unfortunately, the limitations in the software implementation are loose, compared with the available resources to map this algorithm onto a chip. However, the fully parallel Single Instruction Multiple Data (SIMD) architecture of focal plane processors also offers some advantages. Thus, for instance, the chip will benefit from the parallel processing capabilities in the comparison with V_{ref} . In this case, instead of continuously downloading the image and comparing every read pixel with the current value of V_{ref} , we will get these NxM comparisons on-pixel and in continuous-time, without the need for a complete frame readout. This prevents the large evaluation errors to appear in the highest photocurrents, which might arise from different readout times between pixels (similar to what happens when we described the rolling shutter option in chapter 1). On the other hand, the kind of computation that can be performed within a pixel is constrained by the maximum allowable power consumption (directly proportional to the number of pixels), pitch (that, in general, increases as more resources are included on pixel), and accuracy (that, if limited by mismatch, requires using larger devices or calibration blocks, with area penalties in both cases). In addition to that, we must bear in mind that there is a need to provide an aperture in the metal layers to capture as much light as possible in order to maximize sensitivity. This creates further limitations for possible routing among

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

blocks inside the pixel.

The TVHC has been conceived as a complete Vision-System-on-Chip (VSoC). However, we have to remark that this chip is a proof of concept prototype (i.e. the first prototype with this tone mapping algorithm). Therefore, for versatility and security purposes, some non-critical parts of the algorithm have been implemented externally (in an FPGA), as it is shown in chapter 5. This includes the calculation of the levels per bin and the LUT. Thanks to this flexible implementation, the whole TVHC system (TVHC camera + host board) can operate with several LUTs, if necessary, and can change the method for the assignment of levels per bin on the fly. The FPGA's architecture and program, along with the rest of the support circuitry will be explained in chapter 5, where the whole test infrastructure will be described. The system within the FPGA has been coded in Verilog and it includes limited use of internal FPGA resources (a memory and a divider). This may produce a non-optimized code for the FPGA synthesis but guarantees that the same code can be synthesized [120] with reduced modifications in the selected integration technology in future improved versions of the chip.

The chapter is organized as follows: First, the selected integration technology with some optical sensor enhanced features is introduced. Second, the general architecture of the chip is explained along with the functionalities of its blocks. Third, a description of the operation of image capture is included. Finally, there is a general description of tone mapping systems in comparison with our approach.

4.2. Fabrication Process

The present chip has been designed using the Austriamicrosystems (AMS) CMOS 4M/2P 0.35 μ m OPTO technology. This technology is based in the standard AMS 0.35 μ m (C35B4C3) technology, and includes some opto-flavored modifications to provide enhanced light sensing characteristics:

- Inorganic Anti-Reflecting-Coating (ARC) layer
- 14 μ m EPI substrate

Conversely to the usual approach of depositing the ARC layer just on top of the photosensor, this technology adds the inorganic ARC on the top of the passivation layers [121]. In any case, the use of ARC layers results in a higher and smoother spectral response of the sensor. On the other hand, the EPI layer diminishes the dark current. These variations in combination with the base technology make the process to feature these characteristics [122][123]:

- High Sensitivity \Rightarrow Responsivity of Photodiode¹ at 550nm = 290mA/W.
- Low Dark Current \Rightarrow Dark Current of Photodiode < 45pA/cm² (at 27°C).
- Low cost \Rightarrow Only 1 additional process step (Inorganic ARC Layer).
- Capacitor per area for Photodiode: 0.08fF/ μ m.

¹For N-well to P-substrate structures.

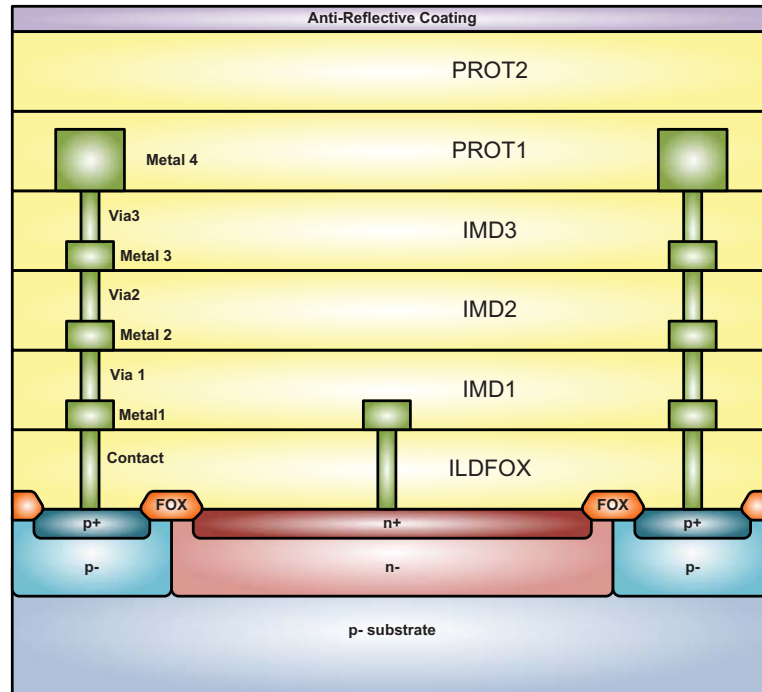


Figure 4.1: Photodiode layer arrangement.

- Minimum Pixel Size: $6\mu\text{m}\times 6\mu\text{m}$.
- Poly-Insulator-Poly (PIP) and Metal Insulator Metal (MIM) Capacitors.
- 4 metal layers.
- Supply Voltage: CMOS 3.3V with optional 5V IOs.
- High driving capabilities and ESD reliable peripheral cells available.
- Process compatibility with mixed base technology, and therefore possible reuse of designs.

Although this technology offers limited modifications for enhanced optical sensors comparing with the UMC $18\mu\text{m}$ CIS technology used in the SCU chip, it does not limit the kind of circuitry that can be allocated next to the photosensors. The recommended physical sensor of this technology is an Nwell-Psubstrate photodiode. This photodiode has a wider depletion region due to the lower doping and a deeper junction than the previously employed (SCU chip) Ndiffusion-Psubstrate structure. However, Nwell layers usually require higher spacing from surrounding circuitry than conventional N+ (Ndiffusion) implants. Therefore, this kind of photodiodes will always result in a larger pitch (for the same technological node). The cross section of the photodiode with the ARC layer is shown in figure 4.1, where FOX is the field oxide, ILDFOX is the oxide between the polysilicon layer and Metal 1, IMD are the Inter-Metal Oxides and PROT indicates passivation layers.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

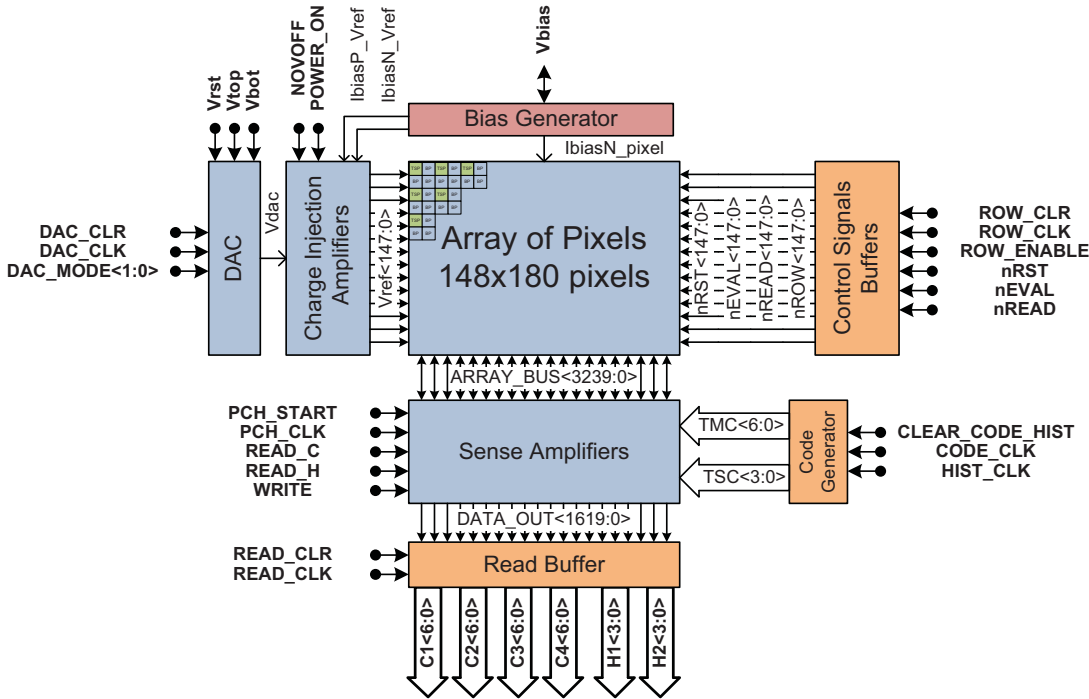


Figure 4.2: TVHC: High-Level Block Diagram.

4.3. Architecture of the TVHC Chip

As in any other imager, the core of the TVHC prototype is an array of pixels whose functionality is supported by many peripheral blocks. These blocks, among many other things, are responsible for biasing, generation of analog references and digital controls, and high-speed image transference. The chip contains all the necessary circuitry to acquire the Tone Mapped Image (TMI) and the Time Stamp Image (TSI) as well as downloading them, already digitized, through a high-speed 36-bits bus (nearly 343Mbits/s¹). The high-level block diagram of the chip is shown in figure 4.2.

The main blocks included in this diagram are:

- Array of Pixels: an array of pixels with QCIF (144×176) resolution. The array contains a border of 2 dummy pixels (i.e. it has 4 additional rows and columns, two on each side), which provide identical surroundings to all pixels in the array.
- DAC: A Digital to Analog Converter generates the analog voltage reference V_{ref} for the pixels.
- Charge Injection Amplifiers: This block distributes the signal generated by the DAC.
- Control Signals Buffers: They drive the digital controls signals of the pixels row by row using a Clock-Tree generated distribution of digital buffers, which guarantee minimum skew (specified at

¹ $\frac{36\text{bits}}{100\text{ns}} \cdot \frac{1}{1024 \cdot 1024} = 343.3\text{Mbits/s}$

a maximum of 300ps when using the automatic Place and Route [P&R] tool) and precise control timing.

- Sense Amplifiers: to acquire TSI and TMI outputs from the pixels. Since the digital codes of the Tone Mapping Image (TMI) and the Time Stamp Image (TSI) are stored in static 6T memories [124] within the pixels, a direct readout of these pixels could produce data flipping on the SRAM (and consequently, a loss of information) due to the high capacitive load of the vertical metal lines that are used to access the pixels. The sense amplifiers are responsible for precharging the column buses at a proper voltage level, retrieve and store the information from the pixels. All the pixels in a row are read at the same time.
- Read Buffer: This is a one row buffer memory, which allows fast image retrieval. We have employed a double buffer topology, such that when we are receiving a row in the Sense Amplifiers block, we are downloading the previous row from the Read Buffer into the FPGA in the TVHC board.
- Code Generator: It generates the digital signals $TMC\langle 6:0 \rangle$ and $TSC\langle 3:0 \rangle$ to the pixels. These buses use gray coding in order to diminish switching at the pixel level to only one bit per write cycle (thus saving a considerable amount of power).
- Bias Generator: This block generates the required biasing for all the analog circuits of the chip. It uses a Proportional To Absolute Temperature (PTAP) current source as reference for a bank of current mirrors, which generates 3 different scaled copies of this reference current (I_{biasP_Vref} , I_{biasN_Vref} , I_{biasN_Pixel}). This block also provides an output voltage V_{bias} to the host board. This voltage can be employed to monitor the temperature of the chip (something that is very useful when adapting the levels of the A-to-D conversion for the attenuation of the effects of the dark current in the output image). Furthermore, this V_{bias} node can be used as startup node for the bandgap reference [125]. The reference current is produced by a bandgap circuit [30], which employs 2 parasitic vertical bipolar transistors. The operation of this subsystem is very well known and documented in the literature and its description is not considered necessary here.

4.4. Pixels

In order to realize the tone mapping algorithm, the pixel must contain an integration node V_{ph} , which is discharged by the photogenerated current. This V_{ph} signal must be compared with an analog reference V_{ref} . When the intersection of the signals V_{ph} and V_{ref} takes place, the values of the Tone Mapping Code (TMC) (7-bits bus) and the Time Stamp Code (TSC) (4-bits bus) must be stored. Two different pixels have been designed:

- Basic Pixels (BP): acquire only the Tone Mapped Image (TMI) $\rightarrow TMC\langle 6:0 \rangle$.
- Time Stamp Pixels (TSP): acquire Tone Mapped Image (TMI) and Time Stamp Image (TSI) $\rightarrow TMC\langle 6:0 \rangle$ and $TSC\langle 3:0 \rangle$, respectively.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

The block diagrams of the BP and TSP are shown in figure 4.3(a) and (b), respectively. Note that the general structure of the pixels is the same but the TSP contains four more static RAM (SRAM) cells to store $TSC<3:0>$.

The integration node V_{ph} is discharged by the current, which is generated by a $3 \times 3 \mu\text{m}^2$ Nwell/Psub photodiode. The photodiode is connected to the analog comparator via a buffer in order to isolate the comparator's kickback noise [126], which is mainly produced by the multiple changes in the voltage reference V_{ref} .

The reset of the photodiode is performed by the feedback loop established by the buffer, the comparator and the reset switch P_1 . Thus, the operation of the circuit is offset-free in a first approach (see Section 4.4.2 for details).

The buffer and the input stage of the comparator (its output signal is amplified by subsequent digital drivers to create a rail-to-rail high gain output) share a common design of a 5T NMOS-input Single Stage

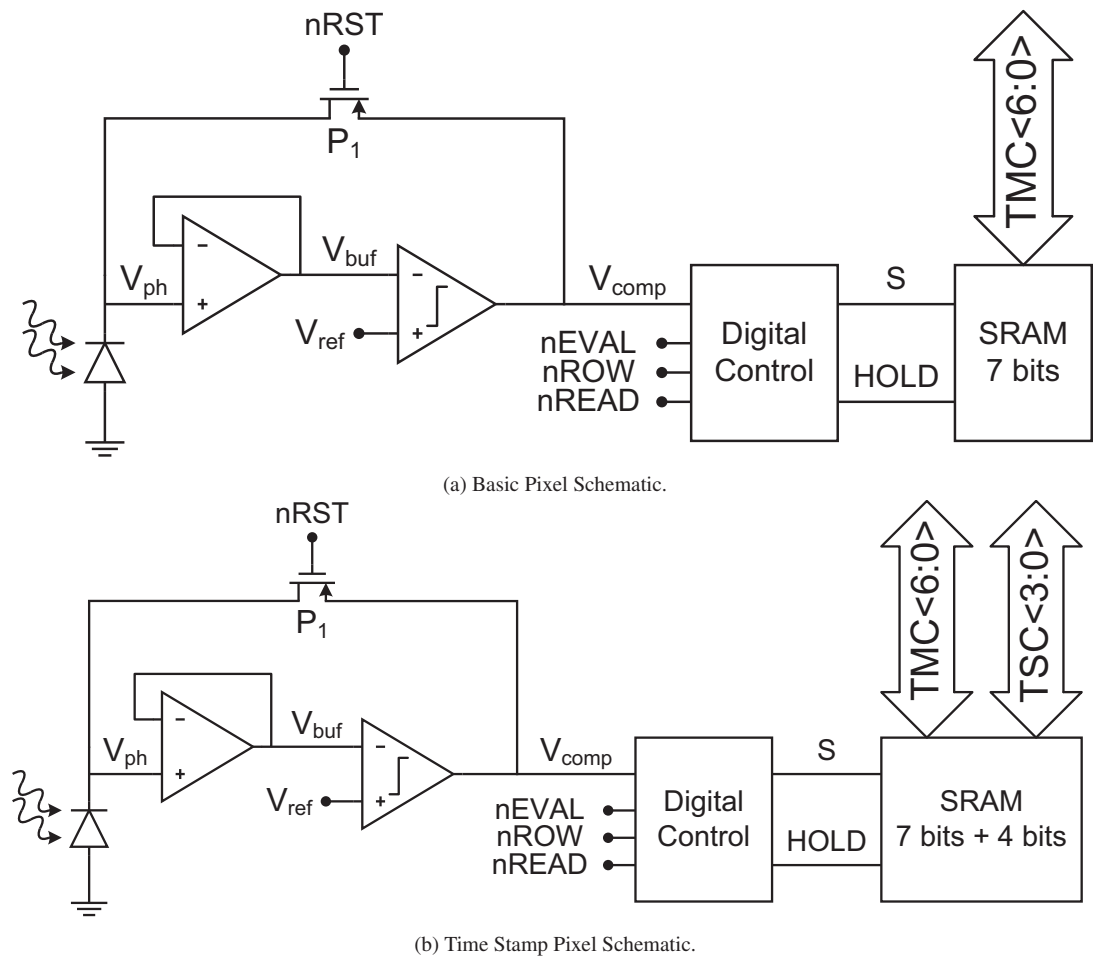


Figure 4.3: Block Diagrams of the Basic and Time Stamp Pixels.

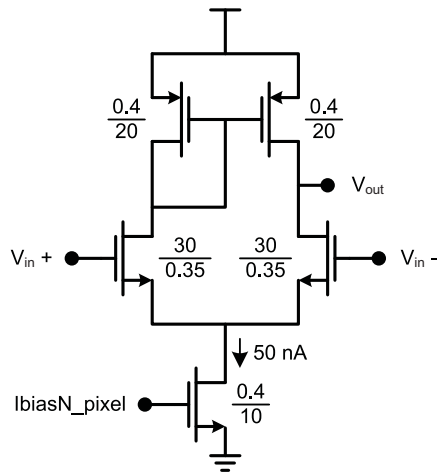


Figure 4.4: 5T Amplifier in the Pixels.

Differential Amplifier. The schematic of this circuit is shown in figure 4.4. Notice that the transistors are relatively large as compared to the minimum transistor that can be fabricated in this technology, $\frac{0.4}{0.35} \mu\text{m}$, as it is dictated by accuracy requirements (see section 4.4.2). Simulations show that the available input range of the buffer and comparator combined is 0.7-2.6V. Figure 4.5(a) shows the simulated DC Gain of the buffer and figure 4.5(b) shows the small signal gain at $V_{in}=1.65\text{V}$.

The output of the input stage of the comparator for several V_{ref} values is shown in figure 4.6(a) and a Monte Carlo simulation of the input-referred offset voltage of the 5T Amplifier for $V_{ref}=1.65\text{V}$ is shown in figure 4.6(b). From the simulation results, we obtain that this offset voltage has a mean value of $320\mu\text{V}$ and a standard deviation of nearly 3.40mV .

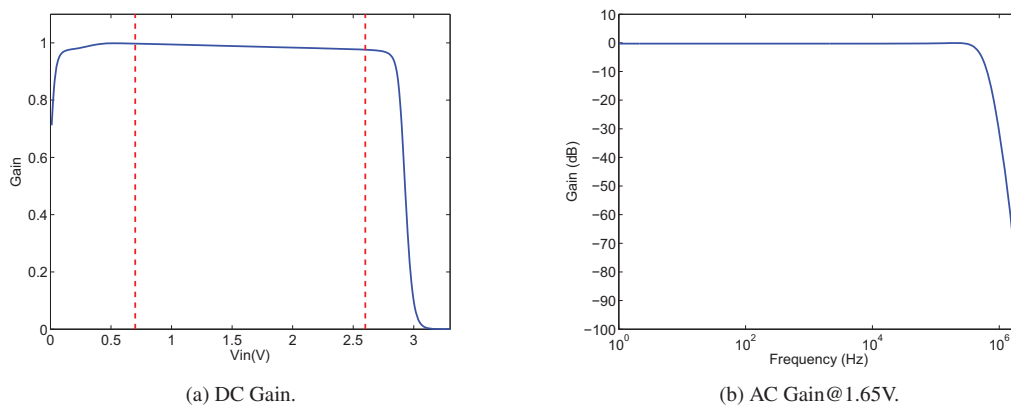


Figure 4.5: 5T Amplifier Gain Plots.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

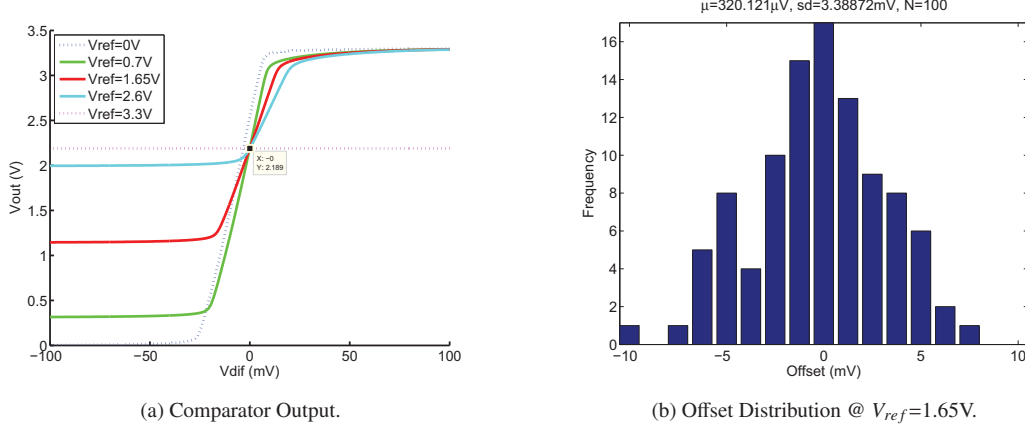


Figure 4.6: 5T Amplifier Comparator Simulations.

4.4.1. Pixel Digital Circuitry

The digital control circuitry within the pixels is responsible for two main functions: (1) it controls the read and write operations of the pixel SRAM cells, and (2) it amplifies the signal from the input stage of the comparator. In this action, the quiescent point of the resulting digital amplifier is optimized by considering the statistics of the possible outputs of the comparator in the chosen range of operation. Therefore, it has been tuned by means of electrical Monte Carlo simulations with several V_{ref} in the range [0.7 2.6].

Although we have mentioned many times that the pursued function is to store the values of TMC and TSC when V_{ph} intersects V_{ref} , what we have implemented in practice is a system that continuously samples the two buses until the intersection takes place. Here, as usually, we have a trade-off between area and power consumption, as the continuous monitoring requires more power (on average) but the sampled option requires one extra register to acquire the comparator's previous output (or a guaranteed delay line) plus additional digital control circuitry. In this design, we have opted for the area reduction option at expense of a higher power consumption (this was indeed the reason for introducing the gray coding in TM and TS codes). Hence, the write operation on the SRAM is directly controlled by the instantaneous output of the comparator (not including any previous sample of it) as in equation 4.1.

$$\begin{aligned}
 V_{ph}(t) = V_{rst} - \frac{I_{pix}}{C_{pix}} \cdot \Delta t < V_{ref} &\Rightarrow V_{comp_{dig}} = 1 \\
 V_{ph}(t) = V_{rst} - \frac{I_{pix}}{C_{pix}} \cdot \Delta t > V_{ref} &\Rightarrow V_{comp_{dig}} = 0
 \end{aligned}
 \tag{4.1}$$

This operation must be inverted so the write switches in the SRAM will be on while $V_{ph}(t) > V_{ref}(t)$. In the schematic of a SRAM cell in figure 4.7, all switches in the SRAM are minimum ($\frac{0.4\mu m}{0.35\mu m}$) NMOS transistors whereas transistors in the inverters (both NMOS and PMOS) are sized $\frac{0.7\mu m}{0.35\mu m}$.

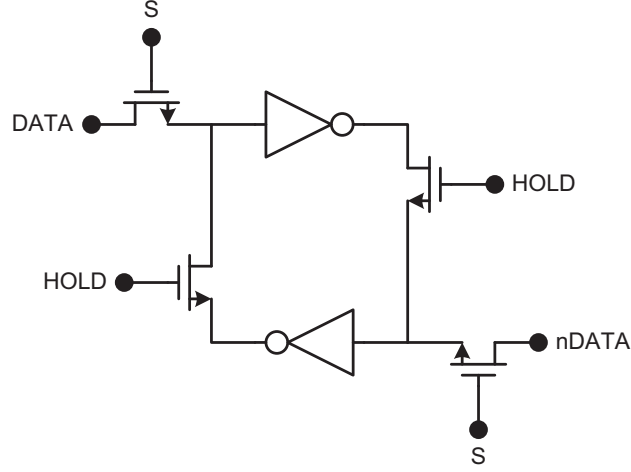


Figure 4.7: Pixel SRAM Cell.

The S switches allow store/retrieve data to/from the SRAM taking into account that $HOLD$ switches must be disconnected for store. These SRAM modules are written/read through the DATA and nDATA (negated version) inputs. To that purpose, the pixels are connected to the bidirectional lines TMC<6:0> and TSC<3:0> and their negated versions.

Write operations are easily performed by configuring $S = 1$ and $HOLD = 0$. This configuration avoids any conflict when flipping the previous data on the SRAM. The situation is not so simple for the read operation. SRAM readouts require $S = 1$ for granting memory access to the bidirectional buses, and $HOLD = 1$ to avoid data loses during the procedure. Besides, DATA and nDATA lines are precharged to a high level in order to avoid any accidental flipping of data in the SRAM due to the differences between the capacitance at the SRAM input nodes (internal capacitors of the memory) and the huge parasitic capacitance associated to the bus lines. Write operation to the SRAM can optionally be synchronized with external signal $nEVAL$ as shown in figure 4.8.

The storage of data row by row is also available (we can upload images to the chip for SRAM testing purposes) by properly using selection signal ROW (figure 4.8). This process can also be employed to write an initial value to all SRAMs at the beginning of the sensing process or just after power up. The logic diagram of the digital circuitry, which generates S and $HOLD$ is displayed in figure 4.8 and its logic behavior is detailed in equations 4.2 and 4.3.

$$\begin{aligned} \mathbf{S} &= \overline{\overline{\overline{(V_{comp} + nEVAL)} \cdot nROW}} = \overline{V_{comp} + nEVAL + nROW} \\ &= (\overline{V_{comp}} \cdot \overline{nEVAL}) + \overline{nROW} \Rightarrow EVAL \cdot \overline{V_{comp}} + ROW \end{aligned} \quad (4.2)$$

$$\mathbf{HOLD} = \overline{S \cdot nREAD} = \overline{S} + \overline{nREAD} \Rightarrow \overline{S} + READ \quad (4.3)$$

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

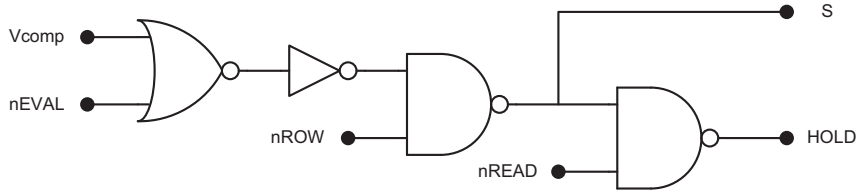


Figure 4.8: Pixels Digital Control.

Since write operations to the SRAM are not only controlled by the analog circuitry (buffer and comparator) previously reported, but also by some digital gates, we need to take into consideration the variations (basically due to mismatch) in the quiescent point of the digital drivers when reporting the total offset of the structure (i.e. how do we detect the sign of $V_{ref}(t) - V_{ph}(t)$ and transmit it through our processing chain). Figure 4.9(a) shows the result of a 100 runs Monte Carlo simulation of the signal S as the output of the combination of the buffer, the comparator and the pixel digital circuitry for $V_{ref}=1.65V$ (DC Response). It must be noticed that these simulations does not include the effect of the reset operation, as they are an illustration of the effect of the comparison chain. Figure 4.9(b) shows the histogram of the measured offset. As expected, the standard deviation of the input-referred offset voltage has increased to 4.2mV and a mean value (basically due to slight differences in the quiescent point of the inverters and the comparator) of 8.22mV, producing offset voltages which vary (for 100 runs in the Monte Carlo simulation) between $[-20, 5]mV$. If we consider that the operation range will be $[0.7, 2.6]V=1.9V$ and we use a typical 3-sigma approach (only 1 out of every 370 cases is out of range), the maximum attainable resolution for the system would be $\log_2(\frac{1.9}{2 \times 3 \times 4.21682 \cdot 10^{-3}})=6.23$ bits.

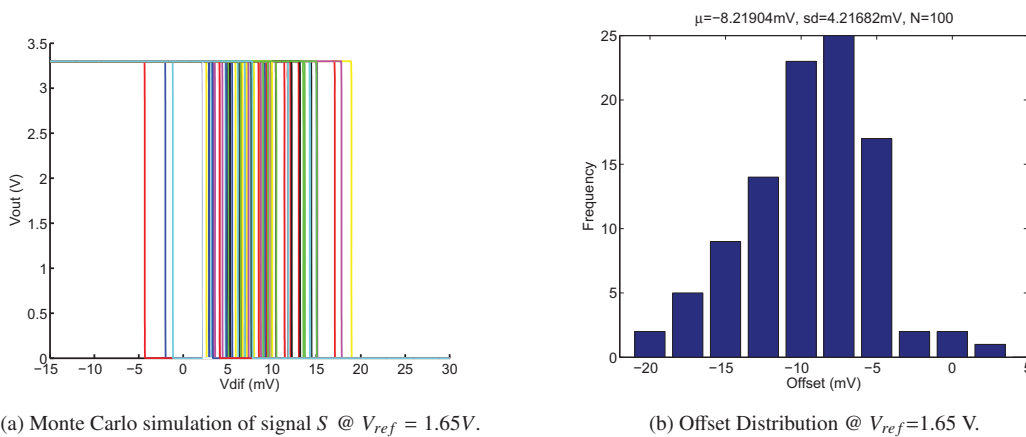


Figure 4.9: Offset of the combination of the buffer, comparator and digital circuitry.

4.4.2. Auto-Zeroing Reset Technique

In opposition to the usual method for photocurrent integration pixels, the operation of our imager is not linear. Therefore, common fixed pattern noise cancellation methods such as Correlated Double Sampling [47] or Double Delta Sampling [48] are not feasible. This is easily understood if we considered that, in our case, the subtraction of reference frames does not provide a correct output due to the implemented adaptive compression. Being aware of this, it becomes clear that all non-idealities cancellation must be performed before the signal is captured and therefore must be part of the capturing mechanism itself. Unfortunately, our area constraints do not allow large analog memories and restrain the options to use (and preferably reuse) simple circuitry.

Since, in our case (leaving aside the Photo Response Non Uniformity [PRNU]) the largest source of error is expected to be the input-referred offset voltage of the processing chain, we have introduced a variation of the typical reset method. The implemented method, which has some similarities to what is employed in Time-To-First-Spike pixels [127], consists on creating a negative feedback loop containing the buffer, the analog comparator and the feedback switch P1, and introduce the reset voltage V_{rst} through the comparator's positive input node. Thus, the photocurrent integration capacitance is initialized at a voltage, which already contains information about the offset of the buffer and the comparator (detailed operation is described below).

