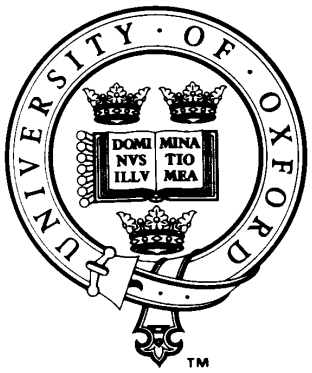
SIMPLIFIED FIXED PATTERN NOISE CORRECTION & IMAGE DISPLAY FOR HIGH DYNAMIC RANGE CMOS LOGARITHMIC IMAGERS



STEPHEN O. OTIM.

LINACRE COLLEGE

Submited for the Degree of
Dphil in Engineering Science,
Department of Engineering Science,
University of Oxford.

Michaelmas Term, 2007.

Abstract

Biologically inspired logarithmic CMOS sensors offer high dynamic range imaging capabilities without the difficulties faced by linear imagers. By compressing dynamic range while encoding contrast information, they mimic the human visual system's response to photo stimuli in fewer bits than those used in linear sensors. Despite this prospect, logarithmic sensors suffer poor image quality due to illumination dependent fixed pattern noise (FPN), making individual pixels appear up to 100 times brighter or darker.

This thesis is primarily concerned with alleviating FPN in logarithmic imagers in a simple and convenient way while undertaking a system approach to its origin, distribution and effect on the quality of monochrome and colour images, after FPN correction. Using the properties of the Human visual system, I propose to characterise the errors arising from FPN in a perceptually significant manner by proposing a novel error measure.

Logarithmic operation over a wide dynamic range is characterised using a new model; $y_{ij} = a_j + b_j \ln\left(\exp\left(\sqrt{c_j + d_j x_i}\right) - 1\right)$, where y_{ij} is the response of the sensor to a light stimulus x_i and a_j, b_j, c_j and d_j are pixel dependent parameters. Using a proposed correction procedure, pixel data from a monochromatic sensor array is FPN corrected to approximately 4% contrast error over 5 decades of illumination even after digitisation - accuracy equivalent to four times the human eyes ability to just notice an illumination difference against a uniform background.

By evaluating how contrast error affects colour, the possibility of indiscernible residual colour error after FPN correction, is analytically explored using a standard set of Munsell colours. After simulating the simple FPN correction procedure, colour quality is analysed using a ΔE_{76} perceptual metric, to check for perceptual discrepancies in image colour. It is shown that, after quantisation, the FPN correction process yields $1-2\ \Delta E_{76}$ error units over approximately 5 decades of illumination; colour quality being perceptibly uniform in this range.

Finally, tone-mapping techniques, required to compress high dynamic range images onto the low range of standard screens, have a predominantly logarithmic operation during brightness compression. A new $Logr'Gb'$ colour representation is presented in this thesis, significantly reducing computational complexity, while encoding contrast information. Using a well-known tone mapping technique, images represented in this new format are shown to maintain colour accuracy when the green colour channel is compressed to the standard display range, instead of the traditional luminance channel. The trade off between colour accuracy and computation in this tone mapping approach is also demonstrated, offering a low cost alternative for applications with low display specifications.

Acknowledgements

I wish to first, thank GOD almighty for all the wisdom and strength that he has accorded to me during my time in Oxford leading to the successful start and completion of this work. Nothing would have come to fruition without his will.

The constant guidance and watchful eyes of my supervisor, Dr. Steve Collins has been immensely enriching. Right from the beginning he has encouraged me to become a better scientist and even when the times were rough he was at hand to guide me through. His keen eye for detail, quality and self-criticism has certainly gone a long way in making me a more analytical mind. After joining the Micro-electronics Circuits and Analogue Devices (MCAD) group in the Engineering Department in the autumn of 2002, Dr. Dileepan Joseph and Dr. Bhaskar Choubey have been very close. I do not think I can thank you enough for making my stay in the laboratory as homely as possible. I enjoyed the debates, arguments, trips (you know where) and ofcourse Bhaskar's influence of spicy (and healthier) Indian vegetarian food.

I am eternally indebted to my family for all the emotional support I received despite being many thousands of miles away from sunny equatorial home. Leaving Uganda to start my PhD, meant meeting my family only a couple of times during the duration of my doctorate, a situation my family (read mum) found difficult, but took only too well. In particular, my lovely and most adorable wife, Dr. Mollynn G. Otim has shown me that patience, faith and love can indeed endure. Despite the many years we were apart, she always encouraged me and most of all believed; I shall forever be beholden.

Lastly, the immense support of the Rhodes Trust towards financing my doctorate remains most appreciated. Since that chilly September month in 2002, the warm welcome of Catherine King, Mary Eaton, Sheila Partridge, Bob Wiley and all at the Rhodes trust helped me settle into Oxford student life almost immediately. My most hearty gratitude to you all.

Dedicated to the family of Mr. Oryang Kinyera (R.I.P) and Mrs Akoth Kinyera.
Dad we miss you so much!