Considering that the amplifiers can be efficiently modeled to our purposes by their input-referred offset voltages and finite DC gains, it is possible to formulate the operation of the autozeroing technique in a simplified manner. First of all, we easily find the output voltages of comparator and buffer to be given by equations 4.4 and 4.5, where V_{ph} is the voltage (referenced to ground) at the photocurrent integration node and V_{ref} is the voltage at the comparator's positive input. Input-referred offset voltages are denoted as V_{oc} for the comparator and V_{ob} for the buffer, whereas A_c and A_b are comparator and buffer gains, respectively.

$$V_{buf} = \frac{A_b}{1 + A_b}(V_{ph} + V_{ob}) \quad (4.4)$$

$$V_{comp} = A_c(V_{ref} - V_{buf} + V_{oc}) \quad (4.5)$$

During the reset phase, V_{rst} is applied at the V_{ref} input, and since the feedback switch P_1 is ON we obtain that the final reset value can be expressed as:

$$V_{ph_{rst}} = \frac{A_c(V_{rst} - \frac{A_b}{1+A_b}V_{ob} + V_{oc})}{1 + \frac{A_c A_b}{1+A_b}} \quad (4.6)$$

Assuming that the gains of the amplifiers are both much higher than 1, we obtain that the value of $V_{ph_{rst}}$ is precisely approximated by:

$$V_{ph_{rst}} \approx V_{rst} - V_{ob} + V_{oc} \quad (4.7)$$

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

During the integration period (exposure), the feedback switch P_1 is OFF, and the photocurrent discharges the integration node from this previous reset voltage. If we consider that turning off the feedback switch introduces a feedthrough error ΔV_{ft} [128] in the integration node, we can express the temporal evolution of this node as:

$$V_{ph}(t) = V_{ph_{rst}} - \frac{I_{pix}}{C_{pix}} \Delta t + \Delta V_{ft} \approx V_{rst} - \frac{I_{pix}}{C_{pix}} \Delta t - V_{ob} + V_{oc} + \Delta V_{ft} \quad (4.8)$$

From equation 4.8, we simply obtain $V_{buf}(t)$ and $V_{comp}(t)$ by substitution (again assuming that the gain is much higher than 1):

$$V_{buf} \approx V_{ph} + V_{ob} \approx (V_{rst} - \frac{I_{pix}}{C_{pix}} \Delta t - \cancel{V_{ob}} + V_{oc} + \Delta V_{ft}) + \cancel{V_{ob}} = V_{rst} - \frac{I_{pix}}{C_{pix}} \Delta t + V_{oc} + \Delta V_{ft} \quad (4.9)$$

$$\begin{aligned} V_{comp} &= A_c(V_{ref} - V_{buf} + V_{oc}) = A_c \left[V_{ref} - (V_{rst} - \frac{I_{pix}}{C_{pix}} \Delta t + \cancel{V_{ob}} + \Delta V_{ft}) + \cancel{V_{ob}} \right] \\ &= A_c \left[V_{ref} - (V_{rst} - \frac{I_{pix}}{C_{pix}} \Delta t + \Delta V_{ft}) \right] = -A_c (V_{ph_{ideal}} - V_{ref} + \Delta V_{ft}) \end{aligned} \quad (4.10)$$

Which shows how, under these assumptions, the effective differential input applied to the comparator does not contain any reference to neither the comparator nor the buffer input-referred offset voltages. Needless to say, this ideal behavior will not occur in practice, where the output will obviously exhibit some dependency on the offset of these two amplifiers. Since our only assumptions were to model the amplifiers as having finite DC gain and input-referred offset voltages, and we neglected (in some cases) the $1/A$ factors in some additions in the previous calculations, we can evaluate the residual error by moving one step backwards and undoing these simplifications. In this case, the effective differential input voltage to the comparator is found to be given by:

$$\begin{aligned} V_{diff_{eff}} &= V_{ref} - (V_{rst} - \frac{I_{pix}}{C_{pix}} \Delta t) + \frac{A_b + 1}{A_c A_b + A_b + 1} V_{rst} - \frac{1}{A_b + 1} \frac{I_{pix}}{C_{pix}} \Delta t \\ &\quad - \frac{A_b}{A_c A_b + A_b + 1} V_{ob} + \frac{A_b + 1}{A_c A_b + A_b + 1} V_{oc} - \frac{A_b}{A_b + 1} \Delta V_{ft} \end{aligned} \quad (4.11)$$

Therefore, the residual¹ error of the autozeroing technique is:

$$\begin{aligned} V_{residue} &= \frac{A_b + 1}{A_c A_b + A_b + 1} V_{rst} - \frac{1}{A_b + 1} \frac{I_{pix}}{C_{pix}} \Delta t - \frac{A_b}{A_c A_b + A_b + 1} V_{ob} + \frac{A_b + 1}{A_c A_b + A_b + 1} V_{oc} - \frac{A_b}{A_b + 1} \Delta V_{ft} \\ &\approx \frac{1}{A_c} V_{rst} - \frac{1}{A_b} \frac{I_{pix}}{C_{pix}} \Delta t - \frac{1}{A_c} V_{ob} + \frac{1}{A_c} V_{oc} - \Delta V_{ft} \end{aligned} \quad (4.12)$$

¹Non-compesated residual error at the input of the comparator.

Clearly, the main contribution to the error is the feedthrough introduced by the reset switch that would also appear in conventional APS pixels, and which, as previously mentioned, is part of the error that is canceled through CDS. Since no cancellation mechanism is provided for this error, its contribution becomes dominant in equation 4.12. However, although it is somewhat an anticipation of a result not shown yet, simulations show that the feedthrough contribution of a minimal dimension transistor such as $P_1(\frac{0.4\mu m}{0.35\mu m})$ is low when compared to the aggregated offset of the processing chain (remember, in the order of $\pm 3 \times \sigma = \pm 3 \times 4.2mV = \pm 12.6mV$). Besides, observing equation 4.10, it is possible to consider the feedthrough as an addition to V_{ref} and it will be the mismatching in ΔV_{ft} (pixel to pixel variation in the introduced feedthrough), the responsible for lowering the maximum attainable accuracy. The electrical design of the pixel has been accomplished under the 3σ constraint for all aggregated errors after auto-zeroing. Therefore, since the operational signal range is $[0.7, 2.6]V$ and the targeted image resolution is 7-bits, the maximum allowable error results in:

$$2 \times 3\sigma \leq \frac{\text{Voltage Range}}{2^{Nbits}} = \frac{(2.6 - 0.7)}{2^7} = \frac{1.9}{128} \rightarrow \sigma \leq 2.5mV \quad (4.13)$$

In order to illustrate how the auto-zeroing technique improves the performance of our processing chain, and that specifications are met, we have executed two different Monte Carlo simulations (100 runs each) with and without auto-zeroing enabled. In both cases, we employ the switch P_1 to reset the photodiode (thus the expected feedthrough contribution should be the same or very similar at least) to the same voltage¹. In both cases, we simulate the photogenerated current by adding an independent current source $I_{ph} = 800fA$ in parallel with the photodiode. Then, we measure the value of $V_{ref}(t)$ (ramping up from 0.7V to 2.6V in 1.28ms), when SRAM control signal S disables SRAM write operation². The results of these Monte Carlo simulations are shown in figure 4.10(a) for the case without auto-zero, and figure 4.10(b) with autozero. As we can see, the average crossing point is 1.4482V in one case and 1.4485 in the other case, thus, the average detected point only differs in 300 μV . However, when we observe the statistics of the crossing point, one clearly sees the improvement provided by the auto-zeroing technique. Thus, the auto-zeroing disabled simulations show a standard deviation of the crossing point of 4.46548mV, which leads to an equivalent resolution³ of only $Nbits = \log_2(\frac{2.6-0.7}{6 \times 4.46548 \cdot 10^{-3}}) = 6.15$ -bits, whereas the auto-zeroing enabled simulations provide crossing points whose standard deviation is 737.272 μV , achieving an equivalent resolution of $\log_2(\frac{2.6-0.7}{6 \times 737.272 \cdot 10^{-6}}) = 8.74$ -bits.

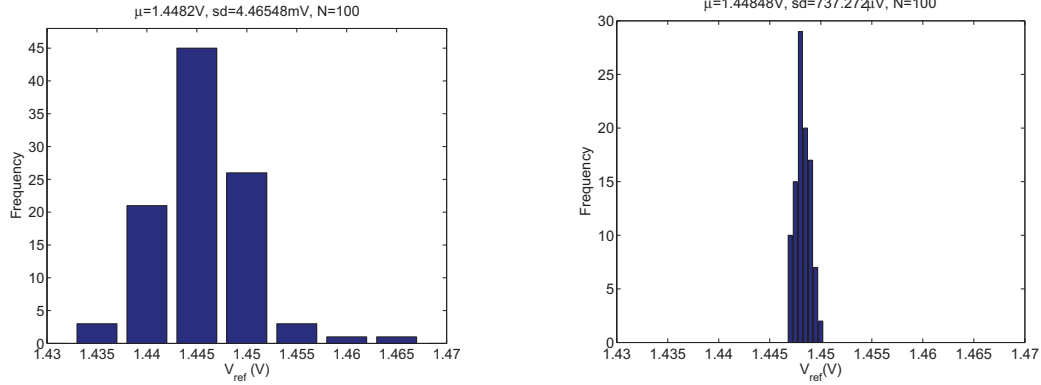
From these results one might conclude that the analog circuitry within the pixel is somewhat over-designed (providing more accuracy than the required and, then, wasting area and power) for our targeted 7-bit images. However, we must bear in mind that this 8.74-bits equivalent accuracy is obtained for the maximum signal swing (1.9V), and we will show in section 4.5.1 how it may be convenient to modify the levels of the A-to-D converter (reduce the voltage swing) in order to reduce the visual effects produced by the dark current in the darkest parts of the image. Thus, these 1.74 extra bits allow us to scale-down the voltage swing to just $6 \cdot 737.272\mu V \cdot 128 = 0.57V$ and still keep 7-bit accurate images.

¹This is an important aspect since there is a signal dependent component on the feedthrough contribution and also the photodiode's parasitic capacitance at the reset point, which both depends on the reset voltage.

²In order to decide the exact point, we actually measure the voltage at the crossing point of S and V_{ref} .

³ 3σ approach.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY



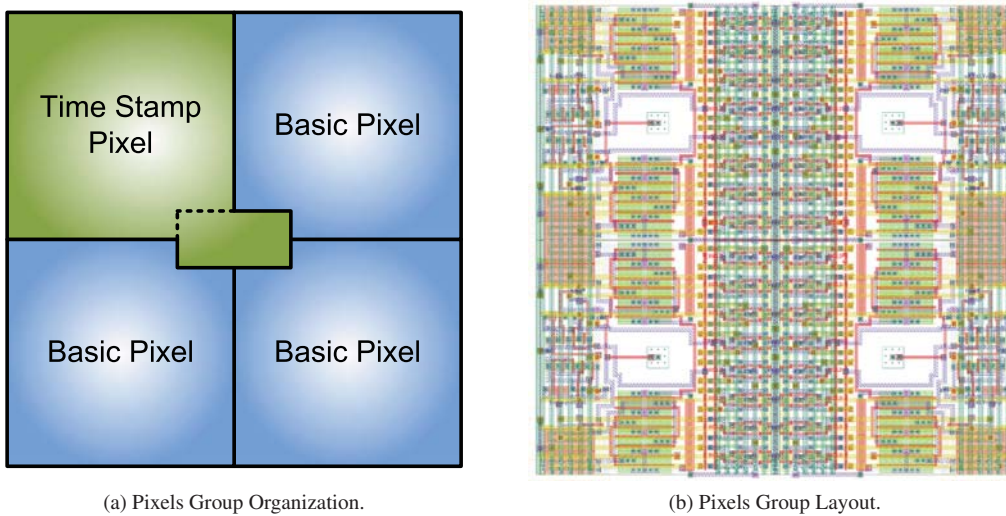
(a) Voltage at the crossing point: Auto-zeroing disabled.

(b) Voltage at the crossing point: Auto-zeroing enabled.

Figure 4.10: Simulation of auto-zero improvement.

4.4.3. Pixels Physical Characteristics

Pixels are arranged in groups of 2×2 elements containing 3 Basic Pixels and 1 Time Stamp Pixel. This allows us to implement the required $\frac{1}{4}$ subsampling for TSI generation in a very simple manner. Figure 4.11(a) shows the physical distribution of Basic and Time Stamp Pixels in the 2×2 arrangement, which occupies an area of $66 \times 66 \mu m^2$ ($33 \times 33 \mu m^2$ per pixel). Notice that, conceptually, each TS Pixel takes some area from its 3 Basic Pixel neighbors in order to allocate its extra SRAM resources. However, in practice, what happens is that both BP and TSP contain 8-SRAM (7TMC+1TSC) modules with the TS related registers being only controlled by the logic in the TS pixel.



(a) Pixels Group Organization.

(b) Pixels Group Layout.

Figure 4.11: Pixels Group.

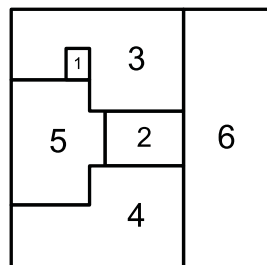
The layout of the group is displayed in figure 4.11(b). Observe that the SRAM modules have been grouped in the central vertical region, sharing digital power and ground lines. In the pixel array, other groups will be placed in every side (left, right, top and bottom). Therefore, the PMOS analog transistors and the digital control circuitry are also grouped side-by-side between pixels, equally sharing their power lines. This allows to vertically route the I/O buses compactly on top of center and left-right sides (metal 2 and metal 4 on top for the data bus and its negated version), which improves the attainable pitch. Figure 4.12 shows the position and relative areas occupied by the different functional blocks on the pixel (left top pixel of the group).

Due to the layout rules, a large separation must be provided between any floating (“hot”) Nwell, such as the photodiode, and the rest of circuitry. In order to take some advantage of this a priori wasted space, the metal aperture over this area has been extended as much as possible. Carriers created within this area can contribute to the pixel photocurrent by reaching the photodiode collection volume through diffusion. The metal opening in the layout is $9.75 \times 7.3 \mu\text{m}^2$ whereas the photodiode is $3 \times 3 \mu\text{m}^2$.

4.5. Digital to Analog Converter

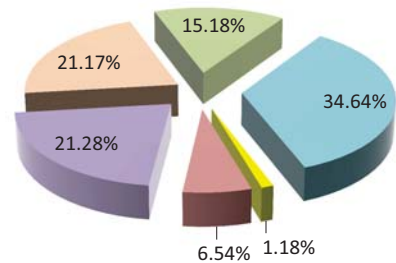
A digital to analog converter (DAC, or D-to-A) generates the analog reference V_{ref} . The temporal evolution of V_{ref} is sketched in figure 4.13. First, it takes the value of the reset voltage V_{rst} during photodiodes initialization, then it goes down to V_{bot} during the first 15 temporal windows (usually most of the exposition time), and finally it ramps up from V_{bot} to V_{top} at the end of the exposition time in order to provide the A-to-D conversion of the low illuminated pixels.

The large initial voltage drop, from V_{rst} to V_{bot} , must be as fast as possible in order to expand the illumination capturing range of highly illuminated pixels. Indeed, if a pixel is so bright that it discharges the integration node before V_{ref} has reached V_{bot} , the SRAM on this pixel will keep its reset value (127). This fact is indeed the limiting factor for capturing high illuminations. There are also some limitations regarding the rate of change of V_{ref} during the A-to-D conversion in the last temporal window. On the one hand, the time between steps (A-to-D clocking) must be long enough to allow proper settling of the V_{ref} signal through the long (almost 6mm) horizontal wires carrying this signal to the pixels. On



(a) Functional Blocks Positions.

- 1-Reset Switch
- 2-Photodiode + Metal Opening
- 3-Comparator
- 4-Buffer
- 5-Digital Control
- 6-Memories



(b) Area percentage.

Figure 4.12: Pixel Area Organization.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

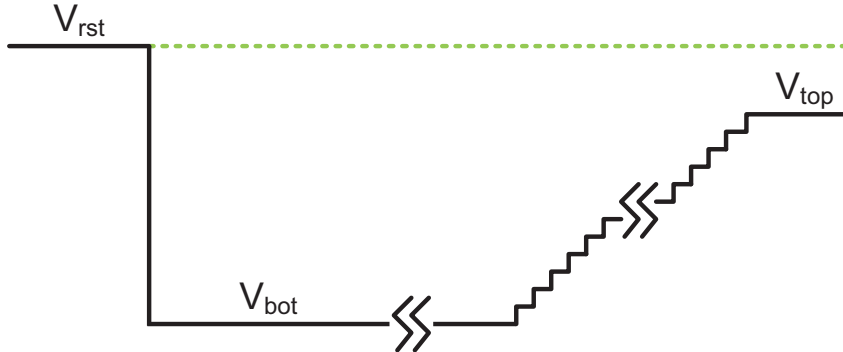


Figure 4.13: Voltages generated by DAC block.

the other hand, since the both V_{top} and V_{bot} are programmable (in order to adapt to the real linear range of the in-pixel amplifiers, and to diminish the effects of dark current on dark pixels), the accuracy of the settling process must be adapted accordingly in order to keep the equivalent number of bits above the 7-bit limit. In other words, if the ADC operating voltage swing is reduced, then the allowable settling error must be lowered, and so the frequency of the A-to-D clock. The rate of change of V_{ref} is controlled by external signals. In order to save IO pins, the ADC uses a counter (with signals DAC_CLK and DAC_CLR as controls) instead of having a 7-bit input to program the ADC output voltage. Besides, a 2-bit control bus DAC_MODE selects whether the DAC has to provide V_{rst} , V_{bot} , or the voltage ramp:

- $DAC_MODE = 00 \Rightarrow V_{rst}$
- $DAC_MODE = 10 \Rightarrow V_{bot}$
- $DAC_MODE = 01 \Rightarrow$ Counting Mode (output voltage ramps up)

The DAC is built using a classical resistor ladder architecture [129] whose basic element contains a $260.87\Omega^1$ poly2 resistor and a complementary CMOS switch. Since this is quite a conventional design, we will not add any further details here. Just for verification purposes, figure 4.14 shows the simulation results for Differential and Integral Nonlinearity Errors [130].

4.5.1. Dark Signal Contribution Attenuation

We have already mentioned how we can attenuate most of the offset contributions from the analog circuitry in the pixel by using an auto-zeroing technique to establish the reset voltage. The ultimate purpose of this calibration process is to improve the uniformity of the captured image. In this section, we will present another possibility provided by the chip to enhance the uniformity of the final images, specially in its darker zones, where noise is usually more noticeable. In this case, the main goal is to reduce the visual effects of the dark current on the darker pixels. To that purpose, we will employ a programmable definition of the voltage levels for A-to-D conversion.

¹Result provided by the extraction tool in Cadence.

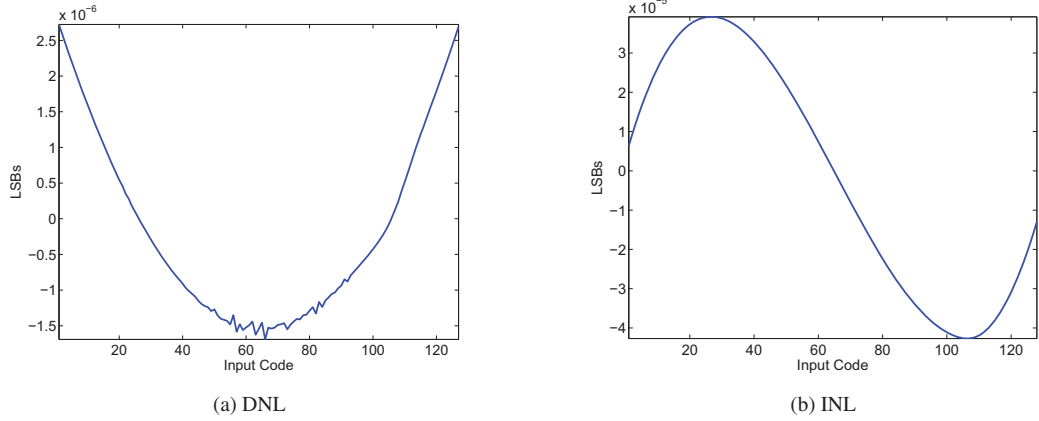


Figure 4.14: DAC Characteristics.

It is easy to understand that the effects of the dark current will be more noticeable in those pixels having a photogenerated current that is similar to the level of the dark current. In such cases, these pixels may look very noisy and provide little extra details to the image (except in some remarkable situations, like automotive applications), while they do worsen the overall quality of the image, which is appreciated by an independent observer. When correctly operated in the HDR mode, our chip will codify these dark pixels within the last bin, and will digitize them using a single-slope A-to-D conversion.

The current I_{pix} , which discharges the integration node V_{ph} in the pixels, can be basically considered as created by the photogenerated current I_{ph} plus the dark current I_{dark} . Figure 4.15 illustrates how we can attenuate the visual effects of dark current over the darker zones of the image. Assume that we can measure (as we will explain in chapter 5) the dark current contribution (basically its average value $\overline{I_{dark}}$ and standard deviation $\sigma[I_{dark}]$) and that we can also measure, at the same time, the temperature of the chip¹, so that we can create a table containing temperature, $\overline{I_{dark}}$, $\sigma(I_{dark})$, and exposition time. Now, if we lower the top voltage of the A-to-D converter such that a pixel whose total discharge current ($I_{ph} + I_{dark}$) is $\overline{I_{dark}} + 3\sigma(I_{dark})$ ² produces a voltage drop below $V_{rst} - V_{top_{new}}$, then these pixels will be in the saturation zone of the A-to-D conversion curve, and consequently will be assigned to the value 0. On the one hand, by doing this, we are clearly neglecting all those pixels whose pixel current is below this limit, and on the second hand, we are reducing the full signal scale of the A-to-D converter as well, or in other words, making its LSB smaller. Thus, we need to guarantee that, for this new voltage swing:

- The ADC keeps its 7-bit accuracy (including cumulative settling errors), see next section.
- The analog circuitry in-pixel keeps its 7-bit accuracy (remember that, accordingly to simulations we can vary the operating voltage range from a maximum of 1.9V to just 0.57V).

¹Remember that we have the V_{bias} output which comes from a PTAP.

²For the current temperature and exposure configuration.

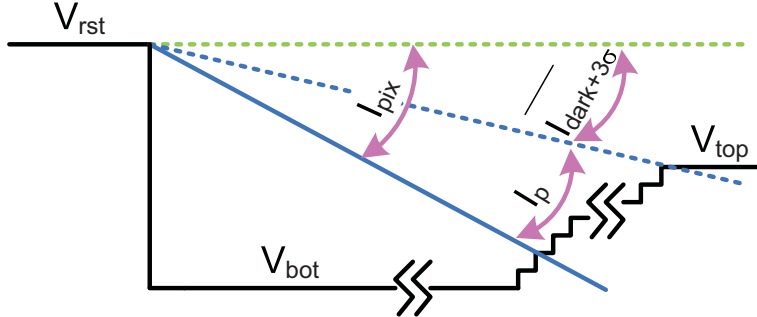


Figure 4.15: Voltage range for Dark Signal elimination.

4.6. V_{ref} Distribution Scheme

Leaving aside the design of the HDR algorithm and the implementation of the in-pixel circuitry, the scheme to precisely distribute the $V_{ref}(t)$ signal to a set of 148×180 (26640) nodes, in an array which is nearly $5mm \times 6mm$, has been by far the hardest design problem in our work. Needless to mention, the accurate distribution of $V_{ref}(t)$ is cornerstone for the correct operation of the whole chip. Indeed, differences in $V_{ref}(t)$ across the pixels (due to either offset or gain errors from the buffers driving this signal to different rows, settling differences due to the different distances, as this signal has to travel to reach every pixel, etc.) directly add to the residual offset of the in-pixel circuitry. Therefore, it can, eventually, ruin the operation of the imager, if they are significant enough.

If one seeks for a precise mismatch-free distribution of the V_{ref} signal, the most straightforward (and logic) approach is to employ a single DAC followed by a powerful output buffer to drive the array. This solution has some important limitations though. First, driving the huge array of $V_{ref_{ij}}$ nodes from a single point implies that one cannot operate faster than the intrinsic time constant of the whole array. Also, since there is only one buffer providing the power, there is a large concentration of power consumption in a very reduced area. This unavoidably leads to the creation of a hot spot around this place, which affects the performance of the pixels nearby (due to the increased temperature in the zone¹).

Let us talk about temporal specifications at this moment. Obviously, we cannot talk properly about temporal specifications without knowing the type of load we are driving. Since the array of pixels was already designed when we started the design of the V_{ref} distribution scheme, we were able to extract its parasitic capacitances and resistances, and compound a parasitic simulation model. Then, this complex model (it contains more than 26000 nodes) was simplified as follows in order to obtain useful design equations and meaningful constraints:

- We extract a row of pixels and excite the V_{ref} node with a constant 1A input current (this is a convenient value which simplifies calculations, and since we are considering linear parasitics it does not create any problem). In this simulation, we measure the slope of the V_{ref} signal in all nodes in the row (to check for discrepancies due to the internal time constant of the pixels) and

¹Observe that dark current has an exponential dependence with temperature [20].

extract the associated capacitance of the row simply as 1/slope. Simulation results show that $C_{row} = 10.82pF$.

- Now, we excite with a voltage pulse and measure how it propagates through the 180 pixels in this row. We measure the time constant at the farthest point, and define it as the row time constant. Simulation results show that $\tau_{row} = 4.54ns$.
- We adjust the R_{row} parameter such that a first order RC system with $C_{row} = 10.82pF$ has the same time constant. The result is that R_{row} needs to be configured to be $R_{row} = 419.59\Omega$.
- To check the correct operation of the 1st order model, we compare the dynamic evolution of the simplified model with that at the farthest point in the array. The only noticeable differences happen at the very first 200ps, however, the dynamic evolution is identical (to the 4th decimal number) from this time on, and therefore, we consider this model valid.
- The same procedure is applied for the whole array (which in addition to the 148 rows of 180 pixels also contains the parasitics associated to the vertical line that interconnects all V_{ref} rows). In this case, we decide to use as input port the point in the middle of the array (that is the optimum if we are distributing the signal from a single point). From simulations, we obtain $\tau_{array} = 787ns$, with $R_{array} = 483.11\Omega$, $C_{array} = 1.629nF$.

Having the load modeled, we need to talk about settling error specifications. Clearly, during the last temporal window, the 7-bit single slope A-to-D conversion establishes a maximum accumulated error that must remain below $\frac{1}{2}$ LSB. In our case, the accumulated error comes not only from the settling error in the DAC, but also from the noise, DAC's non-linearities, offsets (will see later), and the already mentioned residual error in the operation of the auto-zeroing employed to establish the reset. Considering that for the maximum allowable voltage swing (2.6V-0.7V) the 7-bit accuracy LSB corresponds to some 14.84mV, and that the accumulated errors are estimated (from simulations) to be around $3 \cdot 737.272\mu s^1 = 2.21mV$ plus other effects, so we overestimate the total in 3mV, the maximum settling error (E_{max}) must not be larger than some $\frac{14.84mV}{2} - 3mV = 4.42mV$, which giving an additional margin we choose about 4mV. These results in two different situations depending on whether we are within the last temporal window (doing the A-to-D conversion) or moving from V_{rst} to V_{bot} just after releasing the reset signal.

When executing the A-to-D conversion in the last step, the V_{ref} signal to be transmitted to the array moves in steps of 1LSB from V_{bot} to $V_{bot} + 2^7$ LSB. Assuming that each step has a duration T_{step} , the

¹ σ in auto-zero enabled simulations.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

settling error E_k at the k^{th} step can be evaluated using¹:

$$\begin{aligned}
 V(k \cdot T_{step}) &= V_{bot} + k \cdot LSB - E_k \\
 E_k &= LSB \sum_{j=1}^{j=k} e^{-j\alpha} = LSB \left[\frac{1 - e^{-(k+1)\alpha}}{1 - e^{-\alpha}} - 1 \right] \\
 \alpha &= \frac{T_{step}}{\tau}
 \end{aligned} \tag{4.14}$$

Since the settling error accumulates step by step, the worst case (obviously) is committed at the last conversion step, $k = 2^7$. Hence, our constraint becomes:

$$E_{max} \geq LSB \left[\frac{1 - e^{-(k+1)\alpha}}{1 - e^{-\alpha}} - 1 \right] \Rightarrow \frac{1 - e^{-(k+1)\alpha}}{1 - e^{-\alpha}} \leq 1 + \frac{E_{max}}{LSB} \tag{4.15}$$

For $(k+1)\alpha$ values larger than 4, or, in other words, we apply T_{STEP} times that are not shorter than $\tau/32$, this equation converges to:

$$T_{step} \geq \tau \cdot \ln \left[1 + \frac{LSB}{E_{max}} \right] \tag{4.16}$$

which results in $T_{step}(\tau_{array}) \geq 1.22\mu s$, and $T_{step}(\tau_{row}) \geq 7ns$. Therefore, it does not matter how powerful the output buffer of the DAC is, we simply cannot transmit V_{ref} to the array with the necessary accuracy in less than $1.22\mu s$, if we want to keep our specifications within the last step of the analog staircase signal. This also has an important implication regarding the duration of the last temporal window, since it sets a minimum of $T_{ramp}(\tau_{array}) \geq 156\mu s$ or $T_{ramp}(\tau_{row}) \geq 0.9\mu s$.

A second constraint is derived from the first transition of the V_{ref} signal. Just after reset, this signal is changed from V_{rst} to V_{bot} . The time required to establish this change with the required accuracy (E_{max}) defines the maximum photogenerated current that can be assigned a value different from 0 (that is the maximum detectable photocurrent). Thus, if we need to produce a change from V_{rst} to V_{bot} , the minimum time for producing this change with the required accuracy is:

$$T_{min} \geq \tau \cdot \ln \left[\frac{V_{rst} - V_{bot}}{E_{max}} \right] \tag{4.17}$$

A priori, we can assume that $V_{rst} \cong V_{top}$, hence, this limit becomes:

$$T_{min} \geq \tau \cdot \ln \left[\frac{2^N \cdot LSB}{E_{max}} \right] \tag{4.18}$$

which evaluates to $T_{min}(\tau_{array}) \geq 4.85\mu s$, and $T_{min}(\tau_{row}) \geq 27.98ns$. Thus, we can say that the maximum theoretical ratio between largest detectable photocurrents, if we drive V_{ref} to the array from a single point, and if we use one buffer per row is:

¹Here, we use that the Sum of Geometric Progression is $\sum_{j=0}^{j=k-1} r^j = \frac{1-r^k}{1-r}$ [131].

$$R = \frac{\tau_{array}}{\tau_{row}} = 168.99 \quad (4.19)$$

Assuming that the minimum detectable photocurrent is limited by noise factors, driving V_{ref} to the array row by row results in a maximum Dynamic Range improvement of nearly $20 \cdot \log_{10}(168.99) = 44.56dB$ (simply because most of our algorithm relies on making temporal measurements), when compared to the alternative solution of using a single buffer for the whole array. Needless to say, this is a theoretical limit, which assumes that we can incorporate a buffer per row whose response time is significantly faster than τ_{row} (4.54ns), something which is not trivial at all in a 0.35 μ m technology. Clearly, if the time to load a row is limited by the internal dynamics of the buffer (τ_{buf}) rather than for the time constant of the row, the ratio in equation 4.19 becomes:

$$R = \frac{\tau_{array}}{\tau_{buf}} \quad (4.20)$$

hence, as long as we can design a buffer having the pitch of a row¹, which can drive a row much faster than τ_{array} , we would always have a Dynamic Range improvement when using a distributed topology to transmit V_{ref} to the pixels as compared to the single buffer option.

However, distributing V_{ref} using one buffer per row has an important drawback, the offset voltages of the buffers add to the signal being transmitted. Therefore, the specifications may become unreachable not because of the speed of the system, but because the final point is simply incorrect. Consider that, the error must remain below $E_{max} = 4mV$ for the typical operating range [0.7, 2.6] (and it may need to be lowered when using a lower range). If we maintain our 3σ approach, it means that the standard variation of the offset voltage of the buffer should not exceed some $\frac{4mV}{3} = 1.3mV$ in the typical case.

As a first solution, one may consider using any of the typical analog techniques for offset correction in amplifiers (including using very large transistors in the differential pair, common centroid configurations, auto-zeroing, etc.). Our system has a particularity though; the buffer must have the same pitch as the pixel. This, together with the need for a long lasting calibration of the offset voltages (because of the long exposures employed in low-illuminated scenes), made us finally opt for a hybrid solution, which combines, sequentially, a distributed buffer topology, and a single point driving for a precise establishment of the final voltage.

Our solution combines the use of one amplifier per row during the first moments of V_{ref} changes, and driving the array from a single point (and rely on the charge redistribution principle) at the final part of the transient to guarantee that all $V_{ref_{ij}}$ nodes reach the same final voltage, or more precisely, that all V_{ref} nodes are established to the same voltage with the desired accuracy level. Therefore, we will not meet our maximum attainable speed (limited by τ_{row}) but neither we shall be limited² by the time constant of the whole array (τ_{array}). The solution, which is given the name of Charge Injection Amplifiers Block, is inserted between the output of the DAC and the horizontal metal lines that distribute V_{ref} to the pixels row by row. Figure 4.16 shows the schematic of the cell, which is included in every row. Here, V_{dac}

¹And sufficiently small in the horizontal dimension to not affect significantly the final size of the chip.

²Only partially limited.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

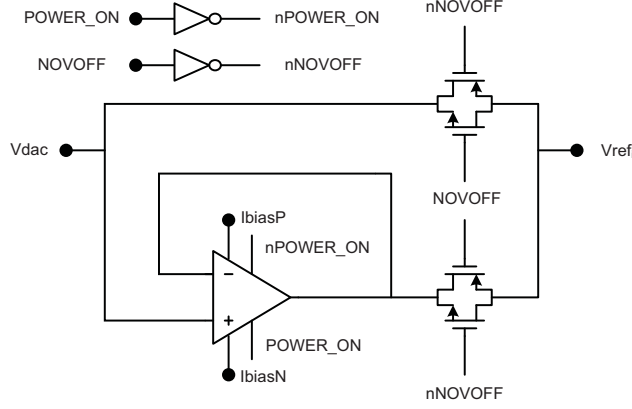


Figure 4.16: Charge Injection Amplifier Cell Schematic.

is the output of the DAC block and V_{ref_i} is the V_{ref} metal line of the i^{th} row. Clearly, two signal paths can be enabled, depending on the state of two complementary transfer gates. One path allows a direct connection between V_{dac} and V_{ref_i} ($NOVOFF=1$), while the alternative path inserts a buffer between these two nodes ($NOVOFF=0$).

The operation of the charge injection amplifier block is illustrated in figure 4.17. At the very end of the reset phase, when the V_{ref} signal must drop very fast from V_{rst} to V_{bot} , the amplifiers are connected allowing this change to happen faster than if we were using a single point distribution scheme. However, the steady state voltages are incorrect due to the introduced offsets. Once V_{ref_i} nodes have stabilized, the amplifiers are disconnected, and the direct path is engaged short-circuiting all V_{ref_i} row lines to the output of V_{dac} . Now, charge redistribution occurs and all nodes will reach the same final voltage once the transient settles down. The key idea here is that we only need to evolve following the time constant of the array the time required to lower the average settling error (because we have many amplifiers) of the first phase (evolving with the time constant of the buffer) below our accuracy limit. Let us illustrate how are we gaining operation speed thanks to this configuration.

Assume that all V_{ref_i} lines reach their steady state value $V_{ref_i}(T_{rst}) = V_{rst} + V_{off_i}$ (where V_{off_i} is the offset voltage of the buffer driving the i^{th} row), and that now we switch from V_{rst} to V_{bot} keeping the buffers connected during a time T_1 . In this case, every V_{ref_i} line reaches a voltage value:

$$V_{ref_i}(T_1) = (V_{bot} + V_{off_i}) + (V_{rst} - V_{bot}) e^{-T_1/\tau_{buf_i}} \quad (4.21)$$

Now, we disconnect the buffers and engage the direct path for a time T_2 . In this case, the evolution is dictated by τ_{array} , such that the final voltage at the farthest point in the array reaches a value¹:

¹Here we assume (validated through Monte Carlo simulations) that:

- (1) $\sum_{i=1}^{i=N_{rows}} \frac{V_{off_i}}{N_{rows}} \approx 0$ due to the large number of rows.
- (2) Redistribution of charge at the beginning of the second phase of the transient is only significant between the very close neighbors of a row and, hence, it is not limited by whole array dynamics. Or, in other words, if we select a row, the charge stored in a very reduced numbers of rows (up and down) around this row is very close to that of the offset free case.

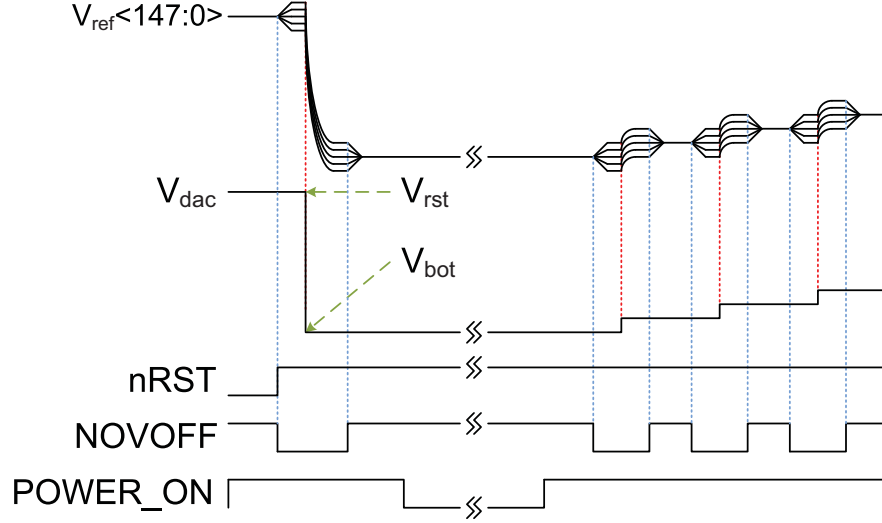


Figure 4.17: Charge Injection Amplifier Operation.

$$V_{ref}(T_2) = V_{bot} + \left[\cancel{V_{bot}} + (V_{rst} - V_{bot})e^{-T_1/\tau_{buf}} - \cancel{V_{bot}} \right] e^{-T_2/\tau_{array}} \quad (4.22)$$

therefore, the settling error at $t = T_1 + T_2$ is given by:

$$E_{set} = (V_{rst} - V_{bot})e^{-T_1/\tau_{buf} - T_2/\tau_{array}} \quad (4.23)$$

whereas if we were driving the array from a single point for the same time, we would have:

$$E_{set}^* = (V_{rst} - V_{bot})e^{-(T_1+T_2)/\tau_{array}} \quad (4.24)$$

and the question is, under what conditions ($E_{set} < E_{set}^*$)?, which results in the following inequality:

$$e^{-T_1/\tau_{buf} - T_2/\tau_{array}} < e^{-(T_1+T_2)/\tau_{array}} \quad (4.25)$$

which can be rewritten in the form:

$$-\frac{T_1}{\tau_{buf}} - \frac{T_2}{\cancel{\tau_{array}}} < -\frac{T_1}{\tau_{array}} - \frac{T_2}{\cancel{\tau_{array}}} \quad (4.26)$$

The solution to this type of inequality is:

$$\tau_{array} > \tau_{buf} \quad (4.27)$$

The inequality results in the logic condition that the buffers response time must be faster than that of the array.

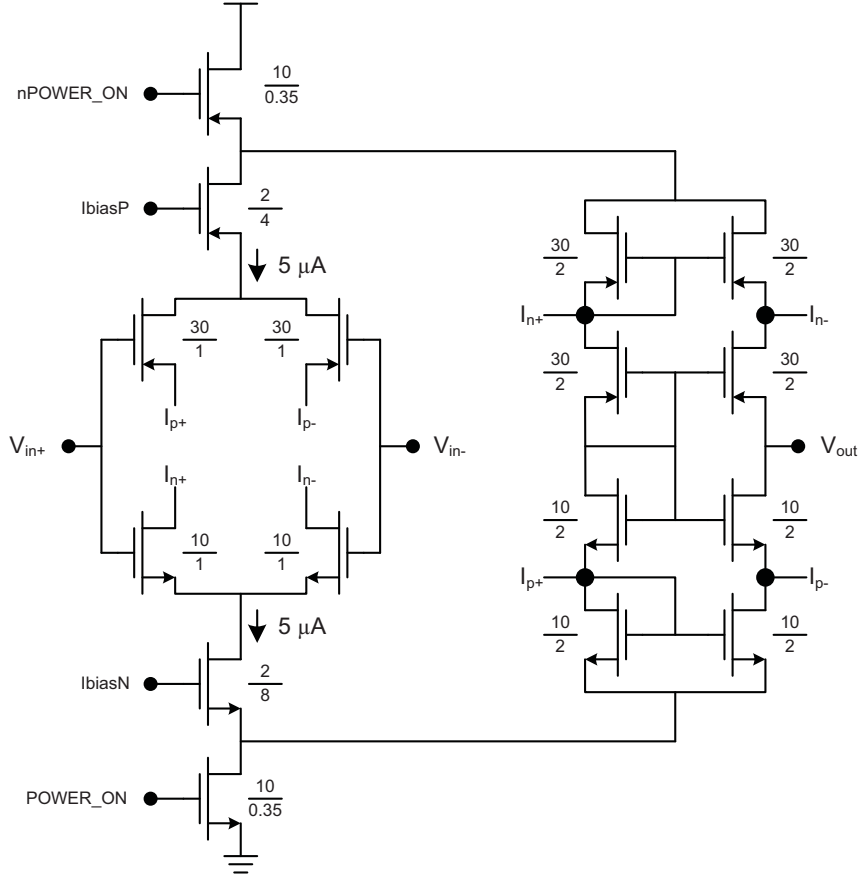


Figure 4.18: Schematic of the Amplifier in the row-wise V_{ref} distribution scheme

The design of the buffer (see figure 4.18) is based on a complementary input differential pair which drives a self-biased complementary folded cascode structure[132][133]. This single stage amplifier has been chosen in order to guarantee feedback stability and relatively high gain. The buffer incorporates a *POWER_ON* signal, which avoids wasting any power in this block when it is not in use. Noticeably, this happens for most of the exposition time, since in common temporal window configurations, the operation of the buffer is only required when V_{ref} experiences fast changes (during the last temporal window and just after reset) but not when $V_{ref} = V_{bot}$, as it happens for the first 15 temporal windows.

Figure 4.19 shows the DC gain and AC response of the buffer (including one row parasitic load model), whereas figure 4.20(a) shows the characteristic of the error at the output, where a red line indicates the maximum error and a blue dotted line indicates the maximum allowed error, and figure 4.20(b) is the input-referred offset voltage histogram at input 1.65V obtained from 100 runs Monte Carlo simulations.

Finally, in order to validate the global operation of the V_{ref} distribution scheme, we have performed different Monte Carlo simulations that include the DAC, the distributed buffers, and the complete par-

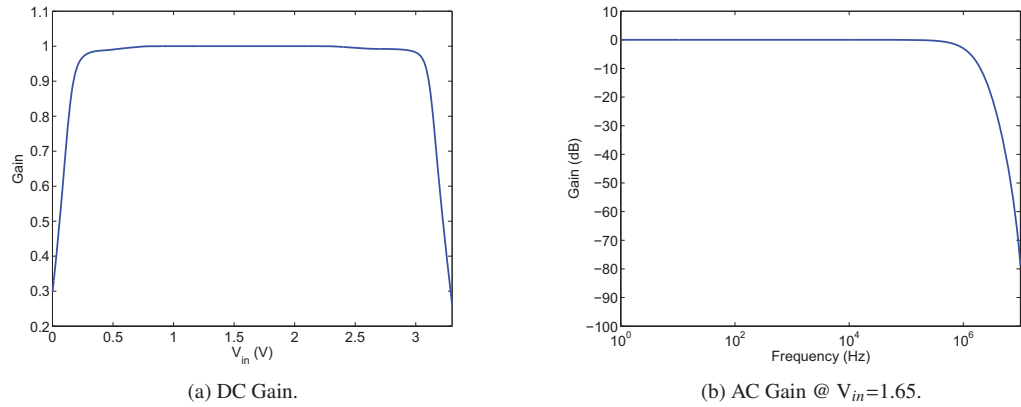


Figure 4.19: Buffer Gain Plots.

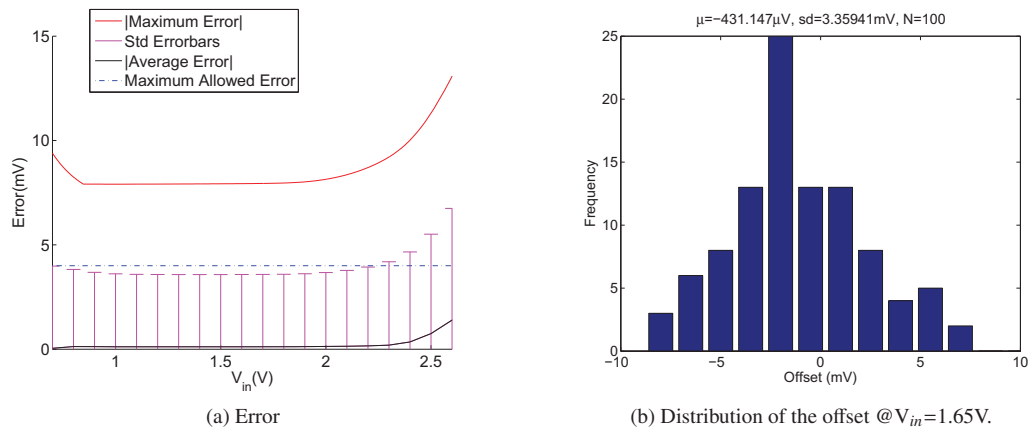


Figure 4.20: Folded Cascode Buffer Errors.

asitic load of the whole array (including the extracted view of each pixel). Figure 4.21 shows the V_{ref} voltages at the end of each row in the latest codes of the ADC operation (near V_{top}), with the theoretical output from the DAC in green and the dotted lines corresponding to the maximum allowable error. Clearly, the amplifiers are able to transmit the signal to the array using a $1.2\mu\text{s}$ per step timing. However, the offset voltages make the signals to spread out of the maximum error limits.

Figure 4.22 shows an example of the operation of the dynamic V_{ref} distribution scheme for a $1.2\mu\text{s}$ per step timing. The amplifiers are enabled for 250ns before the change and 500ns after the change (750ns in total) whereas the lines are connected together to the output of the DAC for 450ns . Observe that charge redistribution happens very fast, and that the transient evolution towards the steady state is interrupted after settling error constraints are met.

Figure 4.23 shows the settling error (maximum as a dash-dotted line and standard deviation bars) for 128 steps. The error is measured 925ns after the change in V_{ref} (just before enabling the amplifiers

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

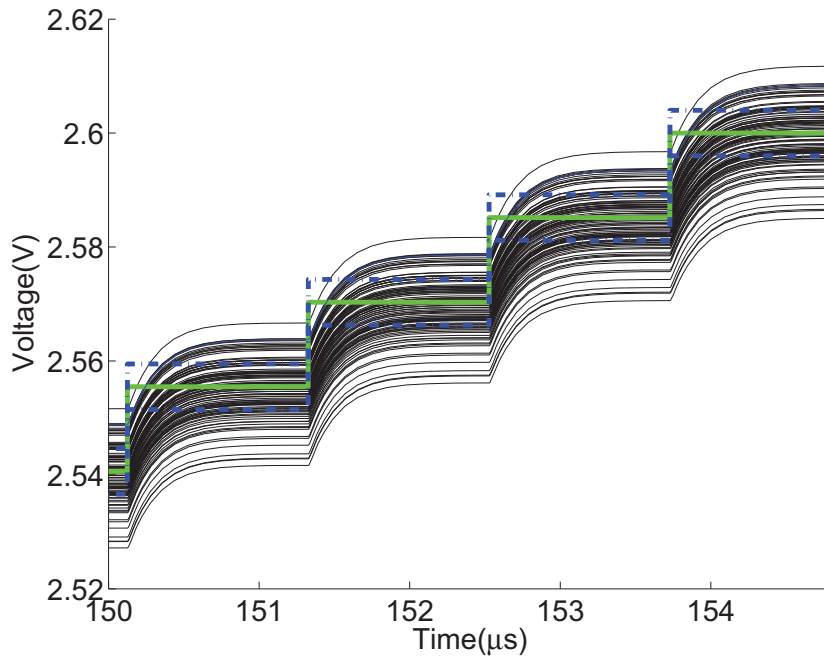


Figure 4.21: Effect of buffers offset in the distributed V_{ref} signal for $T_{step} = 1.2\mu s$.

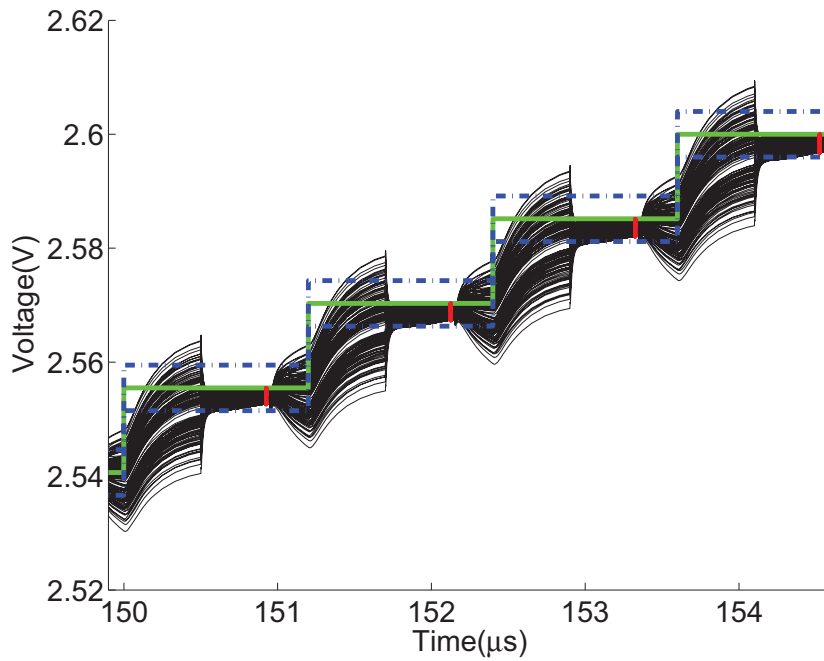


Figure 4.22: Operation of the V_{ref} distribution scheme for $T_{step} = 1.2\mu s$.

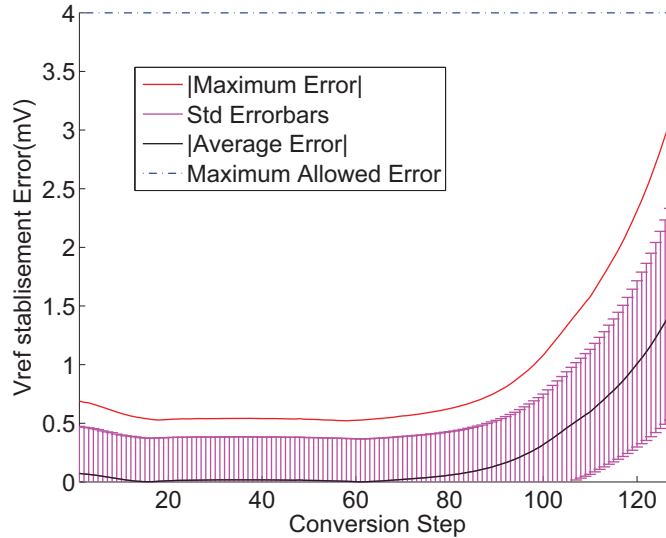


Figure 4.23: Settling Error in the V_{ref} distribution scheme for $T_{step} = 1.2\mu s$.

again). Internally, this is achieved by clocking the operation of the comparator with the in-pixel nEVAL signal that is synchronized with the DAC's clock. As it can be seen, the operation remains within the maximum allowable error limits for all the 128 input codes. The increase in the error towards the higher input codes is due to the fact that both the gain and the AC response of the amplifier worsen for increasing input voltages.

4.7. Code Generator

The sequences appearing at $TMC\langle 6:0 \rangle$ and $TSC\langle 3:0 \rangle$ are internally generated by a specific block located at the periphery of the array. Roughly speaking, this block is just a couple of non-cyclic down-to-zero counters plus a binary-to-gray decoder. The state of the counters (that is generated in the default binary format for convenience) is transformed to gray format by purely combinational logic. These gray codes are then transmitted to the sense amplifier block, which writes them to every column bus, and consequently to the in-pixel SRAMs. Codes are generated in gray format in order to reduce the maximum power consumption peaks, since thanks to using this coding, only one SRAM module flips at every time instead of a maximum of 7 (when changing from 63 (0111111) to 64 (1000000)) in the Basic Pixels, and only 2 instead of 11 (7+4) in the Time Stamp Pixels.

4.8. Pixels Control and I/O Interface

The chip includes several blocks at the periphery of the array to control and manage digital information. These blocks are basically Digital Buffers, which distribute controls to the array under strict

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

timing restrictions, Sense Amplifiers, which write/retrieve information from the in-pixel SRAMs, and a one row Read Buffer, which accelerates image downloading. The reason for needing sense amplifiers to download the contents of the in-pixel SRAM is found in the fact that the metal lines, which connect the SRAM to the external circuitry, run vertically crossing the whole array (nearly 5mm), and consequently the associated parasitic capacitance is huge. Therefore, if we would just simply connect any of the data nodes of the SRAM to these lines, a previous data (opposite) might flip the latch and destroy the stored information. In order to avoid this catastrophic behavior, the metal lines *DATA* and *nDATA* (figure 4.7) are precharged to the same voltage (indeed a high voltage) before connecting the SRAM to them. When the connection is done, one of the terminals of the SRAM does not suffer redistribution effects (since, for sure, one of these two terminals will have a high voltage), whereas the other terminal will suffer a voltage drop due to charge redistribution. However, since the other terminal has been kept, the amplifier in the sense amplifier (that is connected just few nanoseconds after the SRAM is disconnected from the data lines) will easily detect the sign of the difference and will regenerate the levels present in the SRAM.

The Sense Amplifiers [134] block contains $180 \times 7 + 90 \times 4 = 1620$ voltage-mode sense amplifiers. Figure 4.24 shows the schematic of 1 bit cell. Signal *PRECH* initializes the data lines *DATA* and \overline{DATA} to $V_{dd} - V_{th_n}$ to avoid SRAM flip. The control of the sequential precharge of column lines are generated from a single clock *PCH_CLK* and reset input *PCH_START*. Later, the signal *READ* is enabled and the voltage in the (*n*)*DATA* lines will be regenerated. A bus keeper has been added to the terminal *DATA_OUT* in order to allow a simpler synchronized readout by the Read Buffer. Signal *WRITE* is employed to transmit the information produced by the Code Generator to the in-pixel SRAM cells during normal HDR operation.

Since every group of 2×2 pixels uses a vertical 18-bit bus ($TMC_j \langle 6:0 \rangle + TSC_j \langle 3:0 \rangle + TMC_{j+1} \langle 6:0 \rangle$, $7+4+7 = 18$), the sense amplifiers are organized in sets of 9×2 -bit blocks (to match the pitch of the 2×2 arrangement of pixels). The signal *PCH_START* initializes the state machine that controls the precharging sequence, which for power-saving issues is executed in 9 cycles¹ (sequenced by *PCH_CLK*). Then, the SRAM modules in the selected row are connected to their corresponding (*n*)*DATA* lines. Right afterwards, the signal *READ_C* is enabled to read *TMC* data whereas signal *READ_H* is employed (only for odd rows) to read *TSC*.

The Digital Buffers block generates the digital control signals for the array of pixels. These signals, namely *nRST*, *nREAD* and *nEVAL*, are applied to the 148 rows in a row-wise way. Their distribution is carried out through a buffer tree, which has been generated with an automatic layout tool (CADENCE® Encounter), with dimensions of the buffers in the tree and the number of branches such that skew and delay specifications are optimally met. The same distributed scheme is used to create the row selection signal² *nROW*, which activates the rows sequentially.

Finally, the Read Buffer stores a complete row (the one previously read by the Sense Amplifiers). It allows acquiring the next row while the previous one is being downloaded. Since the output of the chip is organized in a 36-bit bus (4 TMI pixels and 2 TSI pixels are provided in every readout cycle), the Read

¹Another 10th cycle is required to end the process.

²It is active low.

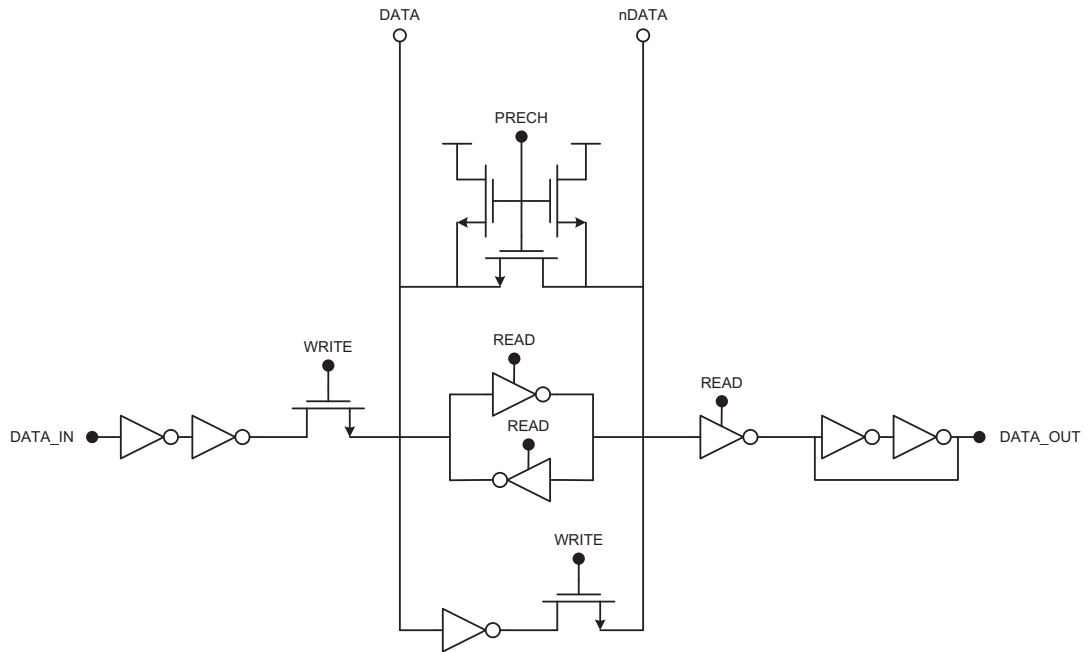


Figure 4.24: Sense Amplifier 1-bit Cell.

Buffer is indeed a big Shift Register, where 36-bit words are moved from left to right until reaching the output pads. A control signal *READ_CLR* makes the Read Buffer to acquire the information from the Sense Amplifiers (it simultaneously reads 1620-bit lines). Then, by pulses on the *READ_CLK* signal, the information is provided to the output pads sequentially. Readouts are (in normal operation) clocked at 10MHz. Thus, a complete image (including dummy pixels) is retrieved in 6660 ($\frac{180}{4} \times 148$) cycles (or 666 μ s for a 10MHz clock), rendering a typical I/O transfer rate of 43MBytes/s.

4.9. Attainable Accuracy

Our algorithm employs two different types of measurements in order to obtain DR enhancement. On the one hand, during the first 15 temporal windows, we somehow codify (non-linearly) the time that takes the photodiode's voltage V_{ph} to reach a fixed reference signal ($V_{ref} = V_{bot}$). On the other hand, during the 16th temporal window, we perform a single-slope A-to-D conversion.

In this later case, we use a 7-bit analog staircase signal in V_{ref} and we digitize the pixel voltage as the code corresponding to the intersection point. It is assumed that the signal (pixel voltage) does not change during this last temporal window (or that its changes are much slower than those of the staircase signal, which is employed for Analog to Digital Conversion). Clearly, since we are targeting a 7-bit representation for the images, the attainable accuracy in this temporal window must be, at least, 7-bit. In other words, all accumulated errors in the operation of the pixel within this last temporal window must remain below $\frac{1}{2}$ LSB of the analog staircase signal. As mentioned already, the design follows a

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

3σ approach. Thus, three times the standard deviation of the total aggregated error (that includes reset error, residual error in the offset cancellation, settling errors, noise, etc.) in the last temporal window must remain below this $\frac{1}{2}$ LSB limit (even when readjusting the operating voltage swing of the ADC for Dark Current effects attenuation, as explained in section 4.5).

Regarding the measurements during the first 15 temporal windows, the situation changes significantly. Indeed, here, the crossing point depends inversely on the photocurrent ($\frac{1}{I_{ph}}$) and thus, the intersection point will spread, as a consequence of aggregated errors, non-linearly as well (depending on the level of photogenerated current). Simply, if the crossing time T_{cross} is given by¹:

$$T_{cross} = \frac{V_{rst} - V_{ref} \pm V_n}{I_{ph}} C_{ph} \quad (4.28)$$

Its variance is easily found to be expressed as:

$$\frac{\sigma^2(T_{cross})}{T_{cross}^2} = \frac{\sigma^2(V_{rst}) + \sigma^2(V_n)}{(V_{rst} - V_{ref})^2} + \frac{\sigma^2(I_{ph})}{I_{ph}^2} + \frac{\sigma^2(C_{ph})}{C_{ph}^2} \quad (4.29)$$

where $\sigma^2(V_{rst}) = \sigma^2 V_{residue}$ (see equation 4.12).

Equation 4.29 gives us some design guidelines to optimize the accuracy in determining the crossing time. First, the residual error for reset establishment must be as small as possible. In addition to that, the larger the reset voltage, and the smaller the reference signal to be crossed, the better the performance (this is what made us use reset and reference voltages defined by the maximum linearity limits of the in-pixel amplifiers). Besides, and obviously, the accuracy in determining the crossing point shows a direct dependency with the mismatch in photogenerated current and integration capacitance (which somehow defines the pixel's PRNU).

The accuracy in determining the crossing time is important since it tells us about the maximum resolution attainable within a temporal window. Let us focus on this in more detail. Consider that we characterize every temporal window, W_j , by its duration, TW_j , and the value of the longest time², T_{end_j} , within this window. As equation 4.29 shows, the error in measuring the intersection point increases with time. Thus, within a temporal window, the worst situation occurs (in most of the cases) at $t = T_{end_j}$. Additionally, since we are using a linear assignment of codes during a temporal window (all codes assigned within a temporal window have the same duration) and we must consider the hypothetical possibility of assigning all available image codes within a single window as well, the maximum allowable spread of the error for window j , $\max[\sigma(T_{cross_j})]$ at $t = T_{end_j}$ is given by:

$$2 \times 3 \times \max[\sigma(T_{end_j})] = \frac{1}{2^7} \cdot TW_j \quad (4.30)$$

Thus, if we want to assure that this 7-bit equivalent accuracy limit is always met, we have a simple algorithm to design the duration of each temporal window. Let us consider that what is fixed is the total

¹ V_n represents the input referred noise, which includes reset and amplifiers noises.

² Which, obviously, corresponds to the starting point for the next window as well.

exposure time T_{exp} , and that the duration of the last temporal window (T_{ramp}) is also fixed (defined by how fast can we perform the single slope A-to-D conversion, or in other words, how fast can we create the 128 levels staircase signal, distribute it to the array, and do the comparisons properly). Let us denote by $TW = T_{exp} - T_{ramp}$ the time available for the first 15 temporal windows. Under these circumstances we would proceed as follows:

1. We would evaluate the error (either theoretically), using equation 4.29, or through numerical Monte Carlo simulations¹ at the end of the last (15th) temporal window. This gives us the value for $\sigma(T_{end_{15}})$.
2. Now, using equation 4.30, we program the duration of the 15th temporal window to be:

$$TW_{15} = 6 \times 2^7 \times \sigma(T_{end_{15}}).$$
3. This also gives us the end point for temporal window TW_{14} , since $T_{end_{14}} = T_{end_{15}} - TW_{15}$.
4. Now we evaluate the error $\sigma(T_{end_{14}})$, and define the duration of this temporal window as we did before:

$$TW_{14} = 6 \times 2^7 \times \sigma(T_{end_{14}}).$$
5. And the process is iterated until defining all the temporal windows.

Two reasonable questions can immediately arise from the method described above. The first one, and the most straightforward is what happens if we run out of available time. That is, we reach a point where the duration of already allocated temporal windows is larger than TW , ($\sum_{j=k;k>1}^{j=15} TW_j \geq TW$). In this case, the answer is simple, we cannot allocate 15 temporal windows within this time frame. Or, in other words, the accuracy of our system is designed for a dynamic range that is much wider than the one required by this setup (we should need to increase the integration time to cover lower illumination ranges in order to have a real need to allocate more temporal windows).