Contents

List of Figures

List of Tables

List of Abbreviations

AC - Absolute Contrast

ADC - Analogue-to-digital Converter

APS - Active Pixel Sensor

CCD - Charge Coupled Device

CDS - Correlated Double Sampling

CIA - Charge Integrating Amplifier

CIE - Commission Internationale de l‘Eclairage

CMF - Colour Matching Functions

CMOS - Complementary Metal Oxide Semiconductor

CMY - Cyan Magenta Yellow

EKV - Enz Krummeanacher Vittoz

FPN - Fixed Pattern Noise

FUGA 15RGB - Camera designed by IMEC

HDR - High Dynamic Range

HVS - Human Visual System

IMEC - Interuniversity Microelectronics Centre

JND - Just Noticeable Difference

LDR - Low Dynamic Range

LSB - Least Significant Bit

MOS - Metal Oxide Semiconductor

MSE - Mean Square Error

MSPS - Mega Samples per Second

MUX - Multiplexer

NMOS - N-channel Metal Oxide Semiconductor

PMOS - P-channel Metal Oxide Semiconductor

PPS - Passive Pixel Sensor

PRNU - Pixel Response Non Uniformity

RC - Relative Contrast

RGB - Red Green Blue

sRGB - Standard Red Green Blue

SNR - Signal-to-Noise Ratio

SSE - Sum Square Error

Chapter 1

Introduction

1.1 Background

In the past decade, analogue photographic film has been quickly declining and continuously being replaced by electronic imaging devices. Today, forecasts by market researchers such as INFOTRENDS predict that after digital camera sales outnumbered classical camera sales in 2003, and with high annual growth rates and ever increasing penetration over 80% of the market sales in 2008 will be digital [16]. Similarly, relatively new imaging applications such as webcams are anticipated to have their market tripling in under half a decade.

The fast decline in the use of analogue modes of image acquisition and the rise of electronic means is very vivid in the digital capture-imaging field. With the emergence of Charged Couple Devices (CCDs) in 1970 [17] and the maturation of Complimentary MOS (CMOS) detectors in the $1990's$, the industry has experienced the increasing benefits of dual but similar technologies. Both provide faster times from capture to an actual tangible image, no specialised skills in the acquisition of digital images, large storage capabilities, lower long-term costs and the possibility to share, manipulate and transfer these images over many platforms in real time.

Charged couple devices, having benefited from longer development and maturation, yield superior quality colour images, higher signal to noise ratios, high photosensitive area and low readout noise [18]. This is due to the specialised fabrication process dedicated to the manufacture of CCDs. On the other hand, CMOS devices that are manufactured in the same process as conventional digital logic and memory arrays, benefit from lower fabrication costs, lower power consumption, random addressability, integration

of several functions on a single chip and faster readout or performance speeds.

In the early 1960′*s* and before the development of CCDs, numerous MOS imaging approaches that coupled NMOS/PMOS or bipolar logic were attempted, receiving varied but not significant levels of success [18, 19]. This presented CCDs as a clear leader while slowly improving on its high and many power levels, quantum efficiency, crosstalk and blooming. However, as special niche markets such as low cost imaging and radiation/IR/automotive imaging started to evolve, coupled with technology scaling, CMOS sensors started to become a relevant part of the industry. In fact, a decade ago, it was predicted that CMOS imaging was destined to overtake CCD technology in most applications [18, 20].

Although this is not true today, both technologies co-exist, with each trying to improve its drawbacks and meet the demands of both previously existing and new imaging applications. Just as CCDs have lowered their power requirements and pixel sizes, CMOS design has shown great strides and flexibility in constantly improving the image quality through the reduction of noise sources and optimising the fabrication process to yield comparable image quality (both monochrome and colour). As a result, CCDs have mainly remained the preserve of high-end scientific, medical and high performance professional cameras while CMOS imagers span the entire spectrum of applications but at the expense of some functionality such as image quality or sensitivity.

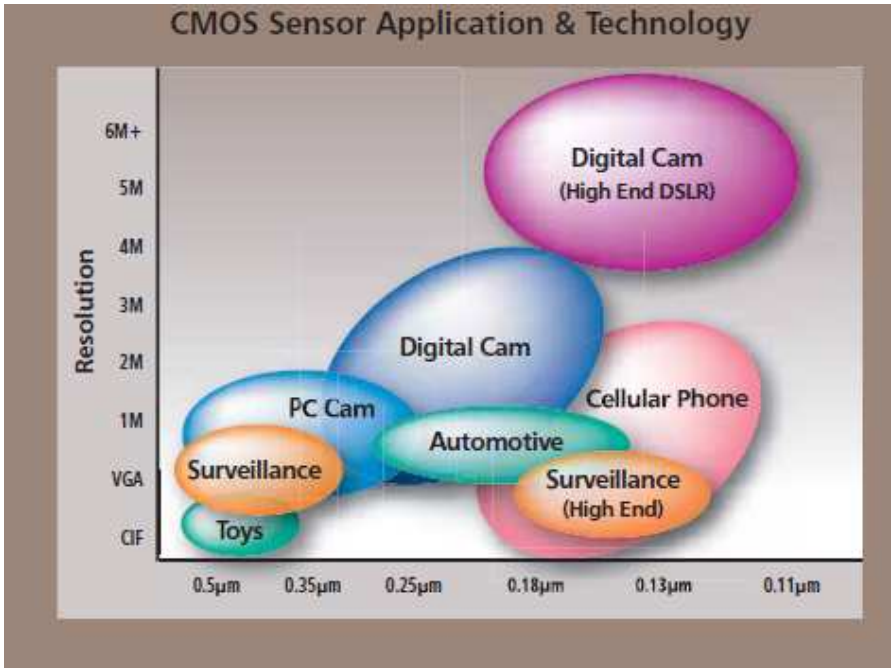


Figure 1.1: The ubiquitous nature and wide span of Complimentary MOS sensor applications implemented in different technologies. [Adapted from Taiwan Semiconductor Manufacturing (TSCM)]

With the possibility of integrating functionality plus analogue and digital signals on a single chip, it is possible to envisage the multitude of applications, as shown in figure 1.1, for which CMOS sensors can deliver sufficient performance on several sub-micron processes. From mobile applications [21], automotive and vehicle safety [22, 23], defence/surveillance, machine/industrial and vision control, medical [24] to high speed large format imaging [25–27], CMOS sensors offer immense flexibility in meeting application specific requirements. Nevertheless, in most of these applications CMOS imagers still fall short on some specifications. In the low cost applications, image quality is mediocre, while in highly integrated systems, costs in terms of design time tend to increase.