The second question is a little bit more controversial. Part of the problem created in the previous situation comes from the fact that we are using the 3σ constraint to allow the system to allocate all possible output codes in the last evaluation cycle (128th) for all the temporal windows (i.e. we apply the 3σ constraint at the worst case $[T_{end_j}]$ in every window). However, this situation is very unlikely. Indeed, for an unbiased distribution of illuminations, the ratio of photocurrents within a temporal window crossing the reference signal in the last (128th) evaluation period is given by:

$$\rho = \frac{1}{128} \frac{T_{end_j} - TW_j}{T_{end_j} - \frac{TW_j}{128}} = \frac{1}{128} \frac{1 - R}{1 - \frac{R}{128}} \quad (4.31)$$

where the factor $R = TW_j/T_{end_j}$ represents the duration of a temporal window versus the accumulated exposure till this point. In order to illustrate this, let us consider the temporal setup in figure 4.25(a), which corresponds to a typical configuration for 40ms exposure. In this case, see figure 4.25(b), the ratio of photocurrents within a temporal window crossing the reference signal in the last evaluation cycle is

¹This is what we really did in practice due to the lack of PRNU data at design time.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

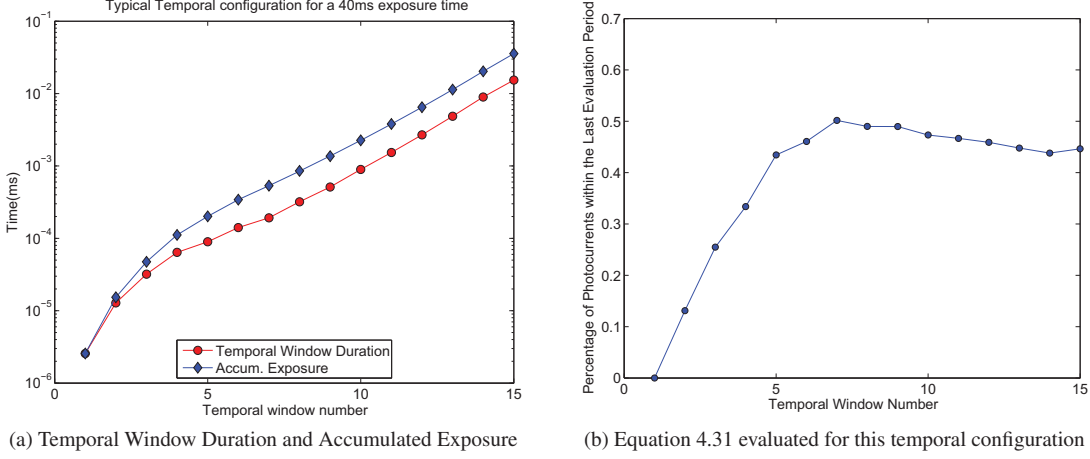


Figure 4.25: Exemplary Temporal Configuration for 40ms Exposure

only about 0.5% for mid to long temporal windows, and even lower for the shorter exposure, suggesting that we could somehow relax the 3σ specifications limit at these points. Let us explain this more deeply.

If we relax the constraints at the last evaluation interval in each temporal window to a $k\sigma$ (with $k < 3$) approach, we define each of the 128 evaluation periods within a frame such that:

$$2 \times k\sigma(T_{endj}) = \frac{TW_j}{128} \quad (4.32)$$

and the probability of committing an error (cases codified within an incorrect evaluation cycle) is increased with respect to the 3σ case by a factor:

$$\rho_{k\sigma} = \frac{\int_{k\sigma}^{\infty} G(x)dx}{\int_{3\sigma}^{\infty} G(x)dx} \quad (4.33)$$

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}}$$

or, using the well-known complementary error functions ($\text{erfc}[x]$).

$$\rho_{k\sigma} = \frac{\text{erfc}(k\sigma/\sqrt{2})}{\text{erfc}(3\sigma/\sqrt{2})} \quad (4.34)$$

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \cdot \int_x^{\infty} e^{-u^2} du$$

We can now quantify this for a few illustrative examples¹. Table 4.1 shows the values of $\rho_{k\sigma}$ for some typical $k\sigma$ configurations. The way to interpret these results is simple. Consider for instance the case $k = 2.326$, the result says that one out of every 50 cases will be out of the interval, and therefore

¹Numbers taken for a few quite often used areas under the Gaussian bell curve.

it will produce a codification error. Besides, since this situation will happen only for some 0.5% of the photocurrents crossing within this temporal window, the probability of committing an error in this temporal window can be estimated to be only $1/50 \times 0.5\%$, or one out of every 10.000 cases. This may, or not, be significant (we have 26.640 pixels in the array) depending on the application and final user. Table 4.1 also tells us approximately how much the situation has worsened as compared to the 3σ case. For this same value of k (2.326), it produces 7.40 pixels with a codifying error ($\pm 1LSB$) for each pixel having an error when using 3σ specification.

Figure 4.26 shows the standard deviation (100 runs of a Monte Carlo simulation of the complete design of the pixel are used for the calculation of sigma) of the crossing time versus its corresponding average value for photocurrents ranging from 50fA to 50nA, 3 points per decade. For most of the time, the standard deviation of the error evolves as expected in equation 4.29, rising linearly with the

k	out of the interval (one in)	$\rho_{k\sigma}$
1	3.151	115.53
1.281	5	74.08
1.645	10	37.05
2	21.977	16.85
2.326	50	7.40
2.575	100	3.70
3	370.398	1

Table 4.1: Some typical values for $\rho_{k\sigma}$

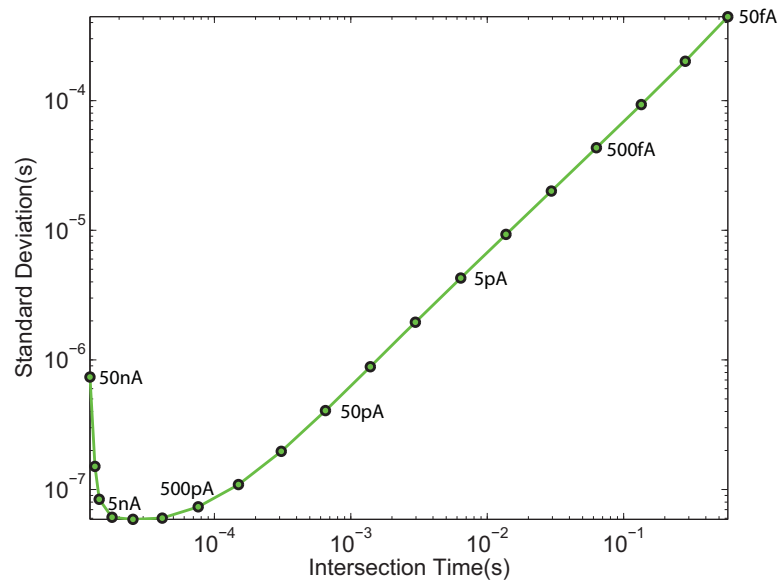


Figure 4.26: Standard deviation of intersection time (T_{cross}) for $V_{ref} = 0.7V$.

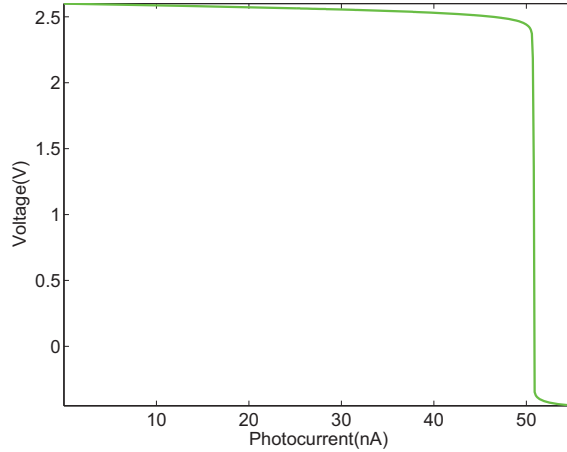


Figure 4.27: Reset voltage versus photogenerated current.

intersection time ($\sigma[T_{cross}] \propto T_{cross}$). However, the situation changes significantly for the very short crossing times. Here, we see that the error starts from a high value and drops until reaching a minimum point and then starts to increase as expected. This unexpected phenomenon is based on the fact that for large photocurrents (short crossing times) the feedback loop in the pixel has more difficulties to reset the photodiode. Indeed, since for low power operation, we lowered the biasing of the amplifiers to just 50nA, the feedback loop is simply unable to establish the reset voltage properly for photocurrents above 10 to 20nA. Figure 4.27 illustrates this effect by plotting the obtained reset voltage versus the photogenerated current.

4.10. Control of the Operation

4.10.1. Image Capture Operation

Though it will be thoroughly described in the next chapter, most of the control of the chip is performed off-chip by an FPGA, which is also responsible for establishing communications through an USB link with the PC where images (or tests) are going to be stored¹. The programmed operation in the FPGA also includes the calculation of the levels per bin from the TSI and the LUT retrieval of the position of decrement for the TM curve (see chapter 3 for details).

Therefore, image capture is possible thanks to the combined action of the FPGA and the TVHC chip. A flow chart summarizing the most important operations for image capture is shown in figure 4.28. The sequence of major steps is:

- To set levels per bin for the first frame. This is 8 by default in the case of using 16 bins ($8 \times 16 =$

¹The system's PCB board also has limited memorization resources that allow for storing nearly 40 images without the need of a PC.

128), as this first frame is only intended for TSI acquisition.

- To set the temporal configuration of the bins. This information is received from the computer via the USB port. Alternatively, it can be stored in the "by default" system setup of the bitstream file, which defines the configuration of the FPGA.
- To program the references for the on-chip DAC. This is done by writing the corresponding registers in an external DAC. This external DAC generates the voltage levels for V_{rst} , V_{bot} and V_{top} , which are the analog inputs for the DAC on-chip.

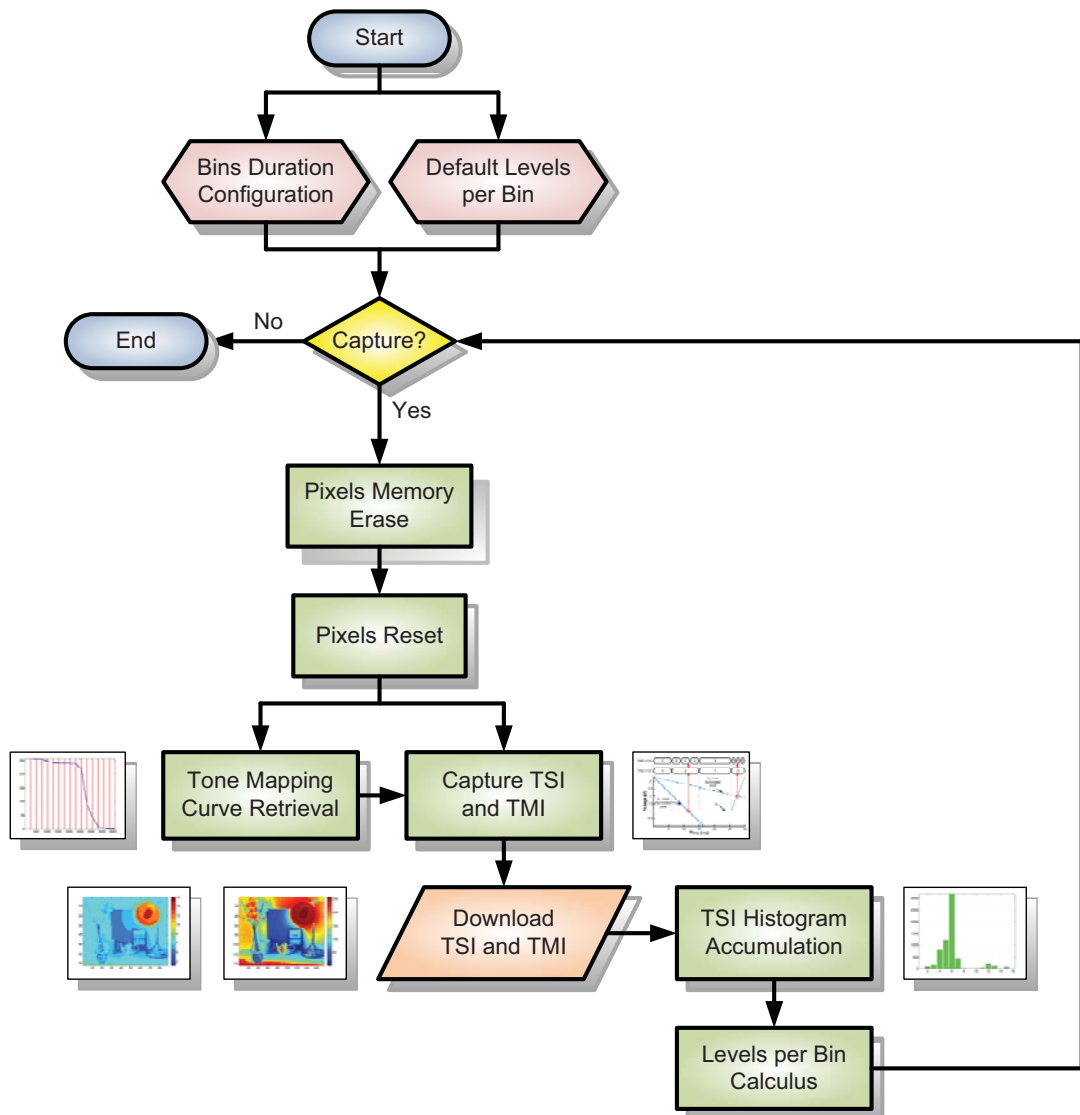


Figure 4.28: TVHC Operation.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

- To initialize pixel memories to the maximum value, 127 and 15 (gray coded), for TMC<6:0> and TSC<3:0>, respectively.
- To reset the photodiode's capacitances.
- To disconnect the direct path from V_{dac} to V_{ref} row lines and to connect it through the row-wise distributed buffers.
- To produce the fastest possible drop from V_{rst} to V_{bot} (this is done by changing the status of the *DAC_MODE* selection bus).
- To connect the direct path from V_{dac} to V_{ref} and to wait for charge redistribution stabilization. When this process settles down, we power off the row-wise amplifiers to save power.
- To start the capture of TMI and TSI. The evaluation of the V_{ph} voltage is usually clocked (by signal *nEVAL*).
- In the last bin, the voltage measurement bin, the following steps are required in every evaluation:
 - To connect the distributed amplifiers and disengaging direct path V_{dac} to V_{ref} .
 - To produce a pulse on DAC_CLK to increase one step the value of V_{dac} , and to wait for the buffers to write this voltage into their corresponding rows with the desired accuracy.
 - To connect the direct path and wait for charge redistribution.
 - To evaluate the $V_{ph} - V_{ref}$ by means of *nEVAL*.
- At the end of exposure, change the configuration of the sense amplifiers from write to read and start downloading.
- The histogram of TSI is calculated during the readout. No extra time is required, and, theoretically, one should not need to store this TS image (though we definitely do it for testing and evaluation purposes).
- Once the histogram is ready (at the end of the download) the FPGA calculates levels per bin for the next capture.
- (Optionally) TM and TS images are sent to the PC via an USB link.
- Repeat for the next frame (with the same setup for temporal windows). If illumination conditions change drastically, the user can scale the duration of each temporal window (except in the case of window 16th, which remains unchanged) keeping the ratios between them. Further actions, as changing the way of assigning levels per bin could also be adopted but are not automatized in the current version of the control software, which runs over Matlab®.

4.10.2. Timing Requirements

Although crucial for the correct operation of the chip, the temporal requirements in the generation of the many different controls that define the operation of the TVHC are likely too low-level issues to be described in this chapter. We refer the interested readers to appendix A for details.

4.11. TVHC Photograph and Layout

The complete layout of the chip and its microphotograph are shown in figure 4.29 and figure 4.30, respectively. It is noticeable that, the pixel array occupies most of the area of the chip. The Sense Amplifiers, the Read Buffer and Code Generator, are placed in the bottom part of the chip. The distributed Control Signals Buffers are placed on the right side of the array, whereas the DAC and Charge Injection Amplifiers are placed on the left. The two small blocks at each side of the Sense Amplifiers are (left) the block that generates the signals for Sense Amplifiers operation, and (right) the Code Generator for the ADC through Sense Amplifiers.

4.12. Other Tone Mapping Hardware Implementations

HDR imagers generally implement tone mapping as a fixed simple compression, typically using logarithmic or exponential curves. This has been explained already in chapter 1, where multiple examples of HDR imagers are depicted. However, this approach is not very well suited because it results in a lack of contrast in highly compressed areas, while some digital codes of the final representation are simply not used (it is quite common to see that the histogram of the images provided by these systems do not span over the whole available range).

If we focus our attention on more complex approaches, most of the existing tone mapping algorithms have been traditionally created by people from the computer graphic research field. Therefore, and most usually, hardware resources are considered "granted and available" since the very beginning, and typically these algorithms are executed on high-end PCs embedding ultra-powerful graphic cards, whose use is justified by the vast amount of operations to be executed. It is also worth to mention that, usually, these algorithms do not pay too much attention to which way the image is obtained, and in general, they take as the starting point a set of image data from several low bit images (multiexposure) or images codified in HDR format (such as radiance or mantissa-exponent).

Leaving aside the pros and cons of having the sensor and processing resources on the same chip, if we only look at the hardware employed for tone mapping execution, we find the following remarkable options in the literature:

- General purpose microprocessor [102].
- PC high-end Graphic Cards [135].
- Dedicated processors (such Graphic Processor Units [GPU] [136]).

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

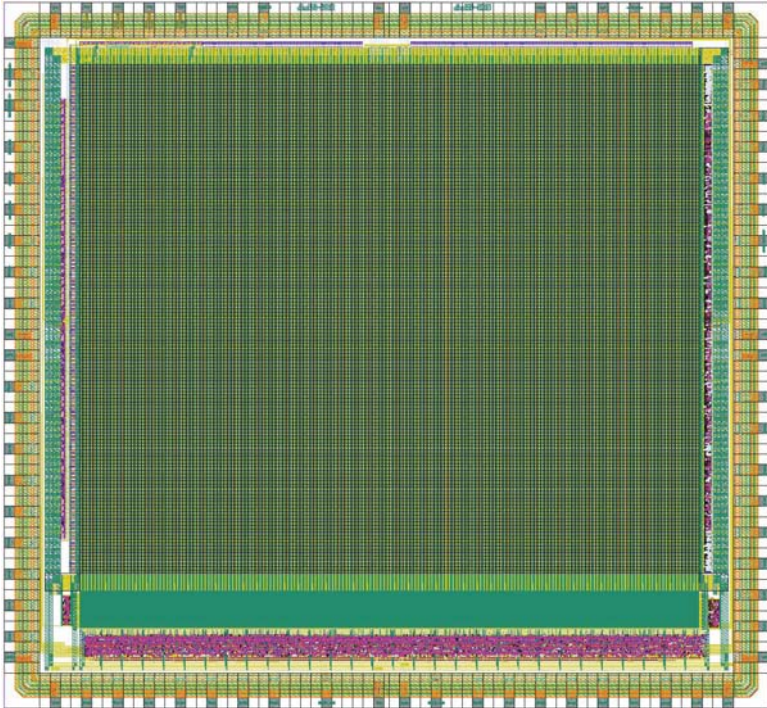


Figure 4.29: TVHC Layout.

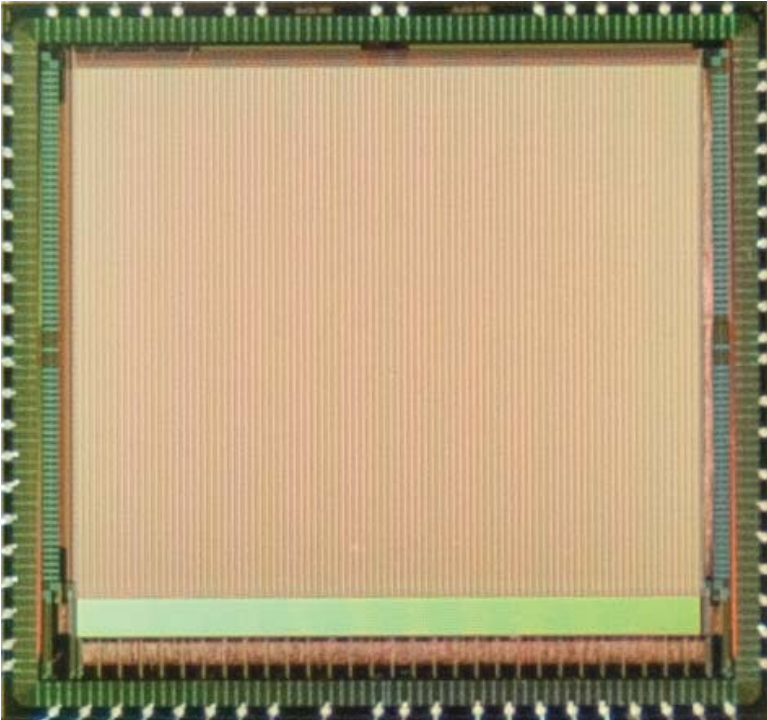


Figure 4.30: TVHC Photograph.

-
- DSP.
 - FPGA [137].
 - Dedicated system (as in our case) [138].

Most of these hardware approaches are generally meant for tone mapping relying on the post-processing of LDR multiple exposure images. Due to their computational complexity, interactive tone mapping techniques have received relatively little attention. Nevertheless, real-time tone mapping functionality embedded within the camera would be desirable, though, for this purpose, the computation should be reduced. It is only in recent times that some commercial systems have started to include HDR capabilities. Some remarkable examples are commented below.

In the field of advanced cameras, Sony A550 and Pentax K-7 offer a HDR multi-exposure mode in which several images are captured, combined, and displayed to improve the final quality of the image in the shadows and high light regions. They use a typical autoexposure and bracketing approach. However, capturing 2-3 consecutive images takes at least one second to be done (this is mostly independent on the real exposure for each frame since times devoted to image download, exposure calculation, etc. do not scale with the exposure). Besides, a reduced extension of the dynamic range is achieved because there is usually a maximum automatic bracketing of $\pm 2EV$ (exposition) variation.

Mobile phones with relatively complex processing resources can also be employed for tone mapping realization. For example, the Nokia N900 model [139] has been studied to include an adapted tone mapping algorithm in its functionality. However, hardware constraints, such as required memory, limit the performance. Another example is found in the iPhone4 [140], which incorporates a commercially released HDR mode at the operating system level (since iOS 4.2). This HDR mode combines 3 images with different exposures. This smart phone provides high computational resources thanks to its Advanced RISC Machine (ARM) processor and a dedicated GPU.

Regarding surveillance applications, Pixim cameras include [141][142] a digital image sensor plus a digital image processor chip. In these systems illumination level, white balance, and scene range estimation, are studied to adjust exposition time from multiple image sampling. Also, the already classic Photonfocus Linlog [143] technique combines linear response at low illumination levels and logarithmic compression at high intensities. However, as already explained, this fixed compression is not optimum as it does not adapt to the content on the scene.

As to other research approaches are considered, to the date and to the knowledge of the author, no other system has been reported with a scene adaptive focal plane in-pixel tone-mapping technique. Therefore, no direct comparison can be made with other approaches, and we find this one of the principal innovations of our work.

Furthermore, the main bottleneck of typical post-processing tone mapping approaches is the large amount of data movements from processor to memory and vice versa, which is needed in order to calculate the final image. In our solution, the final image is the only thing to download, as the hardware is HDR tone mapping oriented. We acknowledge that the user also needs to download the Time Stamp Image in order to calculate the Tone Mapping Curve in the current implementation of the TVHC. However,

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

this extra image is reduced by down-sampling ($\frac{1}{4}$) and only employs 4-bit per pixel.

It is interesting to mention that, our system stores the images in-pixel, which is a great advantage towards vertical integration technologies. In these 3D-IC technologies, memory-to-processor bandwidth will not be a significant problem at all because data will be moved in a highly¹ parallel way from the pixel to a post-processing layer. Subsequently, our system could also be used as the first layer in an approach inspired by the stacked parallel architecture of the human eye. It would mimic the reduction of data² from the first capture layer to a later post-processing layer without losing the details of a HDR scene. In order to have this functionality, the system will have to digitally³ store the image at the pixel level (as is our case) so that the next tier can directly process it. Consequently, no reconstruction of data is necessary, as it will probably be the case in Pixel Event approaches [144]. In our case, the final image does not need to be moved through a bottleneck connection to a memory or an external tone mapping processing block. The data is ready just to be used for the next processing stage in a future multi-tier 3D packaged chip.

4.13. Conclusions

This chapter has introduced the TVHC chip, a QCIF imager that implements the tone mapping algorithm described in chapter 3. The chip has been designed in an almost standard 0.35 μ m CMOS technology (it only adds an ARC layer on top and optimized 14 μ m thick EPI substrate), which implies a substantial lower cost than that of the specialized CMOS Image Sensor Technologies.

The chip has been designed so that most of the HDR tone mapping algorithm is implemented on it. Only TS histogram calculation, level-per-bin assignment and LUT are moved to the FPGA, which controls image acquisition and communications with a PC.

Pixels include an auto-zeroing technique, which reduces the fixed pattern noise (basically mismatch effects) created by the in-pixel analog circuitry. TS and TM information are stored on pixel in Static RAM blocks, allowing for very long exposure operation and featuring virtually zero crosstalk between the stored data and the incident light⁴.

An automatic dark signal contribution mitigation scheme has been implemented to enhance the visual quality in the dark areas. This functionality has been implemented by a simple principle, just adapting the operating range of the ADC which digitizes pixel information.

Global analog reference to the pixels, V_{ref} , is dynamically distributed to allow for low-power, fast, and precise operation. This is a non-classical approach, because it takes advantage of the redistribution of charge in the high number of V_{ref} row lines, and relies on the fact that, for a sufficient large number of amplifiers, their input-referred offset voltages would have a mean which is near to zero.

¹Either completely in parallel or using a multiplexed architecture where few pixels are sequentially downloaded to a pitch-matching processing stage.

²Not based in reducing the number of pixels to be transmitted but into reducing the number of levels to code a wide dynamic range scene.

³This would probably be a must if we need to work with very long exposures.

⁴DRAM or analog memorization approaches may suffer from a certain data degradation due to the incident light. Indeed drain/source diffusions in the switches will act as small photodiodes that may discharge a little bit the storage capacitors.

The in-pixel SRAM storage allows for long exposure shots and parallel computing in future vertical integration technologies. The system is just ready to be incorporated in future multilayer CMOS chips. In this case, the size of the pixel can be scaled-down significantly by moving the digital memorization resources to a tier underneath the sensor using only one TSV¹ between tiers (the result of $V_{ph}-V_{ref}$ comparison).

¹TSV stands for Through Silicon Via.

4. TVHC: A HDR TONE MAPPING IMAGER IN STANDARD CMOS TECHNOLOGY

Chapter 5

Experimental Results

5.1. Introduction

In order to test the performance of the TVHC chip, three different types of experiments have been performed. First, the test environment, which is compound by the TVHC chip inserted in a Printed Circuit Board (PCB), has been placed on an optical table. Then, photometric characteristics have been measured applying a stable light source. These characteristics are spectral response, sensitivity, dark current and dynamic range. Second, a laser beam has been applied on top of the chip in order to test the effect of shadows in this heterogeneous array of pixels. Third, several HDR scenes have been captured in order to illustrate the results of the implemented tone mapping algorithm.

In this chapter, the experimental results obtained in the TVHC chip measurements will be presented. First, a description of the test board and its functionality is provided. Second, the calibration of the system is explained. Third, photometric measurements performed on an optical table will be shown along with shading measurements. Finally, some HDR scene captured images are shown performing a comparative analysis with 3 commercial systems.

5.2. Experimental Setup

In this section, we will present the different setups used for the characterization of the TVHC chip, including the PCB designed to host it. This section is organized as follows: (1) First, the PCB which has been designed to host the TVHC is presented. (2) The optical test setup is introduced afterwards. (3) The test setup for image capture and the effect of the lenses is described.

5.2.1. TVHC PCB Host

The core of test environment is a PCB system, whose control is executed by a Field Programmable Gate Array (FPGA), which also performs the calculation of levels per bin and stores the LUT for TMC generation. Figure 5.1 shows the general architecture of the test environment. Via the USB port, the

5. EXPERIMENTAL RESULTS

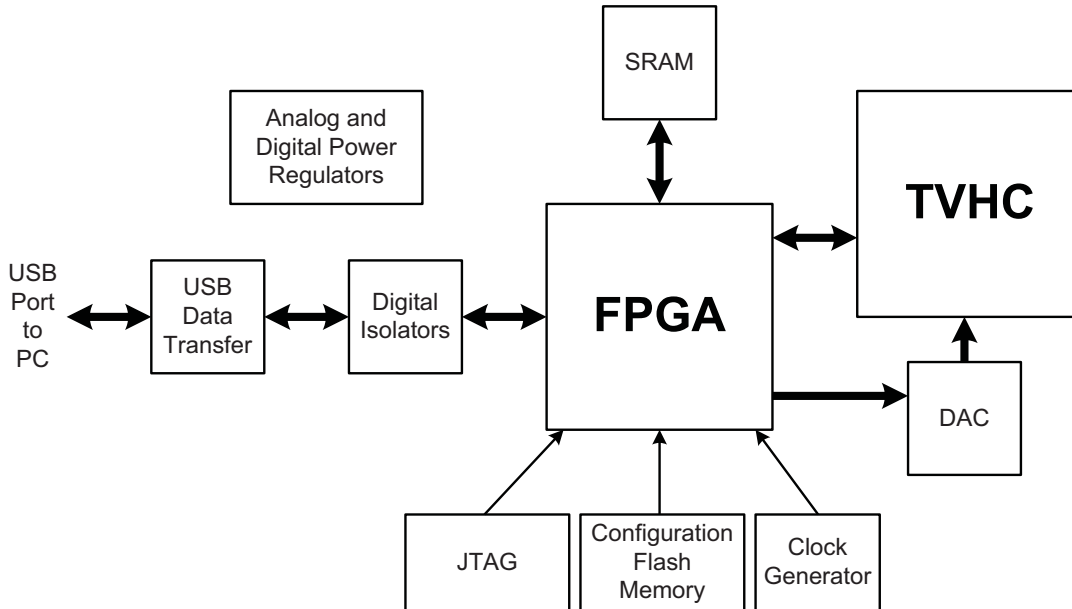


Figure 5.1: TVHC PCB general scheme.

configuration is received from the PC and the captured images (TSI and TMI) are sent to the PC. The test environment is designed so that the PCB receives the following information from the host PC:

- Biasing voltages, V_{rst} , V_{bot} and V_{top} , generated by the external DAC.
- Number of active bins.
- The temporal configuration of the bins.
- Different delays applied to meet timing requirements.

A set of digital isolators have been included between the FPGA and the USB data transfer, in order to avoid malfunctions due to direct connection to the FPGA.

Among many other peripheral devices included, we find an external RAM to store the captured images to be sent later through the USB port, and an external DAC that generates the analog reference voltages for the chip, V_{rst} , V_{bot} and V_{top} .

The FPGA can be programmed directly via the JTAG connector or programmed at start-up from a configuration flash memory. This FPGA receives the clock via an external clock generator. Several voltage regulators supply the board from the general 5V supply, separating digital and analog supplies. The board can also receive the common 5V supply from the USB port, under low power consumption operation.

Figure 5.2 shows the schematic of the TVHC PCB and figure 5.3 is a photograph of the board with assembled lens. The main devices depicted in the schematic are:

-
1. **FPGA:** Xilinx Spartan 3 XC3S400 400K System Gates, 288K Block RAM, 56K Distributed RAM Bits, maximum I/O User 141, PQ208 [145].
 2. **TVHC:** Time Voltage Histogram Camera under test, 84 JLCC (Lead “J” Shaped).
 3. **DAC:** Analog Devices AD7399, Four Rail-to-Rail Channels, Serial-Input 10-Bit DAC [146].
 4. **Digital Isolators:** Analog Devices Quad-Channel Digital Isolators with speed CMOS and monolithic air core transformer technology: Analog Devices ADuM1400 and ADuM1402 [147] plus Texas Instrument Hex Inverter SN74LVC04 [148].
 5. **SRAM:** Brilliance Semiconductor Inc. (BSI) BS62LV8001 SRAM 1M x 8 bit, i.e. 1,048,576 words of 8 bits [149].
 6. **USB Data Transfer:** FTDI FT245BL, Single Chip USB Parallel FIFO bi-directional Data Transfer. Maximum data transfer rate to 300Kilobyte/Sec with Virtual Com Port Drivers. Maximum transfer data rate to 1MByte/Sec using FTDI D2XX direct drivers [150].
 7. **Analog and Digital Voltage Regulators:** this circuitry generates stable 1.2V, 2.5V and 3.3V for digital supply, and 3.3V for analog supply. These voltages are generated by one Linear Technology LT3021a [151] and two National Semiconductor LM1117 [152]. XP Power IL0505 [153] is also included to allow for supply the general 5V via the USB port.
 8. **Configuration Flash Memory:** Xilinx XCF02S 2Mbits Platform Flash In-System Programmable Configuration PROM [154].
 9. **JTAG:** a connector which is an interface to download FPGA programming bitstreams into the FPGA.
 10. **Clock Generator:** the 50MHz clock signal is generated by a quartz crystal Euroquartz XO91 [155].

5.2.2. Optical Setup

In order to accomplish the measurement of photometric characteristics, the camera board with the TVHC Chip has been introduced in an optical test environment similar to that described in chapter 2 for the SCU chip.

In order to measure the high dynamic range of the TVHC chip, metallic neutral density filters were required to be used. These filters, in addition to the described setup, produce a reduction of the amount of light in a controlled way. They have been placed in the same mechanical structures as the band-pass filters. Their densities are 0.10, 0.30, 0.50, 1.00, 2.00, 3.00 and 4.00, (010FN4650 - 400FN4650), which correspond to nominal percental transmission factors of 79.5, 50.0, 31.6, 10.0, 1.0, 0.10 and 0.01, respectively.

5. EXPERIMENTAL RESULTS

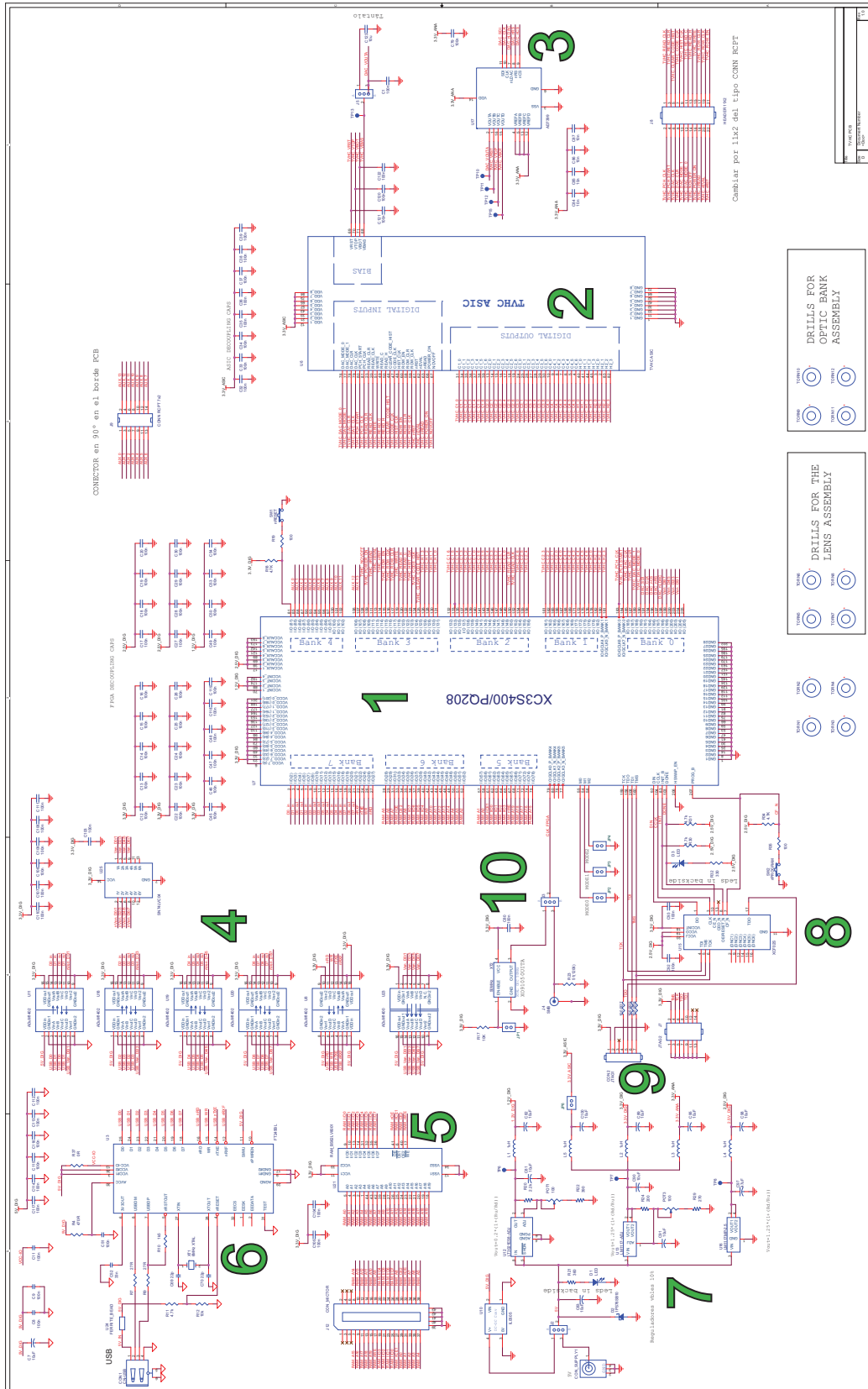


Figure 5.2: TVHC PCB schematic.

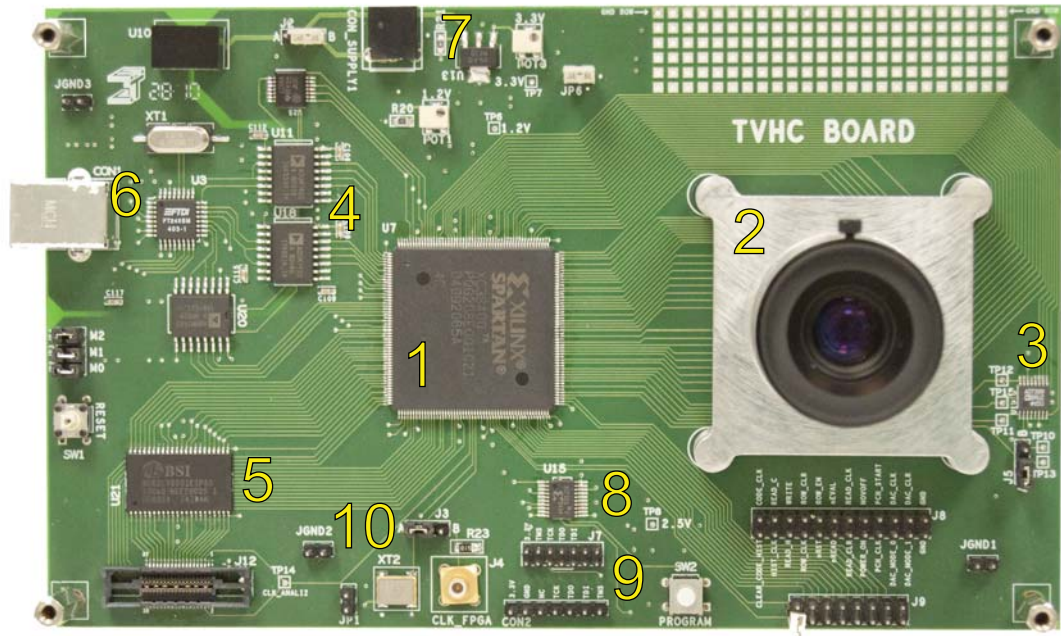


Figure 5.3: TVHC board photograph.

5.2.3. Lens and Optical System

The lenses applied to the system have fixed focus length and are attached to the TVHC PCB by means of a holder over the ASIC chip. The available optical lenses are 2/3" format with focal lengths of 6mm, 8mm, 16mm, 25mm and 50mm, and maximum aperture f-number f/1.4, f/1.4, f/1.4, f/1.6 and f/1.8, respectively.

This optics will have a high impact in the optical behavior of the setup. The light arrives to the chip in a focalized way, which will affect the uniformity. The most common aberrations, distortions or effects of lenses are [156]:

- **Spherical aberration:** Rays at the edge of the lens (marginal rays) come to focus closer to the lens than rays parallel to the axis or center (paraxial rays) do. This causes blurred images.
- **Astigmatism:** The lens does not focus both vertical and horizontal lines on the same plane. It will cause that lines of equal brightness appear different horizontally or vertically.
- **Coma:** It causes parallel oblique rays passing through a lens to be focused not as a point, but as a comet-shaped oval image.
- **Curvature of Field:** It results in the image from the lens being formed not flat but as a curved surface. Therefore, when the center of the image is in focus, the edges are not, and vice versa.

5. EXPERIMENTAL RESULTS

- **Curvilinear Distortion:** This distortion makes straight lines of objects in the scene to appear as curve-like shapes in the captured image. Although distortion can be irregular or follow patterns, most commonly they are radially symmetric due to the symmetry of a photographic lens. The main radial distortions are:
 - Barrel Distortion: Image magnification decreases with distance from the optical axis. The apparent effect is an image which has been mapped around a sphere.
 - Pincushion Distortion: Image magnification increases with the distance from the optical axis. The visible effect is that lines that do not go through the centre of the image are bowed inwards, towards the center of the image.
- **Chromatic Aberration:** It is caused by different behavior depending on the wavelength of the ray of light. There are two types:
 - Longitudinal chromatic aberration: Lens focuses wavelengths (colors) at different planes on the lens axis. Shorter wavelengths come to focus in front of the focal plane, long wavelength behind.
 - Lateral chromatic aberration: is the lateral displacement of color images at the focal plane. This type of aberration is caused by different sizes of images produced by different colors, even though the image is all on the same plane.
- **Lens Flare:** unwanted rays of light which scatter within the lens. It usually creates artifacts such as multiple circles in a line leading from a bright light, such as the sun.
- **Vignetting:** It produces a radial falloff of intensity from the center of the image. It is known that at their maximum aperture settings, even high-quality fixed focal length lenses transmit 30% to 40% less light at the corners of an image than at its center [157].

All image captures in the section 5.4 have been performed on the maximum aperture of the lens in order to get as higher dynamic range as possible in the sensor. However, this issue and the fact that the presented system is optimized to keep details will make more noticeable some optical distortions of the optic lenses. The most noticeable distortions in our case will be the curvilinear barrel distortion, little lens flare and vignetting. Moreover, the high aperture will imply a short field depth and therefore only the target objects are focused while the backgrounds are usually blurred. However, these effects are derived from the physics of optics and it should not be considered as part of the behavior of the sensor. They can only be minimized by special optics designs.

5.2.4. Operation Modes

The TVHC can work in three general modes of operation:

- Tone mapping mode: implementing the proposed algorithm for HDR operation.

-
- Linear mode: images are captured without compression. A 7-bit linear conversion at the end of exposure. After a delay time over exposition time, the last and only bin will be present for a very short period of time performing a fast ramp-up in the reference.
 - Fixed compression mode: the last and only bin will spread over the whole exposition time. The applied compression is the result of the intersection of the discharged signal with a monotonically increasing reference, which is an increasing compression toward the brightest illuminations. However, as it is a fixed compression, it is not optimized and therefore will not be used in the measurements presented in this work.

The linear mode is available by means of proper time bin configuration. If a single bin is chosen, then the capture delay will determine the exposition time. The last bin with the ramp-up voltage will be then the only bin. It will perform a single slope AD conversion, as the typical method for linear image capture [158]. This mode has been used in the radiometric measurements, as non-linear compressed data cannot be used to this purpose.

5.2.4.1. Tonemapping Operation Modes

Some possible modes of operations have been introduced by varying the method for assignment of the levels per bin, which have been described in subsection 3.4.3:

1. Equal distribution
2. Weighted
3. Bin threshold
4. Avoid concentration in one bin
5. Weighted with bright priority
6. Low populated priority
7. Weighted with minimum threshold
8. Non-linear levels adjustment
9. Weighted with minimum low light priority

5.3. Measurements

5.3.1. Spectral Response

The spectral response measurements have been performed by means of a set of bandpass filters from 400 to 750nm with FWHM of 10nm (as in SCU case). These filters have been located between the TVHC chip and the stabilized light source.

5. EXPERIMENTAL RESULTS

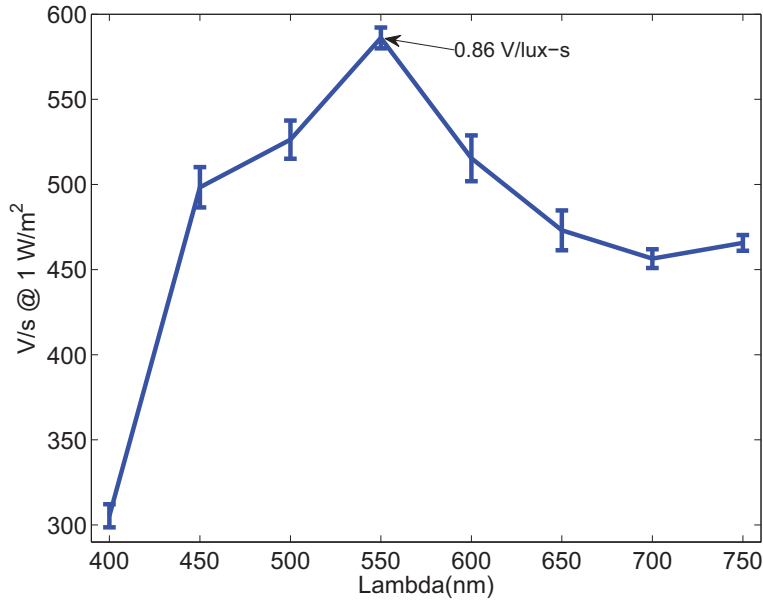


Figure 5.4: Spectral Response.

Figure 5.4 shows the result of the measurements in radiometric unit ($\frac{W}{m^2}$). The spectral response shows a decrement towards the lower part of the visible range, which is consistent with the typical spectral response of a PN junction in silicon. The photometric response at 555nm is usually given as a figure of merit, as this wavelength corresponds to the sensitivity peak of the human eye. In this case, the closer value is at 550nm which is $0.86 \frac{V}{lux-s}$.

5.3.2. Dark Discharge

Due to the architecture of the chip, the capacitance of the integration node of the pixels cannot be directly measured. Instead, the discharge of the integration node due to this parasitic current is measured. This measurement is directly observable in the behavior of the pixel and therefore more useful in terms of control operations, such as compensation or attenuation of its effects.

The contribution of the dark current has been measured covering the chip and in darkness to prevent the influence of light. Figure 5.5 presents the dark discharge over time at ambient temperature (25°C). This measurement implies an average dark signal of $10.8 \frac{mV}{s}$. Figure 5.6 shows the result of the calibration (see subsection 4.5.1) for 1σ (standard deviation) of the previous measurement at 8s. Clearly, it implies a complete cancellation of the dark contribution to the captured image till 3.5ms. Derived from this data, figure 5.7 shows the Signal to Noise Ratio (SNR) in linear mode before and after calibration.

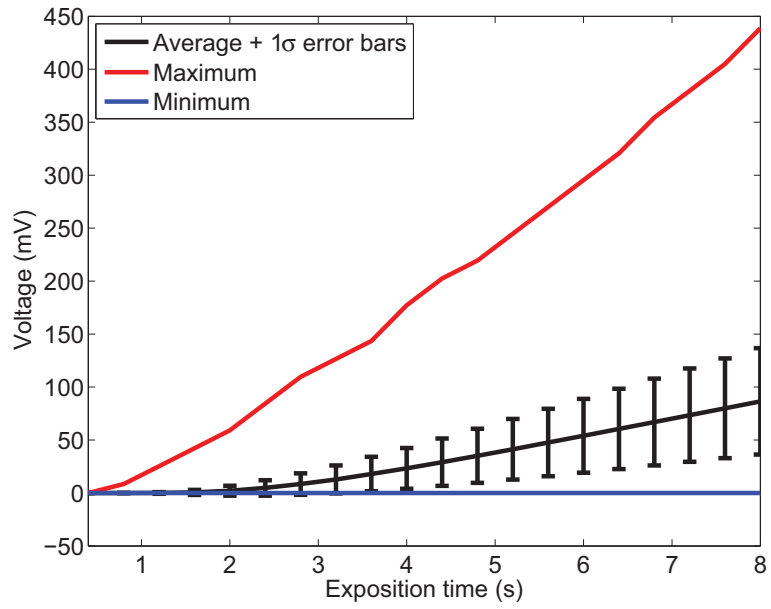


Figure 5.5: Dark discharge.

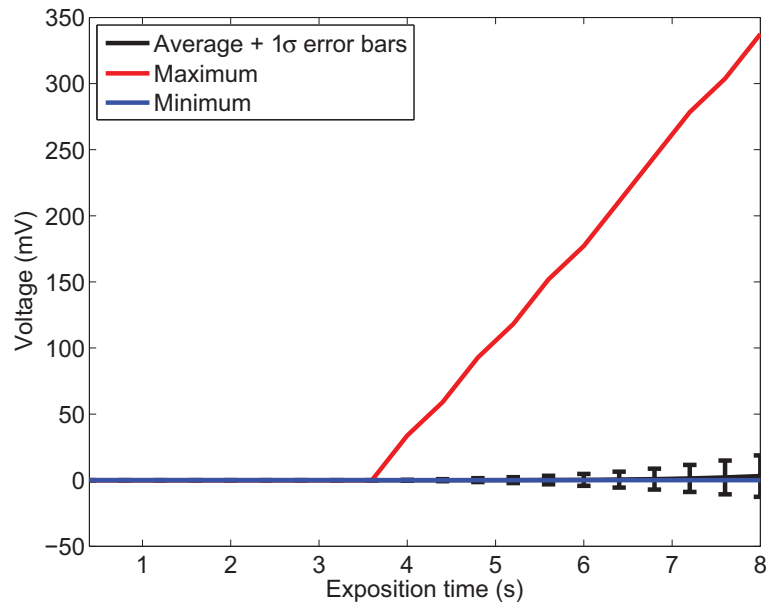


Figure 5.6: Dark discharge after calibration.

5. EXPERIMENTAL RESULTS

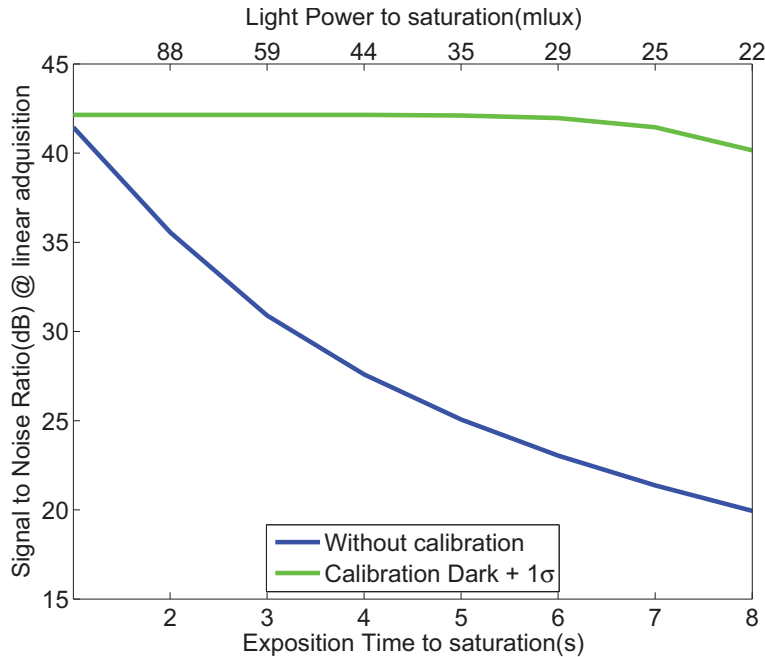


Figure 5.7: Calibration SNR comparative.

5.3.3. Sensitivity

This section presents the sensitivity of the chip under white light excitation. Light source is a 6334 @ 250W QTH Lamp from Newport Corporation, whose spectral irradiance has been shown in chapter 2. The data are provided in photometric units. In order to measure the photometric response to white light, the lux measurement has been performed measuring the spectral distribution of the radiation between 400 to 750nm in radiometric units, and later applying the CIE Spectral Luminous Efficiency Function for Photopic Vision, which is typically named $V(\lambda)$. This operation requires multiplying the radiometric spectral distribution by the photopic response curve, integrating the resulting curve and multiplying the result by a conversion factor of 683. Figure 5.8 shows the measurements for different illuminations, where the error bars are defined for one standard deviation. Considering these data, the average sensitivity results in $5.79 \frac{V}{lux \cdot s}$.

5.3.4. Dynamic Range Measurements

The dynamic range of the chip has been measured via a combination of measurements, applying white light, and considering the relative increments of the dynamic range. The direct measure of the dynamic range of the chip requires creating an artificial image whose DR expands over that of the chip. This is really not possible in our laboratory. Besides, the targeted +140dB conditions are also hard to be obtained indoor. Additionally, the power meter available Newport 1930-C in combination with the silicon detector Newport 918-SL has a lower limit for light power detection of $3 \frac{pW}{cm^2}$ whereas the lowest

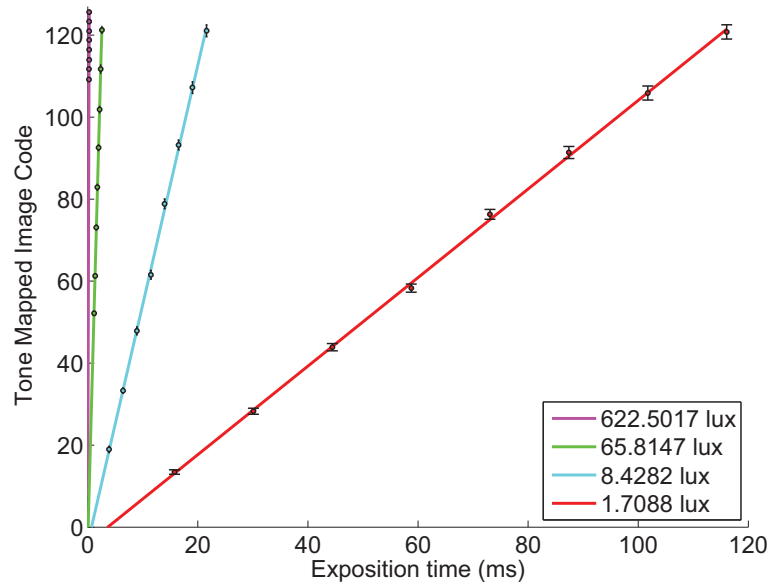


Figure 5.8: Response to white light.

measurable signal by the chip goes below this value.

Due to the dynamic range limitations of the measuring instruments, a combination of neutral density filters has been used. Figure 5.9 illustrates the relative measurements which have been obtained applying the following procedure:

1. A very bright white light has been applied in order to measure the highest limit. This measurement has been performed in the tone mapping mode. Therefore, the highest measurable light corresponds to a value of code 126 with the shortest first bin possible. It is the first distinguishable highest measure. The light power applied in this measurement in photometric units is 55329lux.
2. The light power has been reduced to 5875lux in order to have a lower point of reference to later apply the neutral density filters.
3. The light power is reduced to 311 μ lux by means of neutral density filters. Two neutral density filter with very low transmission are applied and the image is measured in linear mode. The transmission of the neutral density filters has a tolerance. Therefore, the real transmission of the filters been measured. The experimental data of the transmission of the filters indicate that the combination of the two filters OD300-OD400 produces a transmission of $3.73 \cdot 10^{-7}$, and therefore about 128 dB between the photocurrents with and without these filters. The resulting signal crosses the ramp after 8 seconds of exposition time in linear mode. The average value with a ramp of 10 μ s steps is code 105.45.
4. The noise floor has been measured in darkness with linear mode. The average value with same time configuration of the previous measurement is code 10.25.

5. EXPERIMENTAL RESULTS

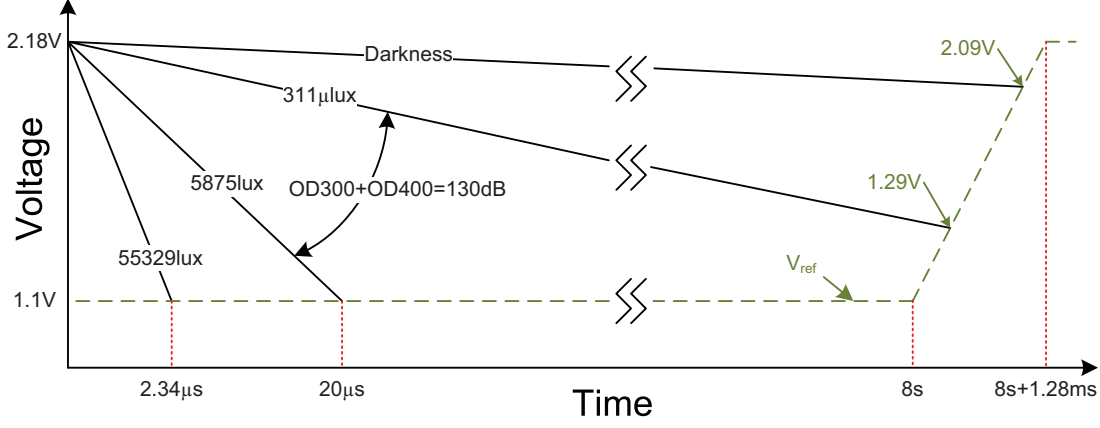


Figure 5.9: Dynamic range measurements.

The expression of the photocurrent is:

$$V_{ph} = V_{rst} - \frac{I_{pix}}{C_{pix}} t \rightarrow I_{pix} = \frac{C_{pix}}{t} (V_{rst} - V_{ph}) \quad (5.1)$$

In order to obtain the relationship between the last two measurements, third measurement (8s exposure) and noise floor, the slope of the ramp signal must be considered:

$$\frac{I_{pix_1}}{I_{pix_2}} = \frac{\frac{C_{pix}}{t_1} (V_{rst} - V_{ph_1})}{\frac{C_{pix}}{t_2} (V_{rst} - V_{ph_2})} = \frac{t_2 (V_{rst} - V_{ph_1})}{t_1 (V_{rst} - V_{ph_2})} \quad (5.2)$$

t_1 and t_2 can be obtained from:

$$\begin{aligned} t_{exp} &= t_{delay} + t_{step}(128 - Code) \\ t_1 &= 8s + 10\mu s(128 - 105.45) = 8s + 225.5\mu s \\ t_2 &= 8s + 10\mu s(128 - 10.25) = 8s + 1.18ms \end{aligned} \quad (5.3)$$

V_{pix} values for the these two measurements are simply given by:

$$\begin{aligned} V_{ph} &= V_{bot} + V_{step}(128 - Code) \\ V_{step} &= \frac{V_{top} - V_{bot}}{128} = \frac{2.18 - 1.1}{128} = 8.44mV \\ V_{ph_1} &= 1.1 + 8.44 \cdot 10^{-3}(128 - 105.45) = 1.29 \\ V_{ph_2} &= 1.1 + 8.44 \cdot 10^{-3}(128 - 10.25) = 2.09 \end{aligned} \quad (5.4)$$

And, therefore:

$$\frac{I_{pix_1}}{I_{pix_2}} = \frac{(8s + 1.18ms)(2.18 - 1.29)}{(8s + 225.5\mu s)(2.18 - 2.09)} = 9.89 \quad (5.5)$$

Finally, the total dynamic range (referred to the SNR=1) is obtained by simply combining the dynamic range between the four measurements, DR_1 , DR_2 and DR_3 .

$$\begin{aligned} DR_1 &= 20 \cdot \log_{10}(55329/5875) = 19.48 \\ DR_2(OD400 + OD300) &= 20 \cdot \log_{10}\left(\frac{1}{3.73 \cdot 10^{-7}}\right) = 128.56 \\ DR_3 &= 20 \cdot \log_{10}(9.89) = 19.9 \\ DR &= DR_1 + DR_2 + DR_3 = 19.48 + 128.56 + 19.9 = 168dB \end{aligned} \quad (5.6)$$

These measurements imply a total dynamic range of 168dB taking the noise floor of SNR=1, and 148dB if we consider that for SNR=10 is the minimum acceptable SNR [159].

5.3.5. Noise Measurements

Noise measurements have been performed by using the standard defined by the Photon Transfer Curve (PTC) method [23]. This method considers the camera system, independently of its complexity, as a black box. In our case, all noise measurements have been done with the chip operating as a conventional linear imager (i.e., we follow the paradigm Reset-Exposure-Single Slope A-to-D conversion-Readout). In order to extract the different parameters, the camera is excited by a white light source or absence of light in order to capture two set of frames. In the combination both set of frames, signal and noise responses are collected in the form of Digital Numbers (camera's digital output codes). The resulting numbers are plotted as a function of the average signal for a variety of exposure times, with a logarithmic distribution, that covers the whole dynamic range, whereas the light power is kept constant. The average signal $S(DN)$ is obtained from the raw images minus the average offset level for every exposure time. Pixel values for each exposure times (with and without light, DN_{ADC_i} and OFF_i) are the result of averaging many samples of the same experiment in order to reduce random effects (as small variations of the incident light power). Thus, the PTC method evaluates the average output code (over the N_{pixels} in an image) simply as:

$$S(DN) = \frac{\sum_{i=1}^{N_{PIX}} DN_{ADC_i} - OFF_i(DN)}{N_{PIX}} \quad (5.7)$$

whereas the total noise for this exposure time is found by calculating the standard deviation of the S(DN) values:

5. EXPERIMENTAL RESULTS

$$\sigma_{TOTAL} = \sqrt{\frac{\sum_{i=1}^{N_{PIX}} [S_i(DN) - S(DN)]^2}{N_{PIX}}} \quad (5.8)$$

Obviously, this total noise contains contributions mainly from read noise, photon shot noise, and Fixed Pattern Noise (FPN). In order to properly identify the sources of noise, results are graphically represented in a log-log axis. In this representation, four regions (or noise-dominated regimes) can be identified:

1. Read noise Regime: Measured under total dark conditions. It is characterized by a line of zero slope for low $S(DN)$ values.
2. Photon shot noise: Noise created by the essential nature of light (Poisson distribution of impinging photons). It is placed at the middle region of the curve. On log-log coordinates, it is characterized by a line with slope $\frac{1}{2}$ (since photon's shot noise is proportional to the square root of the number of photons reaching the sensor).
3. FPN noise: it is associated with in-pixel FPN (as we implement A-to-D conversion at the pixel level). On log-log coordinates, it is characterized by a line with slope 1.
4. Full-well: For increasing signal values, one finds that there is some point where noise start to decrease as saturation of the pixels begins. The beginning of this regime is characterized by a zero slope.

Photon shot noise and the FPN, are expressed as:

$$\sigma_{SHOT} = \sqrt{\frac{S(DN)}{K_{ADC}(e^-/DN)}} \quad (5.9)$$

$$\sigma_{FPN} = P_N \cdot S(DN) \quad (5.10)$$

where P_N is the FPN quality factor. These measurements are initially made in Data Number (DN) units. However, after the PTC curve is measured, these values can be converted to electrons by means of the factor of conversion $K_{ADC}(e^-/DN)$.

The total noise is then defined by:

$$\begin{aligned} \sigma_{TOTAL} &= \sqrt{\sigma_{READ}(DN)^2 + \sigma_{SHOT}(DN)^2 + \sigma_{FPN}(DN)^2} = \\ &= \sqrt{\sigma_{READ}(DN)^2 + \frac{S(DN)}{K_{ADC}(e^-/DN)} + [P_N \cdot S(DN)]^2} \end{aligned} \quad (5.11)$$

Our camera has been exposed to stable light stimulation with exposition times (in ms) taking the values of 1, 2, 3, 4, 5, 7, 9, 12, 15, 19, 24, 31, 39, 49, 62, 78, 99, 125 and 157. Experiments (response under

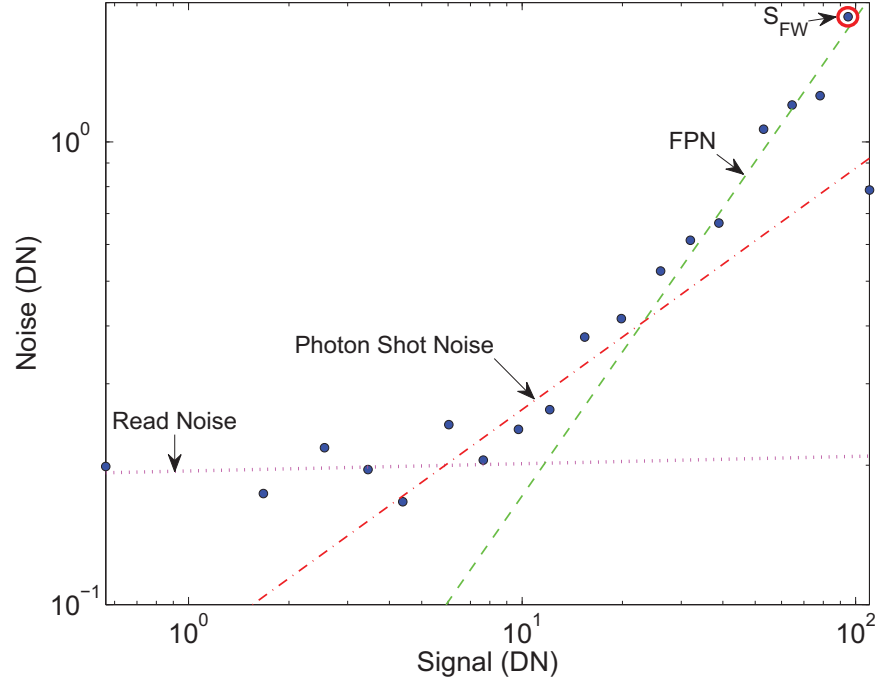


Figure 5.10: Photon Transfer Curve.

illumination and response in darkness) are repeated 100 times for each exposition time and averaged in order to obtain more precise results. The resulting values of $S(DN)$, frame averaged, and standard deviation are shown in the PTC representation of figure 5.10.