However, newer applications in highly specialised fields such as medical imaging (radiology analysis, 3D scanning), surveillance/security (unusual behaviour patterns), space navigation, digital image restoration, colour gamut extension and others, begin to demand more from contemporary sensors. There is a strong need for sensors that can provide high quality, wide dynamic range images at low cost and power while integrated with the rest of the system for high performance in robust environments and conditions. For example automotive applications require 6 decades of illumination as a necessary specification. Even low-end systems such as those embedded in consumer applications as well as low light environments require better images than previously available. It is therefore imperative to develop either new technologies or approaches that achieve better performance, or explore the current sensors and devise ways of overcoming their shortfalls.

In this work, the latter is investigated. First, a review of current electronic imaging design technologies is presented, while comparing their performance in terms of dynamic range, image quality, ease of integration, manufacturing cost, speed and power consumption - section 1.2. It will be shown that, in the pursuit of a low cost, wide dynamic range sensor that yields high quality images, linear CCD and APS sensors fall short of dynamic range and have to employ other cost increasing techniques while logarithmic sensors meet the dynamic range requirements but suffer inferior image quality.

In deciding whether to improve techniques of increasing APS dynamic range or logarithmic sensor image quality, a relevant consideration is what the majority of images are used for in most applications. Since most captured images are viewed and assessed by the human visual system, it is fitting to use the HVS as a reference for determining the mode of capture as well as the quality of the image. Studies have shown that human vision is approximately logarithmic [3, 6, 28] with contrast information being encoded at scene capture over about 10 decades of illumination. With a logarithmic representation providing other benefits that an imaging system can take advantage of such as tone mapping and colour constancy (section 1.3), it will be argued that capturing images logarithmically is indeed logical. To appreciate this

decision fully, existing imaging technologies are first reviewed and then compared accordingly.

1.2 Electronic Imager Designs

Electronic imaging sensors can be broadly divided into Charged Couple Devices and Complimentary MOS based sensors depending on their manufacturing process. They all contain a photoreceptor element which, like photosensitive film, interacts with light energy reflected from an object in a scene. Although photoreceptors are present in many forms (photogates [29], photodiodes and phototransistors [4]) the complete characterisation of the photodiode by Chamberlain and Lee in 1984 [30], and later by Yadid *et al.* [31], cemented their use in many pixel sensors since they provided consistent linearity and were easy to integrate in both CCD and CMOS manufacturing processes. Nevertheless, like other receptors, their basic function remains that of charge capture using energy from photons during charge collection. It is this charge that is transferred and converted into a desirable quantity such as a voltage, time, current or frequency [32], before meaningful illuminant, spectral and spatial information is derived.

This image data, usually pre-processed before analogue to digital conversion, forms the best raw information a particular pixel can derive in any conditions. This is because every sensor has characteristics inherent in its design that pre-determines the best possible performance. For example, a sensor's fill-factor, quantum efficiency, signal to noise ratio and dynamic ratio include some of the performance criteria used to characterise an electronic imager.

1.2.1 Charged Couple Devices

Following their invention in 1970 by Boyle and Smith [17] of Bell Laboratories as a form of circuit memory, charged couple devices have been easily adopted in digital imaging because of silicon's good response to visible radiation and a simple, efficient architecture and design [18]. Photons falling on the photo detector surface transfer their energy, freeing electrons that build up charge in the pixel potential well. Their "bucket bridge" architecture involves this stored charge, being serially transferred along the substrate by a sequence of pulses to an output amplifier [33, 34]. With the linearity of the photo detector and device design, the collected charge is directly proportional to the number of photons captured and the duration of exposure.

CCDs are designed in several configurations, the main ones being full frame CCDs, full transfer CCDs and interline CCDs. In the first, the rows of the captured scene, now represented as collected charge, are shifted down one row at a time along the columns of the array, before being serially sent

into the output amplifier. A high-resolution density is obtained from this simple, easy to fabricate design but comes with some form of shuttering as a requirement to avoid image overlap. The frame transfer mechanism is similar, only that instead of shifting one row at a time down to the serial register, the entire array is read out to a photo-shielded buffer (storage) array of equal dimensions so the actual array is freed to read out another frame. This design does not require shuttering and leads to faster read out times but at higher costs since twice the array size is built. Fortunately, both full frame and full transfer CCDs use photogates as photo detectors which despite having a reduced sensitivity due to the polysilicon gate, enjoy near 100% fill factor. This allows the pixel to hold as much photo generated signal as the well capacity[1] can hold. This is directly linked to the dynamic range[2] of the pixel since a greater well capacity translates to a wider range of illuminations that can be detected.

The above CCDs, particularly the full frame transfer, suffer from image smear, caused when charge collection occurs while the charge associated with the previous readout is still being read, leading to a misrepresentation of the captured image. Interline CCDs successfully avoid this artefact by employing a line of light shielded channels adjacent to the pixels in order to separate image capture and readout as shown in figure 1.2. Here collected charge is transferred from each pixel and immediately pulsed to this interline and then transferred down the column before being serially readout to the output amplifier to form a full image representation. Consequently costs of these CCDs tend to rise due to additional complexity of design and timing. There is also a lowered sensitivity due to a lower fill factor[3], typically by 50%, as the pixel is now reduced in size.

Unlike in full transfer devices where photogates are employed, interline CCDs always utilize photodiodes which suffer the low fill factor but trade that for better sensitivity [19, 35] especially at blue wavelengths. However these photodiodes have image lag when the charge is being relayed from the photodiode to the CCD, as well as reduced potential well capacity as a result of their reduced size. To increase the amount of photon-generated charge collected by the interline CCDs, micro lenses are added to concentrate the incident light onto the photodiode areas of the pixel therefore boosting the effective fill factor to nearly 90%. Micro lenses are also known to shorten the exposure times since more energy is captured quicker but have the problem of becoming less effective as technology scales [19]. This is because in addition to extra cost and complexity, micro-lensed pixels suffer vignetting, cross talk and light scattering all of which affect performance. On balance, they are a welcome addition to optimising

[1] Well capacity, usually measured by the number of electrons, defines the amount of charge, an individual pixel can hold before fully saturating.