The magenta dotted line represents the linear regression for the first regime, slope=0, which is the read noise component:

$$\log_{10}(\sigma_{read}) = 0.0158 \cdot \log_{10}[S(DN)] - 0.7113 \quad (5.12)$$

The y-intercept defines a read noise of $10^{-0.7113} = 0.1944$ DN. The red dash-dotted line represents the linear regression for the second regime, slope=1/2, which is the photon shot noise component:

$$\log_{10}(\sigma_{shot}) = 0.5215 \cdot \log_{10}[S(DN)] - 1.1 \quad (5.13)$$

The factor of conversion (K_{ADC}) is calculated from the x-intercept of the photon shot component, resulting in $K_{ADC}=128.7$ (e^-/DN). The green dashed line represents the linear regression for the third regime, slope=1, which is the FPN component:

$$\log_{10}(\sigma_{FPN}) = 1.0327 \cdot \log_{10}[S(DN)] - 1.7976 \quad (5.14)$$

From the inverse of the x-intercept, the quality factor of the FPN results in 0.018 (1.8%).

Finally, the dynamic range (operating as a linear imager) is calculated, from the boundary of the

5. EXPERIMENTAL RESULTS

full-well regime (slope=0), as the full-well capacity S_{FW} divided by the read noise. We first need to extract the Full Well value. This is simply obtained from the DN which defines the transition between FPN-dominated region and FW region (95DN in our case). Now, it is expressed in e- units, simply by using the conversion factor K_{ADC} :

$$S_{FW}(e^-) = 95DN \cdot 128.7(e^-/DN) = 12226e^- \approx 12ke^- \quad (5.15)$$

$$\sigma_{READ}(e^-) = 0.1944DN \cdot 128.7(e^-/DN) = 25e^- \quad (5.16)$$

$$DR_{PTC}(dB) = 20 \cdot \log_{10} \left(\frac{S_{FW}}{\sigma_{READ}} \right) = 20 \cdot \log_{10} \left(\frac{12226}{25} \right) = 20 \cdot \log_{10}(489) = 53.8dB \quad (5.17)$$

It must be notice that, as the PTC method is aimed for linear image sensors, the DR defined by this method is the ratio between the saturation point and the read noise floor. In our case, it indicates the maximum SNR, while the DR in linear mode is rather $20 \cdot \log_{10}(2^7) = 42.1dB$, as the in-pixel ADC has 7-bits of resolution and the read noise floor is below 1LSB.

5.3.6. Shading Effects due to Heterogeneous Pixels Layout

The heterogeneous arrangement of 2x2 pixels in the chip has led to undesired effect. The asymmetry of the disposition of metal lines around the pixels on the left and right sides of the 2x2 basic structures produces different metal shading effects at the two sides.

Indeed photodiodes are physically separated horizontally by the SRAM modules on one side, and by the PMOS transistors of the amplifiers with the digital control between them on the other side. All data bus lines for the 2x2 arrangement lay on top of both these sides. These lines produce shading effects for non perpendicularly incidents light beams, particularly arriving in the horizontal axis because vertical axis does not contain Metal 4. An illustration of the metals disposition related to the photodiode in horizontal axis (x-axis) in the left top pixel of the 2x2 group is illustrated in figure 5.11. The angle which creates a shadow in photodiode and surroundings is different if it is compared left to right. The metal distribution of the right side pixel is similar but flipped left to right. Therefore, there will be an angle, incident left or right, where one side will be in shadow while the other side still receives some light. This causes the shading effect, which is very noticeable in the horizontal axis due to metal 4 that is not present in the vertical axis of the pixels.

Figure 5.12 shows this effect on real measurements on the chip when it is illuminated with either a class-1 laser or a LED in different incident angles. It is noticeable that the effect disappears when the light is coming vertically. It proves that, the distortion in the image is caused by the high amount of metal vertical lines that connect the pixel memories to the sense amplifier. A zoom over these metal lines is shown in the micrograph displayed in figure 5.13, where it is noticeable the difference between

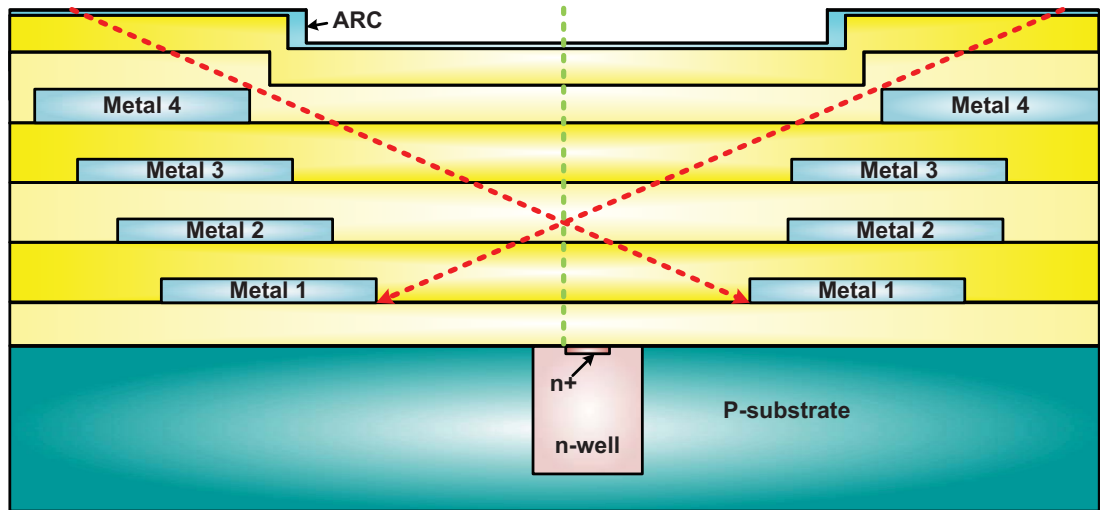


Figure 5.11: Illustration of nearest metals distribution in horizontal axis.

metal lines 2 x 2 (vertical bright and dull yellow lines) ¹.

This effect can be almost eliminated by the use of a 5th metal on top layer covering the entire pixel but the photodiode, as it was made in the SCU chip presented in chapter 2. However, this will reduce the effective sensitivity of the sensor, by degrading its optical efficiency [160]. Moreover, the technology

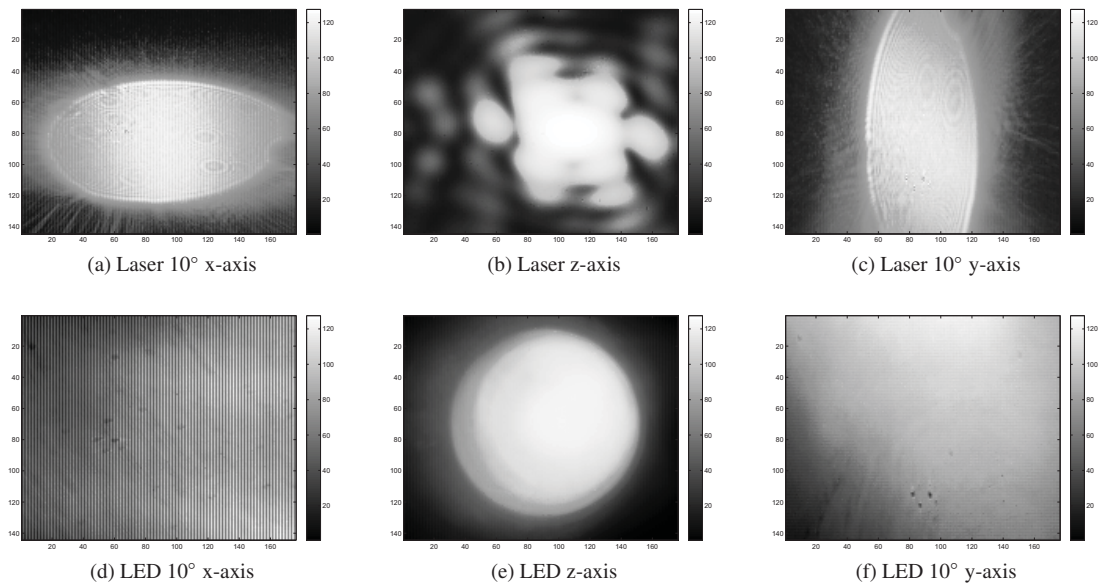


Figure 5.12: Metal shadows image capture.

¹The mark Divisa-2010 is at the top of the chip.

5. EXPERIMENTAL RESULTS

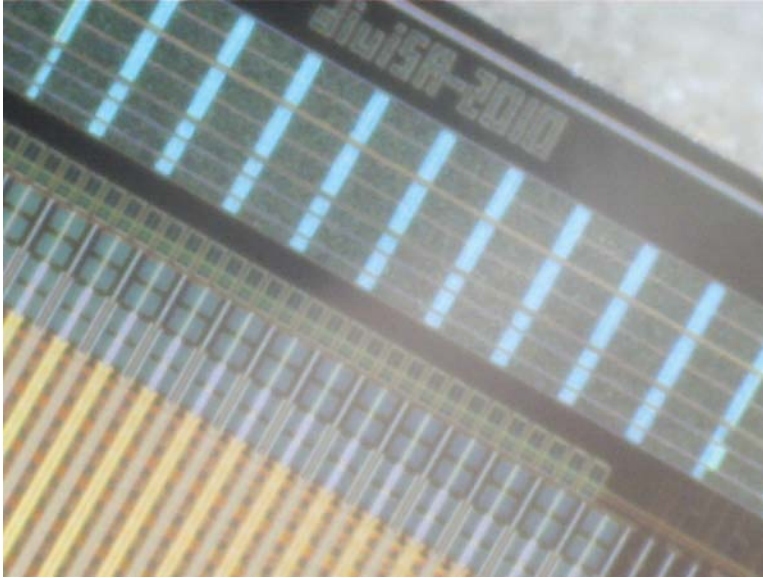


Figure 5.13: Pixel vertical lines zoom micrograph.

used for the TVHC chip has only 4 metals available and the 4th metal had to be used inside the pixel for the routing of necessary signals, so it was not possible to use it for shielding purposes.

Yet another possibility is the use of backside illumination for the pixels, since the metal connections will be underneath the photodiodes, so they will not create different shadows in heterogeneous array of pixels. This solution is aim to be used in the future development of this chip, in a 3D technology. Therefore, it is mostly guaranteed that this effect will not be present.

5.4. Captured Tonemapped Images

In this section, a set of captured images of several HDR scenes are presented. These HDR scenes have been captured using the 9 tone mapping modes previously presented. Besides, a comparison with 3 commercial cameras is also presented to illustrate how the sensor features both HDR and low noise simultaneously. The commercial systems included in the comparison are:

- Sony Cybershot DSC-W80: which uses a CCD array of sensors with enhanced sensitivity (*Super HADTM CCD*) [161].
- iPhone 4: that allows HDR mode by using a combination of 3 pictures (since iOS 4.2) [140].
- Photonfocus Linlog MV-D752E-40-U2-12: which embeds the LINLOGTM technology that uses a linear response at low illumination levels and logarithmic compression at high intensities [162].

A typically used high dynamic range scene created in laboratory is a direct photograph of a light bulb and a set of objects not directly illuminated by it. Standard dynamic range cameras are generally unable

to capture simultaneously the filament or details in the light bulb and the surrounding objects in this kind of scenes. The presented captured images include this scene using lamps of tungsten (figure 5.14), halogen (figure 5.15), fluorescent (figure 5.16) and LED type (figure 5.17), which have different spectral emissivity and emitting surfaces. It has also been included scenes of a very bright ceiling lamp (figure 5.18) and a photograph through a window with very bright natural light outside (figure 5.19). The zones of high details in difficult areas are marked by a blue shape and high loss of details are marked by red shapes. These scenes have been captured with 16 bins, whose subdivisions for evaluation durations are 20ns, 100ns, 200ns, 500ns, 700ns, 1.1 μ s, 1.5 μ s, 2.5 μ s, 4 μ s, 7 μ s, 12 μ s, 21 μ s, 38 μ s, 70 μ s, 120 μ s, ordered in order of occurrence, and finally 10 μ s for the ramp bin. Considering the delay between the states of the FSM of the FPGA and multiplying by 128 evaluations, it gives that the duration of the bins are 38.4 μ s, 48.6 μ s, 61.4 μ s, 99.8 μ s, 125.4 μ s, 176.6 μ s, 227.8 μ s, 355.8 μ s, 547.8 μ s, 931.8 μ s, 1.57ms, 2.72ms, 4.9ms, 8.99ms, 15.4ms and 1.52ms, respectively. This configuration allows a maximum exposition time of 37.72ms.

A set of large characters with decreasing amount of ink has been included in the scenes, as surrounding details. It is not an accurate contrast chart, because large noticeable details where necessary but they give a fast idea of the areas of loss of details due to the low resolution.

It is very noticeable that, first, regarding the images captured by TVHC chip, that despite of being captured with the same configuration, the mode used in the assignation of levels (see subsection 3.4.3) has a large influence in the performance. Second, despite using half the codes for image representation (128 vs. 256), the TVHC chip produces images which are “visually” competitive with the commercial approaches.

In general, it is noticeable that the Cybershot produces images with low noise but it shows both over and underexposed areas in all cases. The iPhone 4 exhibits high noise and the use of only 3 exposures in combination is not enough for very large dynamic range. The combination of images makes a noticeable improvement in the ceiling lamp area and natural light scenes, but it also produces some loss of details in certain areas in other scenes. Regarding the Linlog camera, it also generates an image with a high noise level, comparable with those produced by the iPhone. Nevertheless, it outperforms our approach in the surrounding of the lamps filaments. However, the increment of contrast in this area is at the cost of a loss of details in the low contrast areas (ceiling and low contrast characters).

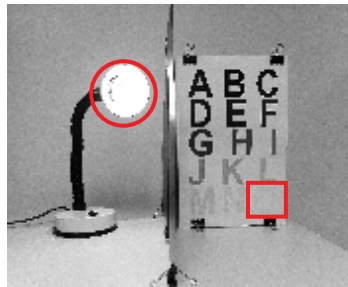
Regarding the performance of the TVHC chip, in general, images exhibit acceptable noise. However, in the surroundings of the light sources, the shading caused by vertical lines is noticeable. Concerning the modes of operation, mode 1 exhibits very bad results in all the scenes because it gives a fixed amount of levels to all the bins without taking care of the population in each bin. However, this mode is only used in the first capture when probability information does not exist and the objective is capturing TSI, not TMI.

In mode 2, where levels are statistically distributed by its contribution to the histogram, it shows very high contrast in the low light areas but it loses details in the surrounding of sources of light that have small area in the image (small amount of pixels \rightarrow small weight). Regarding mode 3, it gives a good performance in the sources of light but it loses details in the dark areas of the image. Mode 4 has

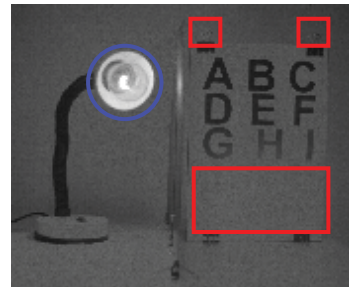
5. EXPERIMENTAL RESULTS



(a) Cybershot



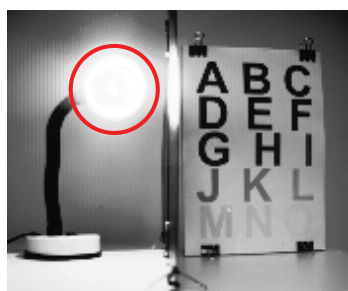
(b) iPhone 4



(c) Photonfocus Linlog



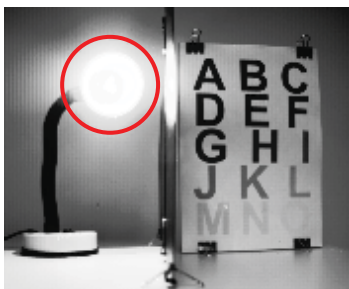
(d) TVHC Mode 1



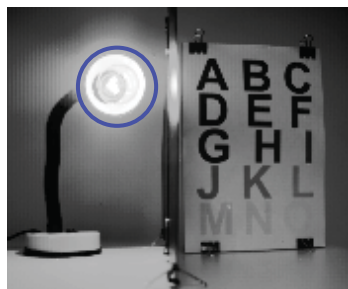
(e) TVHC Mode 2



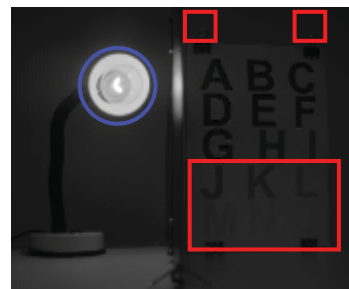
(f) TVHC Mode 3



(g) TVHC Mode 4



(h) TVHC Mode 5



(i) TVHC Mode 6



(j) TVHC Mode 7

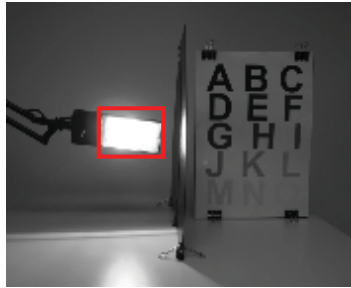


(k) TVHC Mode 8

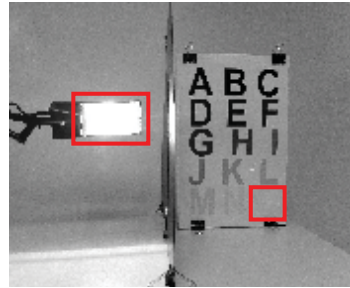


(l) TVHC Mode 9

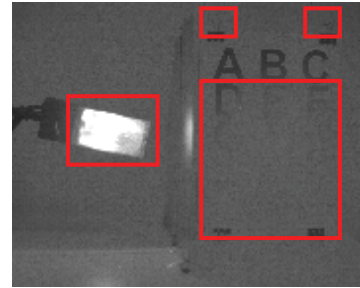
Figure 5.14: Tungsten lamp.



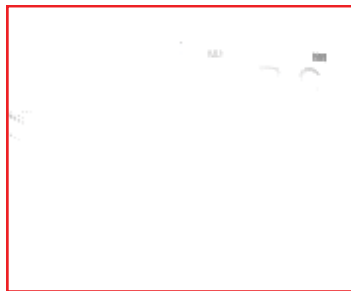
(a) Cybershot



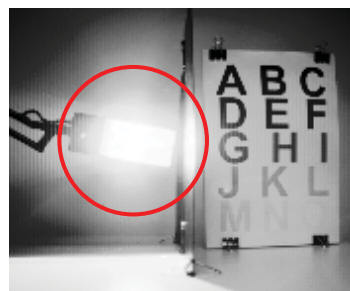
(b) iPhone 4



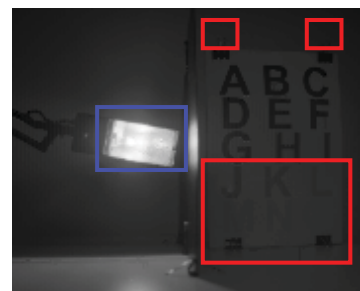
(c) Photonfocus Linlog



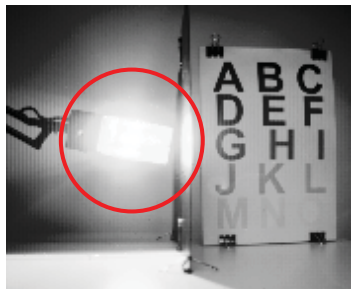
(d) TVHC Mode 1



(e) TVHC Mode 2



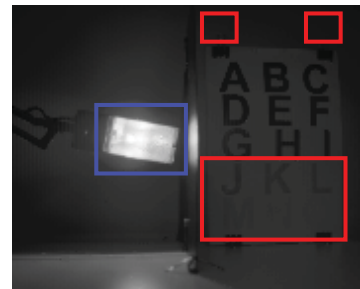
(f) TVHC Mode 3



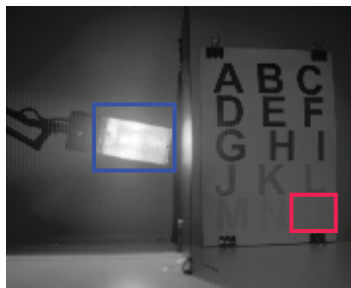
(g) TVHC Mode 4



(h) TVHC Mode 5



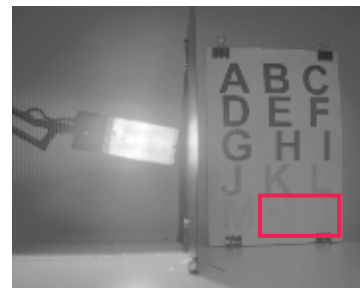
(i) TVHC Mode 6



(j) TVHC Mode 7



(k) TVHC Mode 8



(l) TVHC Mode 9

Figure 5.15: Halogen lamp.

5. EXPERIMENTAL RESULTS

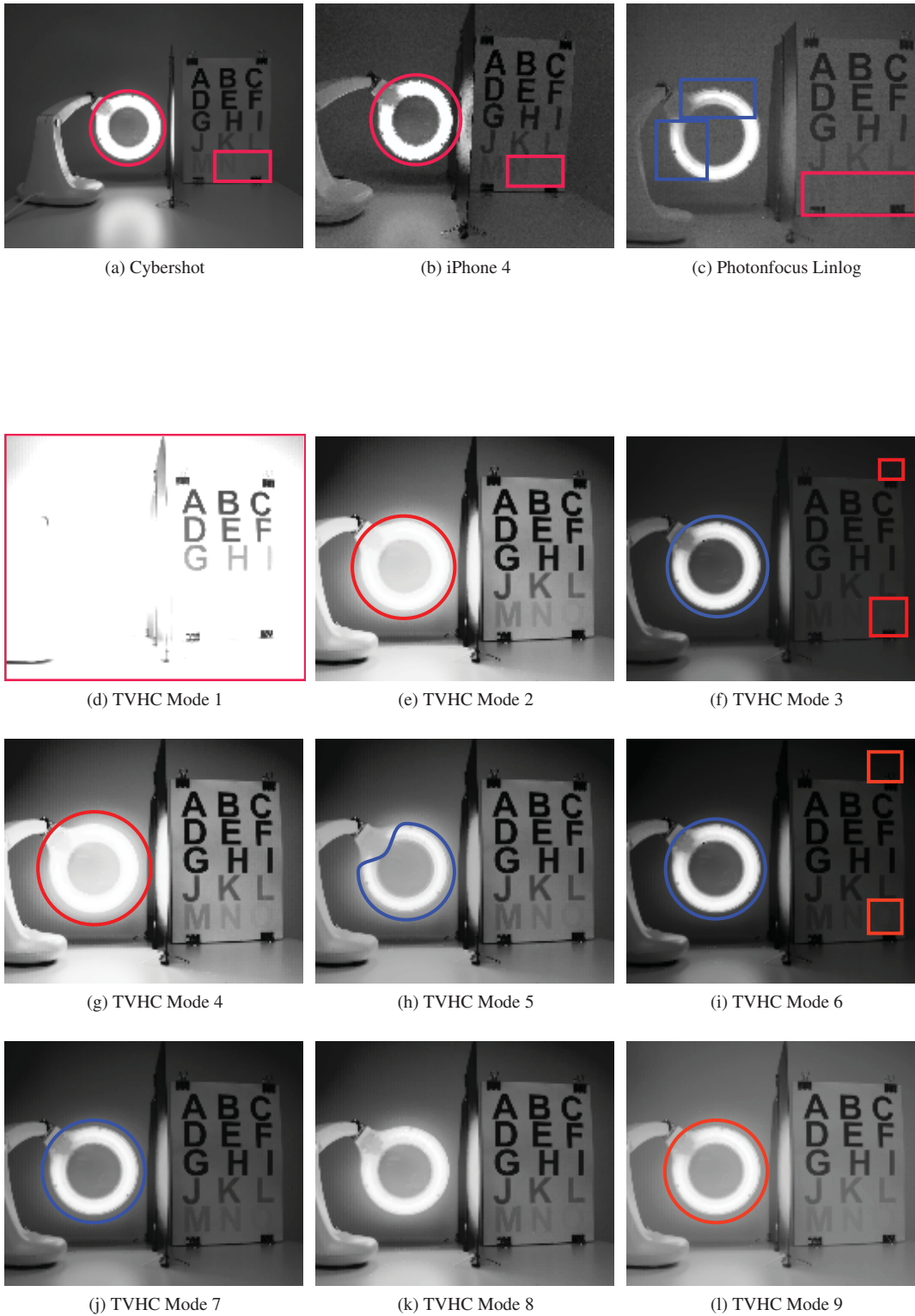
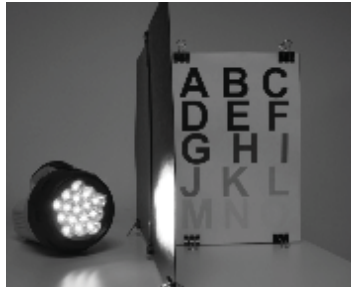
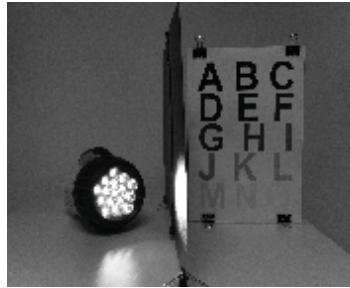


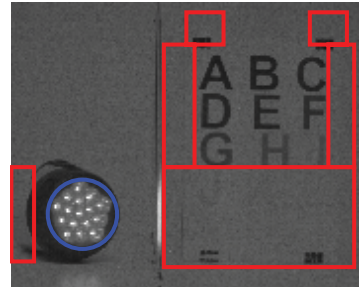
Figure 5.16: Fluorescent lamp.



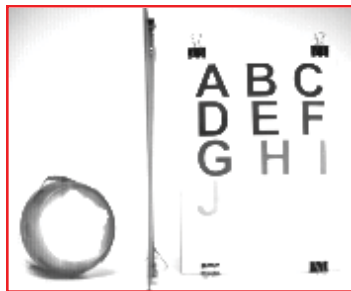
(a) Cybershot



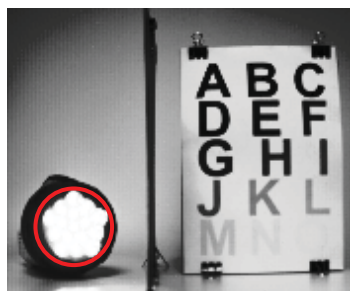
(b) iPhone 4



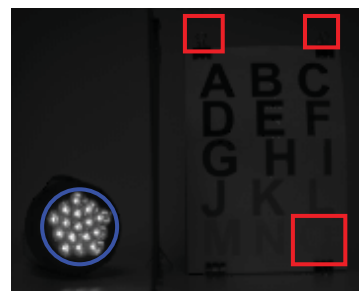
(c) Photonfocus Linlog



(d) TVHC Mode 1



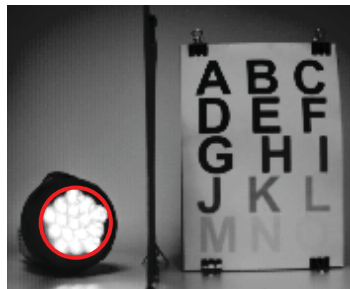
(e) TVHC Mode 2



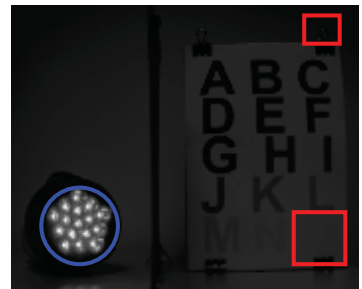
(f) TVHC Mode 3



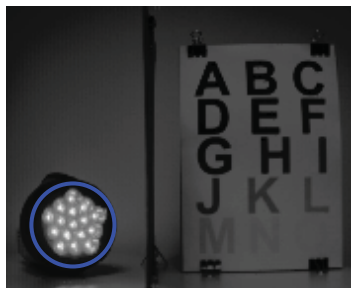
(g) TVHC Mode 4



(h) TVHC Mode 5



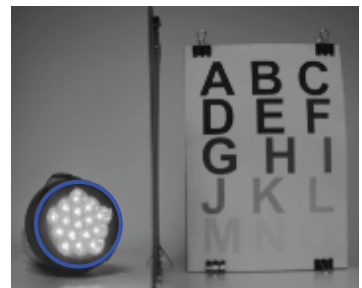
(i) TVHC Mode 6



(j) TVHC Mode 7



(k) TVHC Mode 8



(l) TVHC Mode 9

Figure 5.17: Led torch.

5. EXPERIMENTAL RESULTS

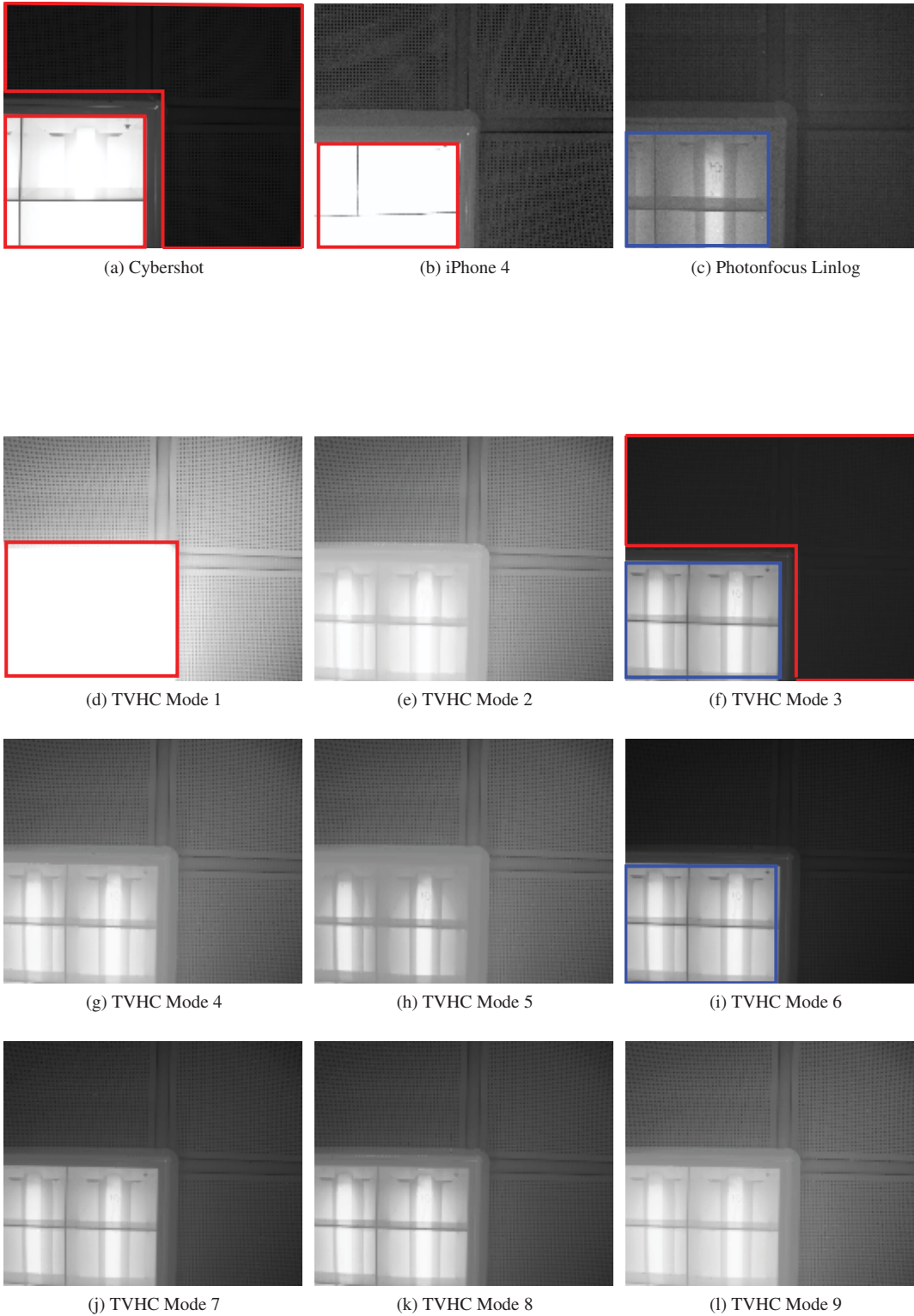


Figure 5.18: Ceiling Lamp.

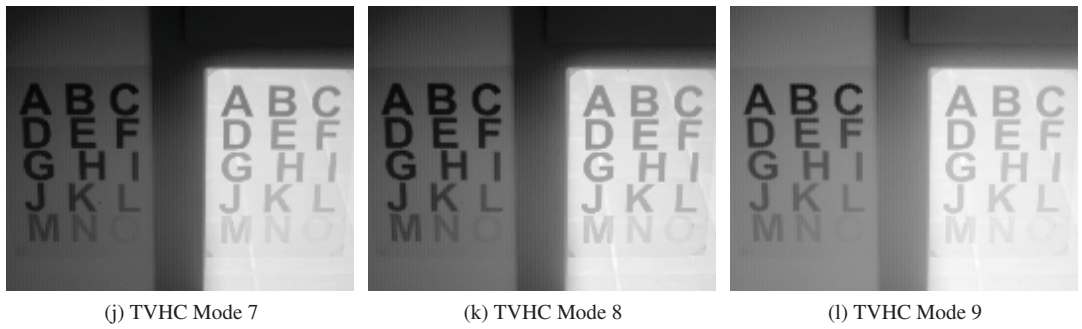
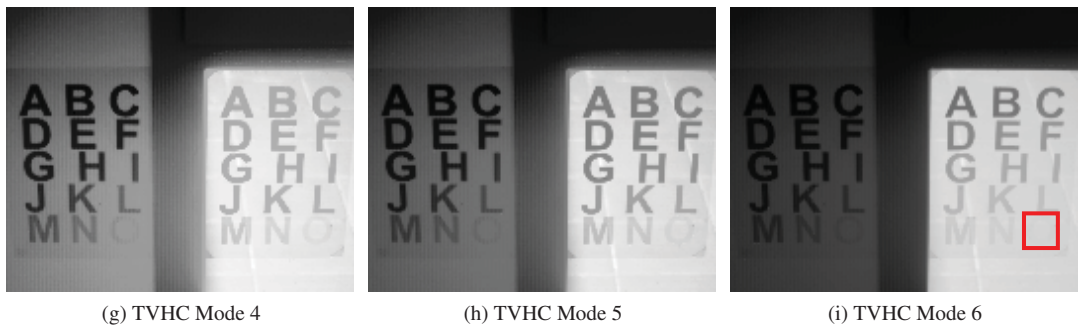
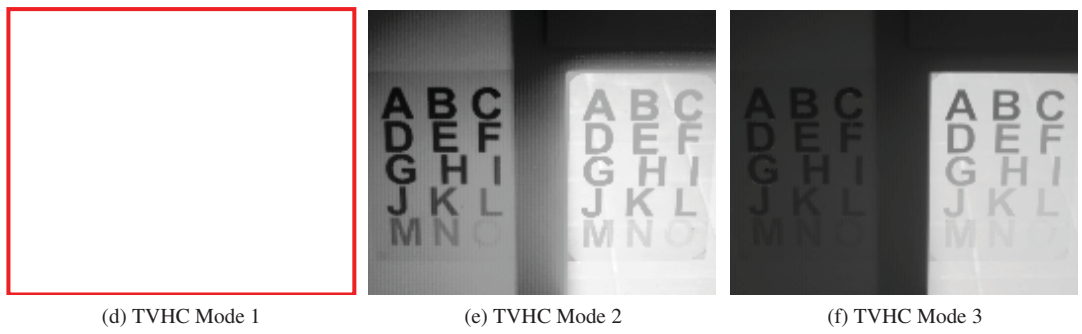
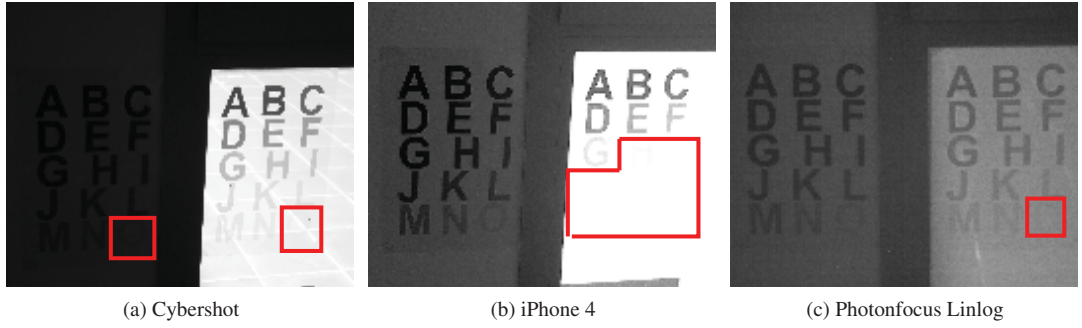


Figure 5.19: Natural light through window.

5. EXPERIMENTAL RESULTS

similar effect as mode 2, but it improves the behavior in the ceiling lamp scene because the distribution to avoid concentration has given levels to this area. Mode 5 improves the performance in the source of light, but it introduces rare halo effects in its surroundings because it has high details in the high light zones and the high gradient present in the area will be noticeable. Mode 6 behaves similar to mode 3 but improves the ceiling in the ceiling lamp scene. Mode 7 is the best mode in performance giving high details in the sources of light as well as in the dark surroundings. Mode 8 loses details in the sources of light in comparison to mode 7. Mode 9 in comparison to mode 7 loses contrast in the image and produces a generally bright image.

It must be notice that these considerations, as explained in chapter 3, are the result of the evaluation by the human vision system from the representation in typical LDR tone reproduction methods (such as printed paper or LCD monitors). This method of evaluation is typically used for the output of electronic image sensors. Besides, our produced tone mapped representation of HDR scenes have a lower dynamic range (7-bit=42.1dB) than LDR standard images (8-bit=48.1dB), and so the DR of reproduction devices will be sufficient in most of the cases ¹. However, in our case, if the output images are not intended for human visualization but for post-processing, the best HDR mode must be chosen by the user depending on the final application and the operations to be performed over the tone mapped images.

Due to the use of several sources of light, it have been proven that the tone mapping algorithm works properly on stable light (tungsten, sun, LED) and pulsing light (fluorescent). The system does not create gradients, which is typical in the presence of pulsing light using rolling shutter. Therefore, as the system acquires similar illumination in similar timings the consequent distortion is not present. The system behaves as having a pseudo global shutter for similar illuminations.

5.5. Summary of TVHC Characteristics

Several measurements have been performed to the TVHC chip. Table 5.1 summarize the main characteristics of the chip including these measurements.

5.6. Conclusions

The measurements indicate a response of $0.86 \frac{V}{lux-s}$ @550nm with a dark discharge of $10.8 \frac{mV}{s}$, which can be efficiently calibrated. The measurable dynamic range, which can be combined in a single frame, is very high. The hardware system is capable of capture 168dB scenes considering a noise floor of SNR=1, this will lead to a linear representation in data of about 28 bits, and 25 bits for noise SNR=10. Here, the improvement of the amount of data in the representation is extensive 7 bits vs. 25 bits. Therefore, this represents that our algorithm uses about a quarter of bits to represent the same scene, with the subsequent decrement in the digital processing circuitry required to analyze this scene.

¹However, they could produce a not linear perception, as the curve of reproduction of the device must be adapted to the curve of standard human perception. So, it usually produces worst perception in brightest and darkest zones.

Technology	AMS 0.35 μm 4M/2P OPTO Technology
Package	84 pins Ceramic Lead "J" Chip Carrier (JLCC84)
Array of Pixels	180(H)x148(V) (QCIF+ 4 dummies)
Pixel Size	33x33 μm^2
Photodiode	3x3 μm^2 NW-Sub
Pixels Metal Opening	9.75x7.3 μm^2
Fill Factor	0.8% (Photodiode) 6.5% (Metal Opening)
Exposition time range	2.34 μs to 8s
Image Coding	7-bit gray code
Spectral Response@555nm	0.86 $\frac{\text{V}}{\mu\text{V}\cdot\text{s}}$
Sensitivity	5.79 $\frac{\text{V}}{\mu\text{V}\cdot\text{s}}$
Dark Discharge	10.8 $\frac{\text{mV}}{\text{s}}$
Full Well Capacity	12.2ke ⁻
Conversion Factor	129(e ⁻ /DN)
PRNU (PTC FPN Quality Factor)	0.018 (1.8%)
Read Noise	25e ⁻
Linear SNR (PTC DR)	53.8dB
Linear DR (7-bit)	42.1dB
Tone Mapping DR (7-bit SNR1)	168dB
Tone Mapping DR (7-bit SNR10)	148dB
Fastest Image Download	666 μs
Maximum Frame Rate	1205.4fps
Static Power Consumption	4.5mW
Acquisition Maximum Power Consumption	562mW
Power supply	3.3V
Die Size	7330x6780 μm^2

Table 5.1: TVHC Chip Main Characteristics

The presented examples of tone mapped scenes show that the system is capable of capturing the details of the objects in HDR scenes in a single frame. Moreover, the flexibility in the assignation of levels per bin imply several working modes which can be chosen depending on the application environment. The system is capable of working in several types of illumination: stable and pulsing light.

The disadvantages of the system are the shading caused by the vertical metal lines. However, the overall performance of the chip is very satisfactory which opens a door for future approaches, such as the inclusion of this chip as the first layer of complex vision systems on chip in the emerging 3D technologies.

5. EXPERIMENTAL RESULTS

Chapter 6

Conclusions and Future Work

6.1. Conclusions

The objective of this doctoral work was to develop new architectures, algorithms, and circuit design techniques, to improve the dynamic range performance of the sensory part in Focal Plane Processors. The work presented in this thesis constitutes an advance in the state-of-the-art of HDR image sensors. It proposes an innovative solution to enhance the Dynamic Range in CMOS image sensors by introducing an on-the-fly tone mapping technique. Furthermore, this technique does not only greatly increase the DR, but also optimizes the final representation of the image in a reduced amount of bits (seven) per pixel. This is accomplished by generating information about the probability of impinging illuminations in the previous frame, and using this information to tune how the current scene is compressed and stored.

This thesis has been structured in three main work packages, namely: (1) exploration of dedicated CIS technologies, (2) design of algorithms for Tone Mapping based High Dynamic Range sensing, and (3) design of a QCIF HDR sensor, in a standard CMOS technology, which serves as a demonstrator of what can be achieved with the developed techniques. The main conclusions from each of these research lines are summarized here:

Regarding the work in CIS technology:

- The use of advanced CIS technologies for Enhanced Dynamic Range sensors reports significant advantages in terms of reduced dark current (29 factor) and increased sensitivity (10 factor) when compared to the equivalent sensor in a standard CMOS mainstream line. Besides, the possibility of adding microlenses on top of the sensors offers further sensitivity improvement (maximum 80% enhancement for low pitch pixels) and improved crosstalk, with no impact on attainable pitch or electrical noise at the sensor level.
- Technological limitations, existing at the time of designing our test chip in a CIS line (and still existing nowadays), excluded the incorporation of PMOS transistors and substrate contacts in the sensor area, thus hardly limiting the amount of “intelligence” which can be incorporated at the pixel level.

6. CONCLUSIONS AND FUTURE WORK

- These limitations produce two important drawbacks for this alternative. First, it almost discards this type of technologies when designing a Focal Plane Processor (that by definition, must incorporate some processing structures at the pixel level), and second, it basically limits DR improvement options to those lowering the noise floor, and in the best of the cases, to the use of multiple sampling techniques (where final images are recovered from the many samples outside the array of pixels). This, again, is not compatible with the FPP approach, since information from the pixels is the starting point of the image processing algorithm and a significant part of it is implemented by on-pixel circuitry.

Regarding the new Tone-Mapping algorithm:

- We have developed a new hardware-aided Tone Mapping algorithm for Dynamic Range improvement which combines the two typical paradigms for acquisition of photogenerated currents. For most of the exposure, we codify the visual information, in a piece-wise linear approach, as the time that takes the voltage level in a photocurrent integration pixel to cross a fixed reference. At the very end of the exposure, we measure the pixel voltage for those pixels not crossing the reference value previously. Commonly, only one of these paradigms are used, however the combination is an innovation inside the algorithm.
- The algorithm generates the compressive tone mapping curve from the histogram of an auxiliary image that is employed as a global descriptor of the distribution of illuminances in the current frame, and requires very little computational resources at the pixel level.
- Though not originally intended to optimize the visualization of HDR scenes in common LDR displays, it can be used to this purpose as well, since the implemented compression is higher at high illuminations, something which is consistent with the human visual system. The characteristics of the developed compression, global and monotonic, produces visually appropriate representations while avoiding the creation of visual artifacts.
- Besides, the algorithm implements a pseudo-equalization of the scene details in the final image, as it avoids gaps in the use of digital output codes when it is properly configured.
- The algorithm is fully compatible with the typical computational resources allocated at the pixel level in Focal Plane Processors, contributing to the state-of-the-art on Tone Mapping techniques, which are very rarely hardware-aided.

Regarding the demonstrator chip, which implements the designed tone mapping algorithm, it is a sort of *rara avis* in the field since, to the knowledge of the author, no other VSoC design with on-the-fly tone mapping in-pixel capabilities have been reported yet. Additionally, the design techniques used to make possible the on-chip implementation of the algorithm bring some innovations to the field of HDR image sensors, in details:

- We have designed a demonstrator in a 0.35 μm standard technology with only two add-ons to enhance light sensing capabilities: an anti-reflection coating and an optimized EPI substrate. This

reduces the cost of the chip in comparison with the use of a CIS technology, and does not introduce any limitation to the kind of circuitry that can be placed close to the sensor.

- The chip has proven that the algorithm works properly. It also confirms that the mathematical method, used for its simulation over HDR scenes, is appropriate to predict the functionality of the chip.
- Pixels include an auto-zeroing technique to minimize the Fixed Pattern Noise (FPN). The scheme does not require any additional storage capacitor or multiple readouts, as it is introduced in the effective photodiode reset voltage.
- We have designed a dynamic analog reference distribution scheme which takes advantage of the fast redistribution of charges in statistically zero averaged signals over large distributed RC loads. This allows for a precise communication of analog signals to an extremely large number of high impedance nodes.
- The chip contains a heterogeneous array of pixels, with the basic unit consisting in a 2x2 arrangement which allows for an easy implementation of the subsampling functionality.
- In-pixel SRAM storage of the final image allows for very long exposure frames, without distortion of the data caused by leakage or surrounding circuitry.
- The chip can be used as a regular linear acquisition imager; in that case, it produces 53.8dB SNR signals.
- The use of the algorithm provides up to 126dB, 168dB (Tone Mapped 7-bits) - 42dB (Linear 7-bits), increment of Dynamic Range when compared to the linear acquisition mode.
- HDR scenes of up to 168dB DR can be mapped by the chip using only 7-bits/pixel.
- In addition to Dynamic Range and SNR measurements, we have conducted a visual quality comparison of the images produced by the chip and those provided by three commercial systems. In our opinion, the chip outperforms all these commercial systems in visual quality since, although it exhibits a higher noise than a commercial CCD sensor, its HDR operation provides details which are simply lost in the scene provided by this commercial system. When compared to a dedicated HDR imager, though this imager provides slightly larger Dynamic Range, our solution offers much better contrast and very much lower noise levels.
- The chip is ideally well suited for future evolutions of the system using vertical stacking 3D-IC technologies.

6.2. Future Work

The satisfactory operation of the system just opens a door for further work oriented to enhance its functionality and optimize its operation. Among the many possibilities we could cite incorporation of

6. CONCLUSIONS AND FUTURE WORK

standardized I/O protocols, more on-chip processing, automatic selection of the tone mapping operation mode, smart power consumption management, etc. Clearly, the images provided by the chip exhibit undesired shadings effect due to the high density of vertical metals crossing over the pixel sensor array in order to download and store the images. This unwanted side effect can be solved by using an additional metal layer to create an identical metal opening for all the pixels. This, however, has some disadvantages since it will also reduce the amount of light reaching the silicon surface, making the use of a Back Side Illumination process (BSI), the natural solution to this problem.

Moreover, in order to increase fill factor, spatial resolution, and possibly to extend the operation to other wavelengths (UV, NIR or IR), a possible evolution of the system is its incorporation as part of a 3D integrated system. The idea is conceptually shown in figure 6.1, where the dies are connected by means of Through Silicon Vias (TSV).

The die on top (Tier 0) would contain the sensors (that might not necessarily be silicon-based), thus allowing for different sensing modalities. In the case of a vision sensor, this top tier would allow us to employ a Back Side Illuminated solution, and to improve sensing capabilities, to avoid shading, to improve fill factor to near 100% (some space will be necessary to place the TSV). Besides it would also permit a significant reduction of pixel size since other circuitry (switches, buffers, etc.) would be allocated in the “hidden” tiers. Thus, the die in the middle (for a 3-Tiers stack) could contain the rest of the circuitry of the pixel. This option would also allow using a technology enhanced for mixed signal circuitry in Tier 1 (or, additionally, a CIS technology at Tier 0). Finally, the die at the bottom (Tier 2) could contain the post-processing circuitry, standardized I/Os, DSP blocks, etc. Figure 6.2 shows a possible redistribution of the pixel circuitry among the different tiers.

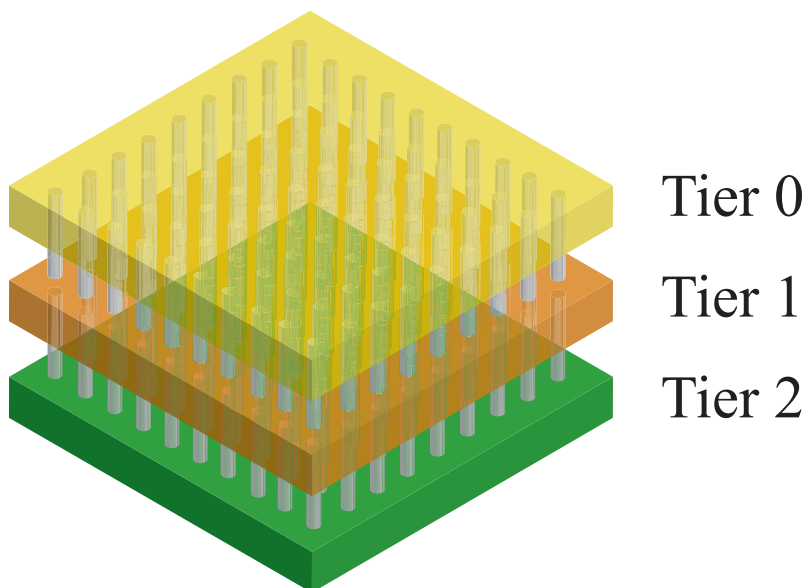


Figure 6.1: 3D-IC layers.

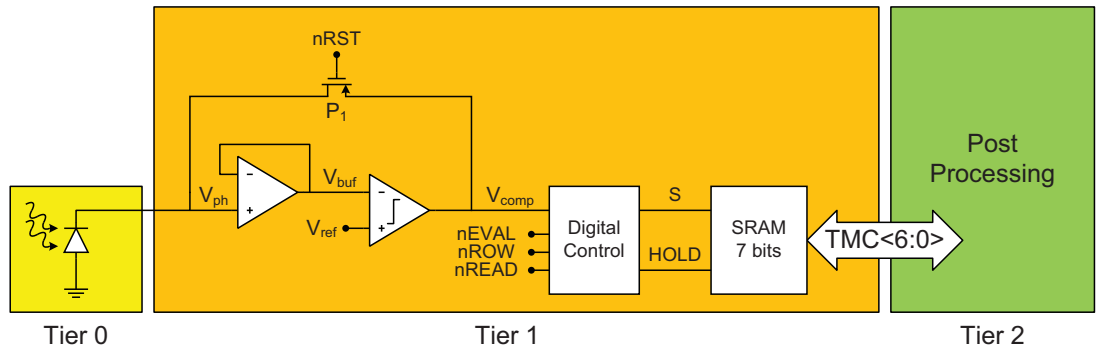


Figure 6.2: 3D-IC pixel schematic.

6. CONCLUSIONS AND FUTURE WORK

Appendix A

TVHC Time Requirements

The image capture operations will have associated time requirements, most of them due to the distributed RC load of the lines which are used for control and reference signals of the array of pixels.

The signals involved in the pixel memories erase operation are depicted in figure A.1. *ROW_CLR* activation will reset the position of the row pointer. *CLEAR_CODE_HIST* will set the signals of the Code Generator block to the maximum data and the signal *WRITE* will connect them to the pixel memories inputs. In order to write the data, the write switch *S* of the pixel must be activated by means of *nROW<147:0>*. The clock ticks in *ROW_CLK* increase the row pointer. The signal *ROW_ENABLE* activates the *S* switch of the row indicated by the row pointer.

Once the pixel memories have been erased, the integration node must be reset and the image capture can start. Figure A.2 shows the evolution of the signals during the reset and exposition time. First, the output voltage of the DAC must be changed from V_{top} to V_{rst} by means of the signal *DAC_MODE<1:0>*. Then the reset operation is performed by lowering the signal *nRST*. Now, the charge injection buffers are powered on and engaged by means of a low *NOVOFF* signal. The signal V_{ref} will decrease from V_{rst} to the value V_{bot} by means of the signal *DAC_MODE<1:0>*. Once $V_{ref}<147:0>$ is stabilized, the direct path to V_{dac} is established due to a high *NOVOFF*. During exposition time, clock ticks in the signals *CODE_CLK* and *HIST_CLK* will decrease the signals in the Code Generator outputs *TMC<6:0>* and *TSC<3:0>*, respectively. These signals will be written to the pixels if the integration node V_{ph} is

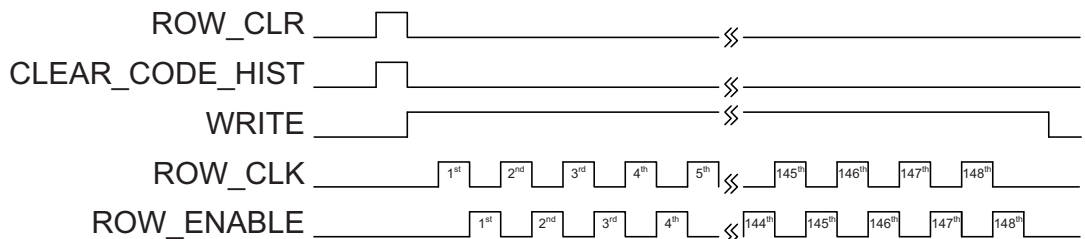


Figure A.1: Pixel memories erase operation.

A. TVHC TIME REQUIREMENTS

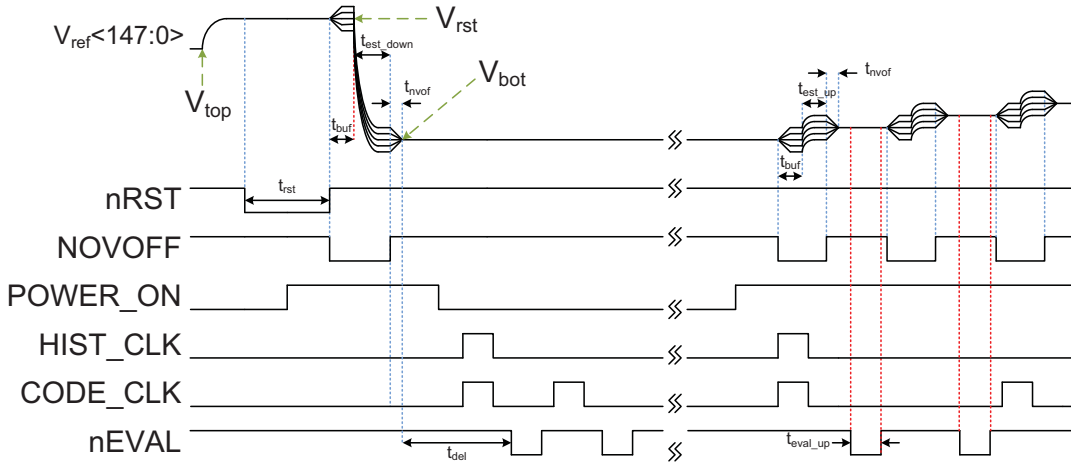


Figure A.2: Reset and exposition operation.

below V_{ref} when the signal $nEVAL$ is low. Therefore, in the ramp up $nEVAL$ can only be activated when V_{ref} has been properly established in all pixels.

When the image is already captured, it must be downloaded. Figure A.3 depicts signals to be applied. In order to read a row by the sense amplifiers block, first, the data lines must be precharged. The state machine which controls this operation should be reset at the start of the download of a frame by a high PCH_START . Then, in every row 10 clock ticks must be performed. Immediately after, the connection switch S of the row indicated by the row pointer must be activated by means of a high ROW_ENABLE . Then, the data is amplified and stored in a bus keeper by the signals $READ_C$ and $READ_H$. The row is then read by the Read Buffer block when the signal $READ_CLR$ is high. Then, every clock cycle in $READ_CLK$ will output 36-bits of this row to the pads. The minimum time requirements values indicated in the figures of this section are shown in the table A.1.

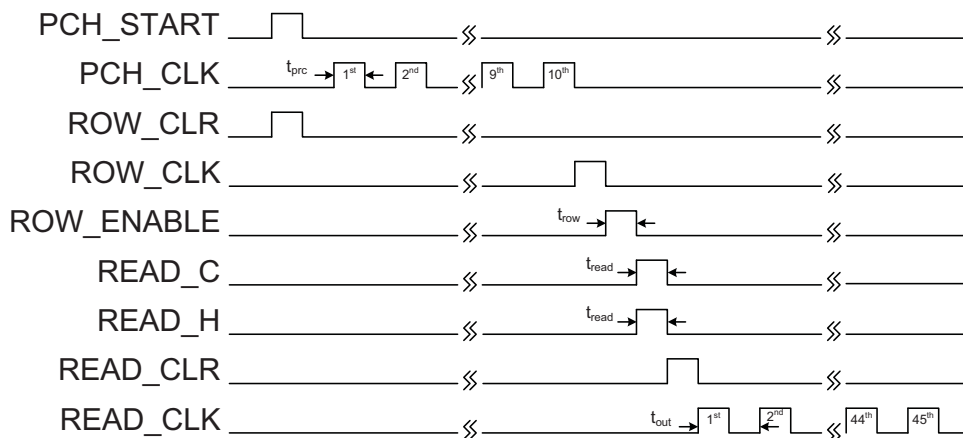


Figure A.3: Image download operation → 1 row.

Value	Signal Operation	Minimum Time Required
t_{rst}	Reset Phase	10 μ s
t_{buf}	Buffer Application time before V_{dac} change	250ns
t_{est_down}	Establishment time for ramp-down of V_{dac}	500ns
t_{est_up}	Establishment time for ramp-up of V_{dac}	250ns
t_{nvof}	Charge redistribution time	250ns
t_{del}	Delay time from NOVOFF application to image capture	0ns
t_{eval_up}	Ramp-up evaluation time for 3τ	1.2 μ s
t_{prc}	Precharge of data lines	250ns
t_{row}	Time for pixel memories to charge data lines	250ns
t_{read}	Time for sense amplification	100ns
t_{out}	Register to register specification of Read Buffer block	100ns

Table A.1: Control signals time requirements

A. TVHC TIME REQUIREMENTS

References

- [1] S.R. Morrison. A new type of photosensitive junction device. Solid-State Electronics, 6(5): 485 – 494, 1963. ISSN 0038-1101. doi: DOI:10.1016/0038-1101(63)90033-9. URL <http://www.sciencedirect.com/science/article/pii/0038110163900339>. 2
- [2] W.S. Boyle and G.E. Smith. Charge coupled semiconductor devices. Bell System Technical Journal, 49:587–593, 1970. 2
- [3] Robert H. Nixon, Sabrina E. Kemeny, Craig O. Staller, and Eric R. Fossum. 128 x 128 cmos photodiode-type active pixel sensor with on-chip timing, control, and signal chain electronics. volume 2415, pages 117–123. SPIE, 1995. doi: 10.1117/12.206529. URL <http://link.aip.org/link/?PSI/2415/117/1>. 2
- [4] Eric R. Fossum. Active pixel sensors: are ccds dinosaurs? volume 1900, pages 2–14. SPIE, 1993. doi: 10.1117/12.148585. URL <http://link.aip.org/link/?PSI/1900/2/1>. 5
- [5] Mats Wernersson and Henrik Eliasson. Sense and sensitivity. 2009 International Image Sensor Workshop, 2009. 6
- [6] Aptina Imaging Corporation. Whitepaper: An objective look at fsi and bsi, 2010. URL <http://www.aptna.com/news/FSI-BSI-WhitePaper.pdf>. 7
- [7] G. Agranov, S. Smith, R. Mauritzson, S. Chieh, U. Boettiger, X. Li, X. Fan, A. Dokoutchaev, B. Gravelle, H. Lee, W. Qian, and R. Johnson. Pixel continues to shrink....small pixels for novel cmos image sensors. 2011 International Image Sensor Workshop, pages 1–4, 2011. 7
- [8] T.H. Hsu, Y.K. Fang, C.Y. Lin, S.F. Chen, C.S. Lin, D.N. Yaung, S.G. Wu, H.C. Chien, C.H. Tseng, J.S. Lin, and C.S. Wang. Light guide for pixel crosstalk improvement in deep submicron cmos image sensor. Electron Device Letters, IEEE, 25(1):22 – 24, jan. 2004. ISSN 0741-3106. doi: 10.1109/LED.2003.821597. 8
- [9] G.A. Antcliffe, L.J. Hornbeck, W.W. Chan, J.W. Walker, W.C. Rhines, and D.R. Collins. A back-side illuminated 400 x 400 charge-coupled device imager. Electron Devices, IEEE Transactions on, 23(11):1225 – 1232, nov 1976. ISSN 0018-9383. doi: 10.1109/T-ED.1976.18583. 8

REFERENCES

- [10] A. Toumier, F. Leverd, L. Favennec, C. Perrot, L. Pinzelli, M. Gatefait, N. Cherault, D. Jeanjean, J-P. Carrere, F. Hirigoyen, L. Grant, and F. Roy. Pixel-to-pixel isolation by deep trench technology: Application to cmos image sensor. 2011 International Image Sensor Workshop, pages 12–15, 2011. 9
- [11] D. Choudhury. 3d integration technologies for emerging microsystems. In Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International, pages 1–4, may 2010. doi: 10.1109/MWSYM.2010.5514747. 9
- [12] Abbas Sheibanyrad, Frédéric Pétrot, and Axel Jantsch. 3D Integration for NoC-based SoC Architectures. Springer New York, 2011. 9
- [13] William J. Greig. Integrated Circuit Packaging, Assembly and Interconnections. Springer US, 2007. 9
- [14] P. Sun, V. Leung, D. Yang, and D. Shi. Development of a novel cost-effective package-on-package (pop) solution. In Electronic Packaging Technology High Density Packaging, 2009. ICEPT-HDP '09. International Conference on, pages 46 –51, aug. 2009. doi: 10.1109/ICEPT.2009.5270798. 9
- [15] J. Brunner, I. Wei Qin, and B. Chylak. Advanced wire bond looping technology for emerging packages. In Electronics Manufacturing Technology Symposium, 2004. IEEE/CPMT/SEMI 29th International, pages 85 – 90, 14-16, 2004. doi: 10.1109/IEMT.2004.1321637. 9
- [16] R. Ulrich and W. Brown. Advanced Electronic Packaging, chapter Electronic Package Assembly, pages 389 –436. Wiley-IEEE Press, 2006. 9
- [17] F.P. Carson, Young Cheol Kim, and In Sang Yoon. 3-d stacked package technology and trends. Proceedings of the IEEE, 97(1):31 –42, jan. 2009. ISSN 0018-9219. doi: 10.1109/JPROC.2008.2007460. 9
- [18] Adi Khakoni, David San Segundo Bello, Koen De Munck, Padmakumar Ramachandra Rao, Piet De Moor, and Georges Gielen. An integration time prediction based algorithm for wide dynamic range 3d-stacked image sensors. 2011 International Image Sensor Workshop, pages 130–133, 2011. 11
- [19] Jun Ohta. Smart CMOS Image Sensors and Applications. 2008. 11, 16
- [20] Junichi Nakamura. Image Sensors and Signal Processing for Digital Still Cameras. 2006. 11, 12, 14, 17, 18, 19, 25, 26, 122
- [21] Orly Yadid-Pecht and Ralph Etienne-Cummings. CMOS Imagers From Phototransduction to Image Processing. Springer US, 2004. 11
- [22] Horst K. Zimmermann. Integrated Silicon Optoelectronics. Springer Berlin / Heidelberg, 2010. 13