[2] The dynamic range of a pixel is the ratio of the highest non-saturating signal to the noise floor signal, usually expressed in dB.

[3] Fill factor is a ratio of the photosensitive area of a pixel to the total area of the pixel and is a measure of the amount of exposure to illumination a pixel is designed to have.

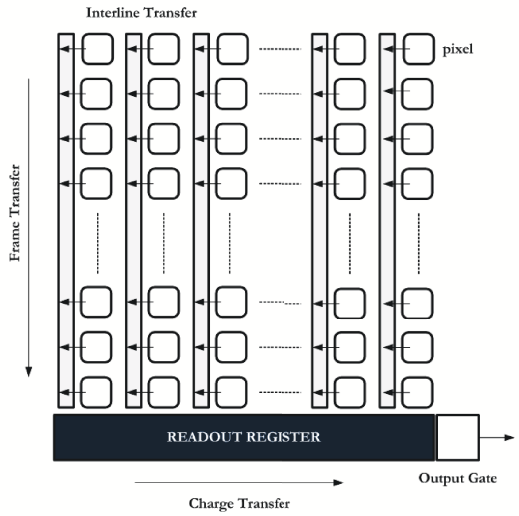
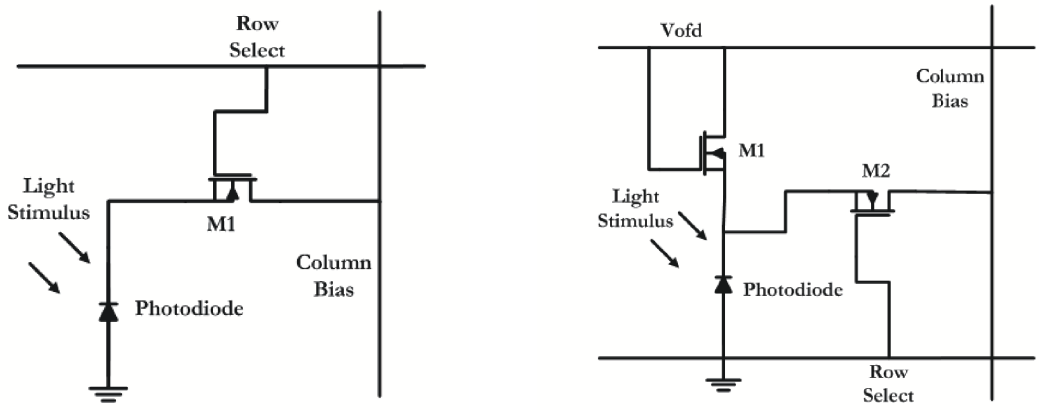


Figure 1.2: The basic structure of an interline Charged Couple Device (CCD).

interline CCD performance since faster, less noisy operation is utilised for the trade off of lower fill factor.

Overall, the CCD's compact but simple nature of design offers high-density spatial resolution sensors since no extra functional integration or processing circuitry is added within the chip. Because of the limited extra structures and a mature specialised process, pixel response non-uniformity (PRNU) is low. Additionally, low dark and readout noise (approximately $23 - 46$ electrons rms [36]), collectively yield high signal to noise ratios that are very favourable; and with today's small pixels sizes, high fill factor (close to 100%) and sensitivity of image capture, CCDs currently set the standard for high quality images in the industry [29].

Unfortunately, manufacturing CCDs is a highly specialised and difficult process to optimize. Extra digital and analogue functionality such as timing control cannot be integrated since they are manufactured in a separate process and even if additional blocks are required for specific designs and applications it often proves difficult. This has been a direct hindrance to the use of CCDs in low cost highly integrated devices and systems. They are also susceptible to charge overflow when over exposed wells overfill with charge that begin to leak to neighbouring substrates, resulting in streaking or blooming which, if uncorrected, appear as bright lines in CCD images. An attempt to control this saturation effect has led to the use of anti-blooming or overflow drain structures, which significantly reduce leakage [36]. For pixels whose wells are filled, the maximum dynamic range is reached and for most typical CCDs ranges between 50dB - 70dB [19]. For commercial and natural scenes, such as those in automotive applications that can span over 120dB [35], this is insufficient and CCDs will cause saturation in the bright parts of



(A) (B)

Figure 1.3: Passive Pixel Sensor (PPS) basic structure - (A) while (B), is a modification with an extra transistor $M1$ to minimise blooming effects.

the image and/or under exposure in the darker parts of the scene.

In addition, the numerous high voltages lines (usually $3-4$) that are used to pulse the charges across pixels and registers use considerable power hence limiting their use in small low cost battery-powered portable devices. Even in moderate applications, power consumption significantly increases costs. Only in high-end professional applications, where image quality is emphasised over costs, is power utilisation usually outweighed. Lastly, the serial output of CCD imagers is a slow approach to charge collection since the charge is sequentially pulsed from pixel location to another along the substrate making CCDs ill suited for fast frame rate applications.

1.2.2 Linear CMOS Devices

Passive Pixels

The use of complimentary MOS technology in imaging sensors resumed in the early $1980's$ after a period of relative failure [18,19]. With the intention of having a sensor that could be manufactured in the existing fabrication process, additional costs associated with a specialised CCD process could be avoided and the added benefit of lower power consumption and integrated functionality, utilised. The first passive pixel sensor basic configuration, figure 1.3(A), incorporated a photodiode and a single row select transistor $M1$, which was used to connect the pixel to the column bus. All the column buses along all rows were then connected to a single charge-integrating amplifier (CIA) at the bottom end of the array, which was used by all the pixels for read out.

A modification of the basic design shown in figure 1.3(B), shows an extra lateral overflow drain transistor $M1$ used to reduce blooming effects. With V_{ofd} chosen a value just above the threshold voltage V_{th}, the overflow drain transistor is switched on and begins to conduct when the voltage at the cathode of the photodiode drops towards $V_{ofd} - V_{th}$, thereby reducing the pixel discharge rate. This way no saturation and blooming efects are realised.