-
- [23] James R. Janesick. Photon Transfer. SPIE Press, Bellingham, WA, 2007. ISBN 9780819478382. doi: DOI:10.1117/3.725073. URL <http://dx.doi.org/10.1117/3.725073>. 14, 159
- [24] Chih-Hung Chen, Bigchoug Hung, Sheng-Yi Huang, Jin-Shyong Jan, V. Liang, and Chune-Sin Yeh. Thermal noise performance in recent cmos technologies. In Solid-State and Integrated-Circuit Technology, 2008. ICSICT 2008. 9th International Conference on, pages 476–479, oct. 2008. doi: 10.1109/ICSICT.2008.4734584. 15
- [25] Bin Luo, Lei Yan, and Fuxing Yang. Research of noise suppression for cmos image sensor. In Measuring Technology and Mechatronics Automation (ICMTMA), 2010 International Conference on, volume 2, pages 1100–1103, march 2010. doi: 10.1109/ICMTMA.2010.261. 15
- [26] Xinyang WANG. Noise in Sub-Micron CMOS Image Sensors. PhD thesis, Technische Universiteit Delft, 2008. 15, 16, 18, 19
- [27] M. Schöberl, S. Fössel, and A. Kaup. Fixed pattern noise column drift compensation (cdc) for digital moving picture cameras. In Image Processing (ICIP), 2010 17th IEEE International Conference on, pages 573–576, sept. 2010. doi: 10.1109/ICIP.2010.5652732. 17
- [28] Chang-Tsun Li and Yue Li. Digital camera identification using colour-decoupled photo response non-uniformity noise pattern. In Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, pages 3052–3055, 30 2010-june 2 2010. doi: 10.1109/ISCAS.2010.5537994. 17
- [29] T. Meyer, R.E. Johanson, and S. Kasap. Effect of 1/f noise in integrating sensors and detectors. Circuits, Devices Systems, IET, 5(3):177–188, may 2011. ISSN 1751-858X. doi: 10.1049/iet-cds.2010.0220. 17
- [30] Behzad Razavi. Design of Analog CMOS Integrated Circuits. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 2001. ISBN 0072380322, 9780072380323. 17, 109
- [31] J. Lai and A. Nathan. Reset and partition noise in active pixel image sensors. Electron Devices, IEEE Transactions on, 52(10):2329–2332, oct. 2005. ISSN 0018-9383. doi: 10.1109/TED.2005.856192. 18
- [32] V. Goiffon, P. Magnan, P. Martin-Gonthier, C. Virmontois, and M. Gaillardin. New source of random telegraph signal in cmos image sensors. 2011 International Image Sensor Workshop, 2011. 18
- [33] Deng Zhang, Hiroaki Ammo, Jegoon Ryu, Hirofumi Sumi, and Toshihiro Nishimura. A modeling and evaluation of the random telegraph signal noise on a cmos image sensor in motion pictures. 2009 International Image Sensor Workshop, 2009. 18
- [34] C. Leyris, J.C. Vildeuil, F. Roy, F. Martinez, M. Valenza, and A. Hoffmann. Response of correlated double sampling cmos imager circuit to random telegraph signal noise. In Devices, Circuits

REFERENCES

- and Systems, Proceedings of the 6th International Caribbean Conference on, pages 109–114, april 2006. doi: 10.1109/ICCDCS.2006.250845. 18
- [35] Yoshi Ohno. Handbook of Optoelectronics, chapter Basic concepts in photometry, radiometry and colorimetry, pages 287–305. Taylor & Francis, 2006. 19
- [36] CIE. Commission internationale de l'éclairage. URL <http://www.cie.co.at/>. 19
- [37] Shigeru Kawai. Handbook of Optical Interconnects, chapter 2, pages 27–64. 2005. 19, 46
- [38] H. Takahashi, T. Noda, T. Matsuda, T. Watanabe, M. Shinohara, T. Endo, S. Takimoto, R. Mishima, S. Nishimura, K. Sakurai, H. Yuzurihara, and S. Inoue. A 1/2.7-in 2.96 mpixel cmos image sensor with double cds architecture for full high-definition camcorders. Solid-State Circuits, IEEE Journal of, 42(12):2960–2967, dec. 2007. ISSN 0018-9200. doi: 10.1109/JSSC.2007.908719. 20
- [39] Takashi Komatsu and Takahiro Saito. Color image acquisition method using color filter arrays occupying overlapped color spaces. 5017(1):274–285, 2003. ISSN 0277786X. doi: DOI:10.1117/12.476743. 20
- [40] Rastislav Lukac and Konstantinos N. Plataniotis. Color Image Processing: Methods and Applications, chapter Single-Sensor Camera Image Processing, pages 379–383. CRC Press, 2007. 20
- [41] Hyungsuck Cho. Optomechatronics: Fusion of Optical and Mechatronic Engineering, chapter Fundamentals of Optics, pages 74–92. 2006. 21
- [42] Gregory Hallock Smith. Camera lenses: from box camera to digital. SPIE Press, 2006. 21
- [43] M.E. Bravo-Zanoguera, J. Rivera-Castillo, M. Vera-Perez, and M.A. Reyna Carranza. Use of the modulation transfer function to measure quality of digital cameras. In Electronics, Communications and Computers, 2006. CONIELECOMP 2006. 16th International Conference on, page 52, feb. 2006. doi: 10.1109/CONIELECOMP.2006.62. 22
- [44] Chia-Kai Liang, Li-Wen Chang, and H.H. Chen. Analysis and compensation of rolling shutter effect. Image Processing, IEEE Transactions on, 17(8):1323–1330, aug. 2008. ISSN 1057-7149. doi: 10.1109/TIP.2008.925384. 23
- [45] Omar Ait-Aider, Nicolas Andreff, Jean Lavest, and Philippe Martinet. Simultaneous object pose and velocity computation using a single view from a rolling shutter camera. In Computer Vision ECCV 2006, volume 3952 of Lecture Notes in Computer Science, pages 56–68. Springer Berlin / Heidelberg, 2006. 23
- [46] M. Wany and G.P. Israel. Cmos image sensor with nmos-only global shutter and enhanced responsivity. Electron Devices, IEEE Transactions on, 50(1):57–62, jan 2003. ISSN 0018-9383. doi: 10.1109/TED.2002.807253. 24

- [47] S.-W. Han and E. Yoon. Area-efficient correlated double sampling scheme with single sampling capacitor for cmos image sensors. Electronics Letters, 42(6):335 – 337, march 2006. ISSN 0013-5194. doi: 10.1049/el:20064189. 24, 115
- [48] R.H. Nixon, S.E. Kemeny, B. Pain, C.O. Staller, and E.R. Fossum. 256 x 256 cmos active pixel sensor camera-on-a-chip. Solid-State Circuits, IEEE Journal of, 31(12):2046 –2050, dec 1996. ISSN 0018-9200. doi: 10.1109/4.545830. 24, 115
- [49] Yu-Chuan Shih and Chung-Yu Wu. An optimized cmos pseudo-active-pixel-sensor structure for low-dark-current imager applications. In Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on, volume 1, pages I–809 – I–812 vol.1, may 2003. doi: 10.1109/ISCAS.2003.1205687. 24
- [50] Yaodong Wang, T. Takaki, and I. Ishii. Intelligent high-frame-rate video recording with image-based trigger. In World Automation Congress (WAC), 2010, pages 1 –6, sept. 2010. 24
- [51] G. Agranov, V. Berezin, and R.H. Tsai. Crosstalk and microlens study in a color cmos image sensor. Electron Devices, IEEE Transactions on, 50(1):4 – 11, jan 2003. ISSN 0018-9383. doi: 10.1109/TED.2002.806473. 25
- [52] Brian A. Wandell. Foundations of Vision. Sinauer Associates, 1995. 26, 68
- [53] Bernd Hoefflinger. High-Dynamic-Range (HDR) Vision. Springer, 2007. 26
- [54] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht. Wide-dynamic-range cmos image sensors - comparative performance analysis. Electron Devices, IEEE Transactions on, 56(11):2446 –2461, nov. 2009. ISSN 0018-9383. doi: 10.1109/TED.2009.2030599. 27, 38
- [55] Liang-Wei Lai, Cheng-Hsiao Lai, and Ya-Chin King. A novel logarithmic response cmos image sensor with high output voltage swing and in-pixel fixed-pattern noise reduction. Sensors Journal, IEEE, 4(1):122 – 126, feb. 2004. ISSN 1530-437X. doi: 10.1109/JSEN.2003.820339. 28
- [56] B. Choubey and S. Collins. Models for pixels with wide-dynamic-range combined linear and logarithmic response. Sensors Journal, IEEE, 7(7):1066 –1072, july 2007. ISSN 1530-437X. doi: 10.1109/JSEN.2007.895959. 29
- [57] N. Akahane, S. Adachi, K. Mizobuchi, and S. Sugawa. Optimum design of conversion gain and full well capacity in cmos image sensor with lateral overflow integration capacitor. Electron Devices, IEEE Transactions on, 56(11):2429 –2435, nov. 2009. ISSN 0018-9383. doi: 10.1109/TED.2009.2030550. 29
- [58] N. Akahane, S. Sugawa, S. Adachi, K. Mori, T. Ishiuchi, and K. Mizobuchi. A sensitivity and linearity improvement of a 100 db dynamic range cmos image sensor using a lateral overflow integration capacitor. In VLSI Circuits, 2005. Digest of Technical Papers. 2005 Symposium on, pages 62 – 65, june 2005. doi: 10.1109/VLSIC.2005.1469334. 29

REFERENCES

- [59] D.G. Chen, D. Matolin, A. Bermak, and C. Posch. Pulse-modulation imaging - review and performance analysis. Biomedical Circuits and Systems, IEEE Transactions on, 5(1):64 –82, feb. 2011. ISSN 1932-4545. doi: 10.1109/TBCAS.2010.2075929. 30
- [60] M. Sasaki, M. Mase, S. Kawahito, and Y. Tadokoro. A wide dynamic range cmos image sensor with multiple short-time exposures. In Sensors, 2004. Proceedings of IEEE, pages 967 – 972 vol.2, oct. 2004. doi: 10.1109/ICSENS.2004.1426333. 31
- [61] O. Yadid-Pecht and A. Belenky. In-pixel autoexposure cmos aps. Solid-State Circuits, IEEE Journal of, 38(8):1425 – 1428, aug. 2003. ISSN 0018-9200. doi: 10.1109/JSSC.2003.811984. 32
- [62] EURO PRACTICE. Europractice ic service. URL <http://www.europractice-ic.com/>. 35
- [63] UMC. Umc cmos image sensor technology. URL <http://www.umc.com/english/process/m.asp>. 35, 36, 37
- [64] Wai-Kai Chen, editor. The VLSI Handbook, Second Edition, chapter CMOS Fabrication, pages 12–2512–27. CRC Press, 2007. 36
- [65] Koichi Mizobuchi, Satoru Adachi, Tomokazu Yamashita, Seiichiro Okamura, Hiromichi Oshikubo, Nana Akahane, and Shigetoshi Sugawa. A wide dynamic range cmos image sensor with resistance to high temperatures. 2007 International Image Sensor Workshop, pages 26–29, 2007. 37
- [66] Hui Tian, B. Fowler, and A.E. Gamal. Analysis of temporal noise in cmos photodiode active pixel sensor. Solid-State Circuits, IEEE Journal of, 36(1):92 –101, jan 2001. ISSN 0018-9200. doi: 10.1109/4.896233. 42
- [67] Bing Sheu, Je-Hurn Shieh, and M. Patil. Modeling charge injection in mos analog switches. Circuits and Systems, IEEE Transactions on, 34(2):214 – 216, feb 1987. ISSN 0098-4094. doi: 10.1109/TCS.1987.1086096. 42
- [68] I. Shcherback and O. Yadid-Pecht. Photoresponse analysis and pixel shape optimization for cmos active pixel sensors. 50(1):12–18, 2003. doi: 10.1109/TED.2002.806966. 43
- [69] Igor Shcherback, Alexander A. Belenky, and Orly Yadid-Pecht. Active-area shape influence on the dark current of cmos imagers. volume 4669, pages 117–124. SPIE, 2002. doi: 10.1117/12.463445. URL <http://link.aip.org/link/?PSI/4669/117/1>. 43
- [70] Franco Maloberti. Analog Design for CMOS VLSI Systems, chapter 4, pages 171–173. 2001. 47
- [71] Newport Corporation. Rpr-46-8 rpr relianceTM industrial and educational grade optical table. URL http://search.newport.com/?q=*&x2=sku&q2=RPR-46-8. 53

- [72] Ching-Chun Wang. A study of CMOS Technologies for Image Sensor Applications. PhD thesis, Massachusetts Institute of Technology, August 2001. 54
- [73] Betsaida Alexandre Barajas. Caracterización de sensores de imagen en una tecnología cmos de 0.18µm. Master's thesis, University of Seville, October 2007. 58, 59
- [74] P.M. Beaudoin, Y. Audet, and V.H. Ponce-Ponce. Dark current compensation in cmos image sensors using a differential pixel architecture. In Circuits and Systems and TAISA Conference, 2009. NEWCAS-TAISA'09. Joint IEEE North-East Workshop on, pages 1–4, 28 2009-july 1 2009. doi: 10.1109/NEWCAS.2009.5290457. 59
- [75] Findlay Shearer. Power Management in Mobile Devices, chapter 2, pages 67–68. 61
- [76] Jonathan Cohen, Chris Tchou, Tim Hawkins, and Paul Debevec. Real-time high dynamic range texture mapping. URL <http://gl.ict.usc.edu/Research/hdrtm/>. 67
- [77] Dee Unglaub Silverthorn. Human physiology : an integrated approach. 2007. 68
- [78] The Colour & Vision Research Laboratory. Webpage. URL <http://www.cvr1.org>. 69
- [79] Sumanta Pattanaik Paul Debevec Erik Reinhard, Greg Ward. High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting. Elsevier / Morgan Kaufmann, 2006. 70, 74, 92
- [80] Ernst Heinrich Weber. De pulsu, resorptione, audita et tactu, Annotationes anatomicae et physiologicae. Leipzig, 1834. 70
- [81] Gustav Theodor Fechner. Über ein wichtiges psychophysiches Grundgesetz und dessen Beziehung zur Schätzung der Sterngrößen. 1858. 70
- [82] S.S. Stevens. On the psychophysical law. Psychological Review, 64(3):153–181, 1957. ISSN 0033-295X. doi: DOI:10.1037/h0046162. URL <http://www.sciencedirect.com/science/article/B6X04-4NN6WD7-1/2/c2f430472bd6dbfdb7b13849e1988c28>. 71
- [83] G. S. Miller and C. R. Hoffman. Illumination and reflection maps: Simulated objects in simulated and real environments. Technical report, SIGGRAPH 84 Course Notes for Advanced Computer Graphics Animation, July 1984. 74
- [84] J. Tumblin and H. Rushmeier. Tone reproduction for realistic images. 13(6):42–48, 1993. doi: 10.1109/38.252554. 74
- [85] G. Ward. A contrast-based scalefactor for luminance display, chapter Graphics Gems IV, pages 415–421. Boston: Academic Press, 1994. 74
- [86] H. Richard Blackwell. Studies of psychophysical methods for measuring visual thresholds. Journal of the Optical Society of America, 42(9):606–614, Sep 1952. doi: 10.1364/JOSA.42.000606. URL <http://www.opticsinfobase.org/abstract.cfm?URI=josa-42-9-606>. 74

REFERENCES

- [87] James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg. A model of visual adaptation for realistic image synthesis. In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, SIGGRAPH '96, pages 249–258, New York, NY, USA, 1996. ACM. ISBN 0-89791-746-4. URL <http://doi.acm.org/10.1145/237170.237262>. 74
- [88] F. Drago, K. Myszkowski, T. Annen, and N. Chiba. Adaptive logarithmic mapping for displaying high contrast scenes. Computer Graphics Forum, 22:419–426, 2003. 75, 98
- [89] Jr. Stockham, T.G. Image processing in the context of a visual model. Proceedings of the IEEE, 60(7):828–842, 1972. ISSN 0018-9219. doi: 10.1109/PROC.1972.8782. 75
- [90] E. Reinhard and K. Devlin. Dynamic range reduction inspired by photoreceptor physiology. Visualization and Computer Graphics, IEEE Transactions on, 11(1):13–24, 2005. ISSN 1077-2626. doi: 10.1109/TVCG.2005.9. 75, 98
- [91] Donald C. Hood, Marcia A. Finkelstein, and Eugene Buckingham. Psychophysical tests of models of the response function. Vision Research, 19(4):401 – 406, 1979. ISSN 0042-6989. doi: DOI:10.1016/0042-6989(79)90104-4. URL <http://www.sciencedirect.com/science/article/B6T0W-4846GT7-BN/2/d86f728b55a5b681d4e60484fc51bda2>. Visual Sensitivity and Adaption, The British Photobiology Society and The Association for Research in Vision and Ophthalmology Inc. 75
- [92] J. von Kries. Sources of Color Science, chapter Chromatic Adaptation(1902), pages 120–126. 1970. 75
- [93] G.W. Larson, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. Visualization and Computer Graphics, IEEE Transactions on, 3(4): 291–306, 1997. ISSN 1077-2626. doi: 10.1109/2945.646233. 75
- [94] G. Sakas, P. Shirley, and S. Müller. Photoreslistic Rendering Techniques, chapter Quantization Techniques for High Dynamic Range Pictures. Christophe Schlick., pages 7–20. Springer-Verlag, Berlin, 1995. 75
- [95] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In Proceedings of Graphics Interface 93, pages 245–253, 1993. 76
- [96] Jobson Daniel J., Rahman Zia-ur, and Woodell Glenn A. Retinex image processing: Improved fidelity to direct visual observation. Technical report, 1996. 77
- [97] Edwin H. Land, John, and J. Mccann. Lightness and retinex theory. Journal of the Optical Society of America, pages 1–11, 1971. 77

- [98] Sumanta N. Pattanaik, Mark D. Fairchild, James A. Ferwerda, and Donald P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. pages 287–298, 1998. 77
- [99] Mark D. Fairchild. A revision of ciecam97s for practical applications. Color Research & Application, 26(6):418–427, 2001. ISSN 1520-6378. doi: 10.1002/col.1061. URL <http://dx.doi.org/10.1002/col.1061>. 77
- [100] Marc Ebner. Color Spaces, pages 87–101. John Wiley & Sons, Ltd, 2007. ISBN 9780470510490. doi: 10.1002/9780470510490.ch5. URL <http://dx.doi.org/10.1002/9780470510490.ch5>. 78
- [101] Michael Ashikhmin. A tone mapping algorithm for high contrast images. In Proceedings of the 13th Eurographics workshop on Rendering, EGRW '02, pages 145–156, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association. ISBN 1-58113-534-3. URL <http://portal.acm.org/citation.cfm?id=581896.581916>. 78, 98
- [102] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. In Proceedings of the 29th annual conference on Computer graphics and interactive techniques, SIGGRAPH '02, pages 267–276, New York, NY, USA, 2002. ACM. ISBN 1-58113-521-1. URL <http://doi.acm.org/10.1145/566570.566575>. 78, 98, 141
- [103] Ansel Adams. The Ansel Adams Photography series: The camera, The negative and The print. Little, Brown and Company. 78
- [104] A. V. Oppenheim, R. W. Schaffer, and Jr. Stockham, T. G. Nonlinear filtering of multiplied and convolved signals. 56(8):1264–1291, 1968. doi: 10.1109/PROC.1968.6570. 78
- [105] Frédo Durand and Julie Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. In Proceedings of the 29th annual conference on Computer graphics and interactive techniques, SIGGRAPH '02, pages 257–266, New York, NY, USA, 2002. ACM. ISBN 1-58113-521-1. URL <http://doi.acm.org/10.1145/566570.566574>. 78, 98
- [106] Prasun Choudhury and Jack Tumblin. The trilateral filter for high contrast images and meshes. In ACM SIGGRAPH 2005 Courses, SIGGRAPH '05, New York, NY, USA, 2005. ACM. URL <http://doi.acm.org/10.1145/1198555.1198565>. 78
- [107] Berthold K.P. Horn. Determining lightness from an image. Computer Graphics and Image Processing, 3(4):277 – 299, 1974. ISSN 0146-664X. doi: DOI:10.1016/0146-664X(74)90022-7. URL <http://www.sciencedirect.com/science/article/B7GXF-4S26XJT-1/2/8fe855c8112b8116f5ad6195143a19ab>. 79
- [108] Raanan Fattal, Dani Lischinski, and Michael Werman. Gradient domain high dynamic range compression. ACM Trans. Graph., 21:249–256, July 2002. ISSN 0730-0301. URL <http://doi.acm.org/10.1145/566654.566573>. 79, 98

REFERENCES

- [109] Mohammad T. Rahman, Nasser Kehtarnavaz, and Qolamreza R. Razlighi. Using image entropy maximum for auto exposure. 20(1):013007, 2011. ISSN 10179909. doi: DOI:10.1117/1.3534855. 91
- [110] Nikon. Nikon d90. URL <http://imaging.nikon.com/lineup/dslr/d90/>. 91
- [111] dpreview. Nikon d90 review, . URL <http://www.dpreview.com/reviews/nikond90/>. 91
- [112] DXOMark. Tests and reviews for the camera nikon d90. URL <http://www.dxomark.com/index.php/Cameras/Camera-Sensor-Database/Nikon/D90>. 91
- [113] Adobe Photoshop CS5 Help. Merge images to hdr. URL http://help.adobe.com/en_US/photoshop/cs/using/WSfd1234e1c4b69f30ea53e41001031ab64-78e5a.html. 92
- [114] Qtpfsgui. Luminance hdr. URL <http://qtpfsgui.sourceforge.net/>. 98
- [115] Open Source Photography. Parameters for tone mapping operators. URL <http://osp.wikidot.com/parameters-for-photographers>. 98
- [116] Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept.*, 3:286–308, July 2006. ISSN 1544-3558. URL <http://doi.acm.org/10.1145/1166087.1166095>. 98
- [117] Rafal Mantiuk, Scott Daly, and Louis Kerofsky. Display adaptive tone mapping. In *ACM SIGGRAPH 2008 papers, SIGGRAPH '08*, pages 68:1–68:10, New York, NY, USA, 2008. ACM. ISBN 978-1-4503-0112-1. URL <http://doi.acm.org/10.1145/1399504.1360667>. 98
- [118] Sumanta N. Pattanaik, Jack Tumblin, Hector Yee, and Donald P. Greenberg. Time-dependent visual adaptation for fast realistic image display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques, SIGGRAPH '00*, pages 47–54, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co. ISBN 1-58113-208-5. URL <http://dx.doi.org/10.1145/344779.344810>. 98
- [119] Karel Zuiderveld. *Contrast limited adaptive histogram equalization*, pages 474–485. Academic Press Professional, Inc., San Diego, CA, USA, 1994. ISBN 0-12-336155-9. URL <http://dl.acm.org/citation.cfm?id=180895.180940>. 98
- [120] Sunil Khatri and Narendra Shenoy. *EDA for IC Implementation, Circuit Design, and Process Technology*, chapter 2, pages 2–1 – 2–26. CRC Press, 2006. 106
- [121] H. Angus Macleod. *Thin-Film Optical Filters, Fourth Edition*, chapter Antireflection Coatings, page 105184. CRC Press, 2010. 106
- [122] Europractice IC. Datasheet: austriamicrosystems 0.35 mm cmos opto (c35b4o1), . URL http://www.europractice-ic.com/docs/austria_datasheets/035umcmos_OPTO35_v2.pdf. 106

- [123] Europractice IC. Datasheet: austriamicrosystems 0.35 mm cmos (c35), . URL http://www.europractice-ic.com/docs/austria_datasheets/035umCMOS_C35_2.pdf. 106
- [124] V. Gupta and M. Anis. Statistical design of the 6t sram bit cell. Circuits and Systems I: Regular Papers, IEEE Transactions on, 57(1):93 –104, jan. 2010. ISSN 1549-8328. doi: 10.1109/TCSI.2009.2016633. 109
- [125] Q.A. Khan, S.K. Wadhwa, and K. Misri. Low power startup circuits for voltage and current reference with zero steady state current. In Low Power Electronics and Design, 2003. ISLPED '03. Proceedings of the 2003 International Symposium on, pages 184 – 188, aug. 2003. doi: 10.1109/LPE.2003.1231859. 109
- [126] G.A. Fahmy, R.K. Pokharel, H. Kanaya, and K. Yoshida. A 1.2v 246 μ w cmos latched comparator with neutralization technique for reducing kickback noise. In TENCON 2010 - 2010 IEEE Region 10 Conference, pages 1162 –1165, nov. 2010. doi: 10.1109/TENCON.2010.5686392. 110
- [127] Xiaochuan Guo, Xin Qi, and J.G. Harris. A time-to-first-spike cmos image sensor. Sensors Journal, IEEE, 7(8):1165 –1175, aug. 2007. ISSN 1530-437X. doi: 10.1109/JSEN.2007.900937. 115
- [128] Ran Zheng, Tingcun Wei, Deyuan Gao, Feng Li, and Huiming Zeng. Optimizing techniques for charge injection effect of pixels in cmos image sensor. In Computer Application and System Modeling (ICCASM), 2010 International Conference on, volume 10, pages V10–258 –V10–261, oct. 2010. doi: 10.1109/ICCASM.2010.5622800. 116
- [129] Nihal Kularatna. Electronic Circuit Design From Concept to Implementation, chapter Data Converters, pages 333–337. CRC Press, 2008. 120
- [130] Wai-Kai Chen, editor. The VLSI Handbook, Second Edition, chapter Nyquist-Rate ADC and DAC, pages 58–158–33. CRC Press, 2007. 120
- [131] Adedeji B . Badiru and Olufemi A . Omitaomu. Handbook of Industrial Engineering Equations, Formulas, and Calculations. CRC Press, 2010. 124
- [132] Chih-Wen Lu. High-speed driving scheme and compact high-speed low-power rail-to-rail class-b buffer amplifier for lcd applications. Solid-State Circuits, IEEE Journal of, 39(11):1938 – 1947, nov. 2004. ISSN 0018-9200. doi: 10.1109/JSSC.2004.835821. 128
- [133] R. Hogervorst, J.P. Tero, R.G.H. Eschauzier, and J.H. Huijsing. A compact power-efficient 3v cmos rail-to-rail input/output operational amplifier for vlsi cell libraries. Solid-State Circuits, IEEE Journal of, 29(12):1505 –1513, dec 1994. ISSN 0018-9200. doi: 10.1109/4.340424. 128
- [134] M. Sinha, S. Hsu, A. Alvandpour, W. Burlison, R. Krishnamurthy, and S. Borkar. High-performance and low-voltage sense-amplifier techniques for sub-90nm sram. In SOC Conference, 2003. Proceedings. IEEE International [Systems-on-Chip], pages 113 – 116, sept. 2003. doi: 10.1109/SOC.2003.1241474. 132

REFERENCES

- [135] Nolan Goodnight, Rui Wang, Cliff Woolley, and Greg Humphreys. Interactive time-dependent tone mapping using programmable graphics hardware. In ACM SIGGRAPH 2005 Courses, SIGGRAPH '05, New York, NY, USA, 2005. ACM. URL <http://doi.acm.org/10.1145/1198555.1198783>. 141
- [136] Ching-Te Chiu, Tsun-Hsien Wang, Wei-Ming Ke, Chen-Yu Chuang, Jhih-Rong Chen, Rong Yang, and Ren-Song Tsay. Design optimization of a global/local tone mapping processor on arm soc platform for real-time high dynamic range video. In Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on, pages 1400–1403, oct. 2008. doi: 10.1109/ICIP.2008.4712026. 141
- [137] Firas Hassan and Joan Carletta. An fpga-based architecture for a local tone-mapping operator. Journal of Real-Time Image Processing, 2:293–308, 2007. ISSN 1861-8200. URL <http://dx.doi.org/10.1007/s11554-007-0056-7>. 10.1007/s11554-007-0056-7. 143
- [138] Ching-Te Chiu, Tsun-Hsien Wang, Wei-Ming Ke, Chen-Yu Chuang, Jhih-Siao Huang, Wei-Su Wong, Ren-Song Tsay, and Cyuan-Jhe Wu. Real-time tone-mapping processor with integrated photographic and gradient compression using 0.13 μm technology on an arm soc platform. Journal of Signal Processing Systems, 64:93–107, 2011. ISSN 1939-8018. URL <http://dx.doi.org/10.1007/s11265-010-0491-8>. 10.1007/s11265-010-0491-8. 143
- [139] Natasha Gelfand, Andrew Adams, Sung Hee Park, and Kari Pulli. Multi-exposure imaging on mobile devices. In Proceedings of the international conference on Multimedia, MM '10, pages 823–826, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-933-6. URL <http://doi.acm.org/10.1145/1873951.1874088>. 143
- [140] Apple. Iphone 4 user guide. URL http://manuals.info.apple.com/en_US/\iphone_user_guide.pdf. 143, 164
- [141] Pixim Products and Technology. URL <http://www.pixim.com/products-and-technology/technology>. 143
- [142] W. Bidermann, A. El Gamal, S. Ewedemi, J. Reyneri, H. Tian, D. Wile, and D. Yang. A 0.18 μm high dynamic range ntsc/pal imaging system-on-chip with embedded dram frame buffer. In Solid-State Circuits Conference, 2003. Digest of Technical Papers. ISSCC. 2003 IEEE International, pages 212 – 488 vol.1, 2003. doi: 10.1109/ISSCC.2003.1234271. 143
- [143] Photonfocus. Linlog technology, . URL <http://www.photonfocus.com/html/eng/cmos/linlog.php>. 143
- [144] C. Posch, D. Matolin, and R. Wohlgenannt. A qvga 143db dynamic range asynchronous address-event pwm dynamic image sensor with lossless pixel-level video compression. In Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International, pages 400–401, feb. 2010. doi: 10.1109/ISSCC.2010.5433973. 144

-
- [145] Xilinx. Spartan-3 fpga family datasheet, . URL http://www.xilinx.com/support/documentation/data_sheets/ds099.pdf. 149
- [146] Analog Devices. Ad7399: Quad, serial-input 10-bit dac, . URL <http://www.analog.com/en/digital-to-analog-converters/da-converters/ad7399/products/product.html>. 149
- [147] Analog Devices. Quad-channel digital isolators adum1400/adum1401/adum1402, . URL http://www.analog.com/static/imported-files/data_sheets/ADuM1400_1401_1402.pdf. 149
- [148] Texas Instrument. Hex inverter sn74lvc04. URL <http://www.ti.com/product/sn74lvc04a>. 149
- [149] BSI. Bsi bs62lv8001 datasheet. URL <http://www.brilliancesemi.com/product/BS62LV8001.pdf>. 149
- [150] FTDI. Ftdi ft245bl data sheet. URL http://www.ftdichip.com/Support/Documents/DataSheets/ICs/DS_FT245BL.pdf. 149
- [151] Linear Technology. Lt3021. URL <http://www.linear.com/product/LT3021>. 149
- [152] National Semiconductor. Lm1117. URL <http://www.national.com/mpf/LM/LM1117.html#Overview>. 149
- [153] XP Power. Il series. URL <http://www.xppower.com/orderPriceList2.php?seriesid=100138&lang=EN>. 149
- [145] Xilinx. Platform flash in-system programmable configuration proms, . URL http://www.xilinx.com/support/documentation/data_sheets/ds123.pdf. 149
- [155] Euroquartz. Xo91 oscillators datasheet. URL <http://www.euroquartz.co.uk/Portals/0/xo91.pdf>. 149
- [156] Bruce Fraser and Jeff Schewe. Real World Image Sharpening with Adobe Photoshop, Camera Raw, and Lightroom. Peachpit Press, second edition, August 2009. 151
- [157] D.B. Goldman. Vignette and exposure calibration and compensation. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(12):2276 –2288, dec. 2010. ISSN 0162-8828. doi: 10.1109/TPAMI.2010.55. 152
- [158] Seunghyun Lim, Jeonghwan Lee, Dongsoo Kim, and Gunhee Han. A high-speed cmos image sensor with column-parallel two-step single-slope adcs. Electron Devices, IEEE Transactions on, 56(3):393 –398, march 2009. ISSN 0018-9383. doi: 10.1109/TED.2008.2011846. 153
- [159] Juha Alakarhu. Image sensors and image quality in mobile phones. 2007 International Image Sensor Workshop, 2007. 159

REFERENCES

- [160] Peter Catrysse, Peter B. Catrysse, and Brian A. W. Optical efficiency of image sensor pixels. Journal of the Optical Society of America A, 19:1610–1620, 2002. 163
- [161] dpreview. Sony cyber-shot w80 review, . URL <http://www.dpreview.com/reviews/sonyw80/>. 164
- [162] Photonfocus. Photonfocus linlog mv-d752e-40-u2-12, . URL <http://www.photonfocus.com/html/eng/products/\products.php?prodId=55>. 164