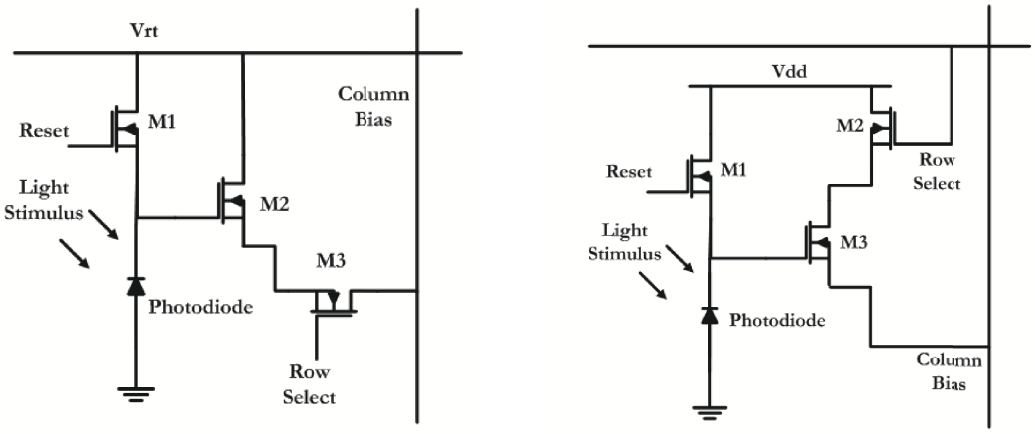
Irrespective of the configuration, passive pixel sensors were built in the same technology processes as other memory and logic circuitry and therefore yielded quite large pixels [18, 19]. Nevertheless, this design offered the largest fill factor for a given pixel size and the smallest pixel size for a specific process. They consumed less power since only a single power line was necessary for the entire chip. Artefacts such as blooming and streaking were also easily eradicated in a simple, easy to fabricate process [33]. Extra functionality that was harder to include in CCD designs, found no difficulty being incorporated into the chip; for instance adding $X - Y$ addressing transistors for random pixel selection and windowing for applications that offer zoom functionality.

However, the reliance of passive pixel sensor on a single amplifier readout does make them slow relative to CCDs, hence creating a big difficulty if large resolution arrays are to be built for use in video applications. In spite of a small pixel size and large fill factor, this design also has low sensitivity and highlights the susceptibility of the entire array to noise due to the large column capacitance. As a result signal to noise ratios are low with readout noise in the order of 250 electrons rms, about ten fold the value for CCDs. Moreover, the passive pixels become impractical with technology scaling since the large arrays result in higher capacitance which in turn create slower response times [33].

Most importantly, pixel responses from passive pixel sensors have an illumination dependent non-uniformity that manifests itself as a time invariant noise. This non-uniformity, called fixed pattern noise (FPN), results from variations in the fabrication of identical pixels, amplifiers, readouts and extra circuitry, intrinsically due to varying doping densities and manufacturing defects. This is very significant because the uncorrected images appear inferior, often as non-uniform line patterns, when a uniform scene is imaged. To reduce the effects of fixed pattern noise, a technique called correlated double sampling (CDS) [36], is employed which minimises FPN while reducing kT/C reset noise[4]. In this procedure, pixel readings when the mechanical shutter is closed are subtracted from readings from the same pixel when a scene has been imaged thereby removing additive components of FPN. This is made possible using a single sample and hold circuit and a comparator that is usually located just before the amplifier.

The dynamic range of passive pixel devices, just like CCDs, is limited to around 60dB, a value

[4] kT/C noise is considered a form of thermal noise which occurs as a result of thermal agitations of charge carriers.



(A) (B)

Figure 1.4: Variations of Active Pixel Sensor (APS) design structure.

still inadequate for high dynamic range imaging if high dynamic range (HDR) image quality is to be guaranteed. While in CCDs, this is attributed to limitations of the well capacity, passive pixel devices are mainly limited by the maximum voltage swing attainable and high noise floor levels. The lack of adequate amplification of the converted charge is seen as the main hurdle to raising the low signal to noise ratio and dynamic range. In addition readout noise tends to impair the quality of the images further, making them inferior to those from CCDs.

Active Pixels

Active pixel sensors (APS) are a direct result of a modification of the passive pixel sensor. The addition of an amplifier source follower circuit to each pixel in the early $1990's$, greatly increased the signal to noise ratios and operating speed. With an extra transistor switch, every pixel could now be addressed spatially making the $3T$ APS a popular configuration. Represented in figure 1.4(A), the reset transistor $M1$ has, connected to its source, a photodiode and a source follower amplifier $M2$ which together with the row select switch $M3$, allow the pixel to be read to the column bus. A variation of this design - figure 1.4(B), in which the position of the source follower and row select are reversed, is sometimes used as a design option.

In both circuits, pixel operation begins with the reset at the gate of $M1$ being pulsed high thereby integrating this charge onto the photodiode capacitance. This capacitance is then gradually discharged by the light incident on the photodiode junction. The source follower transistor then buffers this small capacitance from the large column capacitance before the row select allows readout to the column bus.

Not only does this architecture increase SNR, power dissipation is minimal compared to both PPS and CCDs, since readout capacitance is low and these buffers are on only for the duration of the readout [18]. The lower fill factor as a result of the extra transistors, is also more than compensated by better dynamic range and higher SNR. Note that at approximately 30%, APS fill factor is already similar to those attained by some high performance interline CCDs. To further increase the overall sensitivity of APS pixels, they also utilise polyimide micro lenses to increase the amount of light captured by refracting light away from the photo insensitive areas onto the detector region and also overcome the losses in surface reflections. Some manufacturers such as DALSA corporation have even shown ways to improve image capture capabilities by adding exposure control while eliminating rolling shutters, image lag and anti-blooming [37].

APS sensors have benefited from techniques such as column level analogue to digital conversion [38] and advances in fabrication methods to create architectures with better sensitivity and responsivity to light [18]. The subsequent increases in pixel size due to more circuitry and larger size compromises cost effectiveness, sometimes potentially increasing noise levels. New sub-micron processes, compatible with already existing CMOS processes, have also led to increasing functionality and compatibility at reduced costs. Yadid *et al.* [31], for example, broke new ground in 1991, when they designed a pixel in a 3μm process for complete random scan addressability. Today more functionality than random scanning exists in 0.5μm - 0.13μm processes including motion detection, multiresolution imaging and dynamic range enhancement [18, 39, 40] among others. More complex examples exist where a combination of functionality is employed in a single chip at low cost while on chip colour processing functions such as gamma correction, white balance and colour correction have been demonstrated [21].

A factor affecting APS performance, and considered to be one of the main hindrances to them becoming ideal imagers, is noise [29]. Present in various forms, fixed pattern noise, dark current noise[5], readout noise and temporal noise represent the strongest impact on CMOS imager performance [41]. As explained in section 1.2.2, fixed pattern noise results from variations in the responses of identical pixels due to manufacturing variations such as minutely varying aspect ratios, doping variations and varying capacitances. These differences in the pixel responses need to be accounted for in a calibration process. To date, correlated double sampling constitutes the main way of minimising this noise. True correlated double sampling in APS involves reading all pixels twice, at any illumination within any frame starting with a reset operation. Usually situated along the array columns [42, 43], CDS circuits read the pixel response at the beginning and end of every integration cycle. This ensures that, in addition to pixel and column-to-column FPN, the kT/C and $1/f$ flicker noise associated with this integration cycle are re-

[5]Dark current noise is caused by reverse currents in non-identical photodiodes.

moved. In addition, dark current noise, which appears as part of FPN, is not suppressed by correlated double sampling. This is because the dark current appears not in the reset level, but only in the signal level of the pixel signal.

Despite the success of the FPN reduction methods such as CDS techniques in minimising additive FPN to 0.24% of the saturation signal [42], little effort has gone into the removal of any other components. Liu's [44] attempt at determining and removing gain FPN, using optical flow estimation and linear minimisation numerical algorithms from motion sequences, resulted in significant visual improvement in image quality but acknowledged the delicate balance between complexity and performance.

Nevertheless, with technology scaling, new challenges to pixel design arise. One demerit is the lower voltage swing to accommodate the smaller sizes of the devices. Since the size of the saturated signal is one of the determinants of dynamic range, low voltage swings limit the dynamic range that can be claimed. This compounds the fact that APS pixels already suffer limited dynamic range resulting from the definite amount of charge that can be held at the pixel. Nevertheless, low voltage designs of complimentary APS have been shown to obtain a dynamic range of 70dB, attained with a novel read out method, utilised to increase the potentially low SNRs [43]. Despite the higher capacitances and lower readout gains of the PMOS components of this design, a multitude of other ways to increase the dynamic range in active pixel sensors have also been devised. For example, a stepped reset level on a lateral overflow capacitor was used to create a customised piecewise compression [42]. Using a gate voltage, excess charge was collected at the overflow (analogous to well capacity adjustment) allowing an increase from 70dB to 94dB but at a cost of increased fixed pattern noise and capacitance from the lateral capacitors themselves. Schanz *et al.* [45] tried a unique combination of multiple exposure and multi gain readout to deliver a 120dB range, obtaining a good SNR of 56dB at about 50 frames per second, ideal for applications in the automotive industry. In this approach, the sensor was subjected to different frame exposures ranging from short times for bright parts of the scenes to longer times for the capture of the darker areas of the scene. In addition these low illumination images were allowed higher gains while the bright scenes, lower gains. The resulting image, achieved by stitching together its individual parts, then more accurately represented the entire range of the captured scene. More complicated designs emerged later with Lule *et al.* [46] attempting pixel level automatic shutter, adapting the integration time to the local intensity to achieve 120dB. Stoppa *et al.* [47] even used a combination of a piece-wise linear design, comparators and analogue memory to constantly adapt the single pixel integration time in proportion to the local illumination condition. Despite the 132dB range, the low fill factor and high number of transistors reduced viability. Similarly, the coding of pixel integrating times instead of the

conventional pixel voltages [48] to avoid possible saturation, yielded 100dB; but like most of the above dynamic range improving methods, added complexity, extra circuitry and costs.

Overall, linear APS sensors show a big improvement in image quality from passive pixels in addition to their capability of fast speed performance. However, their dynamic range is still as limited as those of CCDs and although numerous mechanisms of dynamic range expansion have been shown above, extra design or cost seems to be associated. Recent work by Chuang *et al.* [49,50] appears to suggest a changing approach to handling dynamic range. By converting the linear response of an active pixel sensor to a logarithmic one using a piecewise linear approximating analogue to digital converter, they were trying to mimic the human visual systems' approximately logarithmic response to light stimulus [3,6,9]. Allotting the finer resolution to the lower illumination values and gradually less to the higher illuminations, the multi resolution ADC vastly improved contrast at lower illumination but increased quantisation error and ADC cost concurrently. Although, no independent or qualitative/quantitative assessment of the resulting images was made, this contrast encoding imaging concept bears a striking resemblance to another class of CMOS sensors that are non-linear in operation.

1.2.3 Logarithmic CMOS Pixels

All the CMOS devices discussed above operate by integrating a charge onto a capacitance to give a voltage, which is discharged at a rate directly proportional to the integrating time. However non-linear CMOS sensors that have existed since the early $1990's$, generate a voltage that has a logarithmic relationship to the pixel photocurrent. These logarithmic pixels mimic the human eyes response [3, 6, 9] to photo-stimuli and have similar properties to photographic film, which is pseudo-logarithmic. Responses from a logarithmic configuration are independent of the absolute light intensity of a fixed contrast image and encode this contrast information by compressing the dynamic range. Unlike linear sensors that implement real-time vision applications with added design effort, the building of architectures that easily combine time continuous visual analogue processing with photoreception are better suited to logarithmic sensors.

The adaptive photoreceptive circuit, reported by T. Delbruck and C. Mead [51, 52], was the first to contain a logarithmic pixel as an alternative device for the imaging field. Today, logarithmic pixels can be obtained by PMOS or NMOS transistors acting as load transistors in weak inversion and a photosensitive device usually a photodiode [13, 15, 53–55].

The basic logarithmic pixel configuration is shown in figure 1.5, where $M1$ is the load transistor operating in weak inversion with $M2$ as a row select device and $M3$ as a source follower readout. With

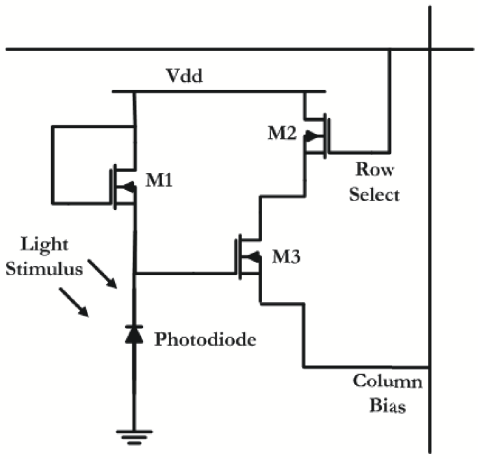


Figure 1.5: The configuration of a 3 transistor logarithmic pixel with source follower readout.

$V_{gs} < V_{th}$ for $M1$, the voltage at source of transistor $M1$ drops as the amount of light increases at the photodiode pn junction, with the relationship between the drain source current I_{ds} and drain source voltage V_{ds} generally given by [54, 56, 57],

$$I_{ds} = \frac{W}{L} I_{do} \exp\left(\frac{qV_{gs}}{nkT}\right)(1 + \lambda V_{ds}) \tag{1.1}$$

$$\therefore \quad V_{out} = V_{dd} - \frac{nkT}{q} \ln\left(\frac{I_{ds}L}{I_{do}W}\right) \tag{1.2}$$

for $1 >> \lambda V_{ds}$. V_{out} is the output voltage, W/L the aspect ratio of transistor $M1$, k the Boltzmann constant, T the absolute temperature, q the electronic charge, n the subthreshold slope parameter and λ represents the channel modulation coefficient. It is clear that for a near unity gain source follower readout, the voltage readout of the column bus is a logarithmic compression of the current flowing through the photodiode. The lack of a time dimension in any of the terms in equation (1.2) highlights one of the strong merits of continuous time operation of logarithmic pixels. This eliminates the complexity of designing multisampling pixels, fast ADCs and extra buffer/memory that is required in most high dynamic range linear sensors. Moreover, the logarithmic compression of the captured images leads to imaged scenes having greater dynamic range than those of linear sensors of any kind, often up to 6 decades. This may not be a very significant factor when imaging low dynamic range scenes but is quite influential when high dynamic range scenes have to be captured and displayed without saturating or underexposure.

This is well demonstrated by the images on the left of figure 1.6 which show both over-exposure and under-exposure in the different parts after being captured by low dynamic range CCDs while the

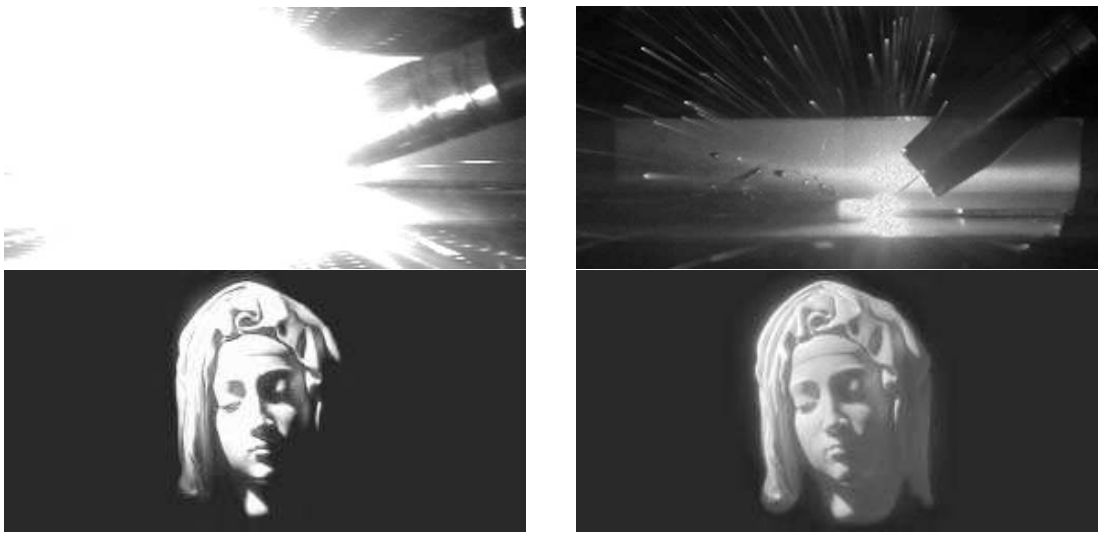


Figure 1.6: Top images: A 256×128 welding scene first captured with a CCD and then with a digital 0.8μm logarithmic CMOS HDRC camera without any exposure control at 120 frames/sec and with a constant access time of 140ns per pixel. At the bottom, another high contrast 640×480 image, and the more natural image when captured from a 0.25μm CMOS logarithmic VGA Imager [1].

images on the right, captured by logarithmic sensors, have both areas with better visibility and clarity[6]. Such capabilities form an essential requirement in the automotive industry where, for vehicular assisted driving, wide ranges of illumination need to be covered, typically up to six orders of magnitude and containing vital images that have to be relayed to the driver as they develop [23, 45]. Logarithmic pixels also meet similar requirements in the machine control industry, X-ray & medical imaging, surveillance and robotics [19, 22, 24].

The high dynamic range suggests a high number of bits required to cover a wide range. Logarithmic pixels, however, need fewer bits per pixel to cover the same wide dynamic range than linear sensors. For instance for a five decade dynamic range, only 10 bits of digital data are required for a 1% accuracy [22] whereas 24 bits per pixel are needed in linear sensors for a similar level of accuracy. This is very significant and invaluable to high dynamic range applications as the higher bit numbers represent a rapid increase in pixel circuits, power consumption, cost and possible complexity. Logarithmic imagers therefore combine the special capabilities of constant contrast sensitivity with the high dynamic range that is ideal for system integration with wide dynamic range applications. They also leverage the other advantages of CMOS process technology such as easy integration of extra functions and circuitry, low cost and power performance, faster readout speeds and the ability of random addressability and sub-windowing.

Due to the low sensitivity of logarithmic sensors, small voltage swings are managed and even for an

[6]Note that the limited dynamic range of printed material/paper further degrades the quality of print media since they are limited to less than 2 decades

average noise floor, the signal to noise ratio is lower than those attained by linear pixels. At a sensitivity of 50mV/decade, a logarithmic pixel spans about 300mV over six decades while an APS may have an output swing of about 1.5V over 3 decades. It is unsurprising therefore that logarithmic sensors are more noise prone and susceptible to image quality degradation. More importantly, this sensitivity drops in low illumination becoming almost unresponsive in near dark environments. This is because in very low illumination the photo detector current is dominated by the dark leakage current noise. Coupled with the fact that temporal noise becomes very significant in raising the noise floor, good sensor operation in low illumination is even made harder to achieve. All the issues above mean that noise significantly limits logarithmic pixel performance.

However, above all noise sources, fixed pattern noise plays the greatest role in limiting the quality of images from these non-linear pixels. As mentioned before, non-uniform responses to a uniform scene from identical pixels cause inferior image quality. Because of manufacturing defects caused by variations in doping densities and threshold voltages, an illumination dependent noise appears as random speckled non-uniformity sometimes even as pronounced vertical lines. If uncorrected, fixed pattern noise is known to cause pixels in the images to appear as high as an order or two magnitudes, brighter or darker than expected [14, 15, 58].

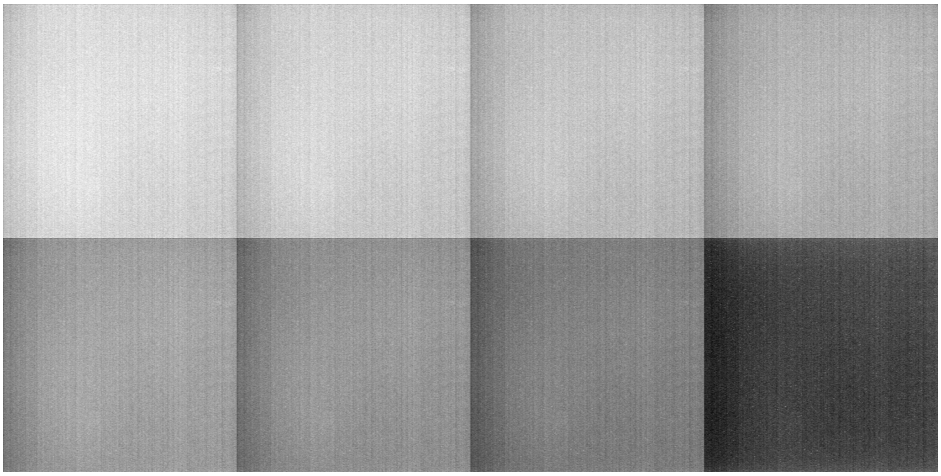


Figure 1.7: Uniform images of a white sheet of paper taken using a FUGA 15RGB logarithmic camera at 8 different illuminations represented by f-stop numbers; 2.0, 2.8, 4.0, 5.6, 8.0, 11.0, 16.0 and ∞.

A simple example of this kind of noise is shown in figure 1.7, where a white piece of paper is uniformly illuminated under room conditions and then imaged by an HDRC FUGA 15RGB logarithmic camera [53] at different exposures. All the images appear degraded and not as uniform as a white sheet of paper should. Pronounced vertical lines that are more distinct at the edges create an unrealistic im-

pression of the originally captured image. FPN therefore needs to be minimised if not eliminated for logarithmic pixels to yield comparable quality images.

The correction of fixed pattern noise has been a major subject in the study of these pixels with several efforts to calibrate logarithmic sensors for this noise. In general the correction procedures such as correlated double sampling [35] and in-pixel subtraction methods [59] that are used in linear sensors, cannot be applied to logarithmic sensors because of their continuous mode of operation. In logarithmic operation, the absence of a reset prevents the use of the any double sampling procedure. Consequently, alternative methods have had to be developed to suit this unique situation, some classified as internal to the pixel structure and operation while others have tended to be external.

In 1995, Ricquier *et al.* [60] attempted to reduce FPN by employing hot carrier injection to alter threshold voltages in a traditional three transistor pixel using a pre-programmed offset thus compensating for their variation. However, the long stressing times together with the high voltages of up to $16V$ outweighed the benefit of the 5mV peak-to-peak FPN that was achieved. The calibration needed approximately 10 minutes per pixel, making the procedure untenable for large arrays or imagers that require fast operations. Diericks *et al.* [61] on the other hand, suggested the use of a post chip software solution to remove previously determined and stored individual pixel offsets. Their implementation had the impediment of slow performance as memory access times and calculations reduced sensor performance. A hardware solution involving an EEPROM and dedicated co-processor was, however, developed in a large chip that allowed for random addressability, trading size for faster frame rates. This performed well, but the extra circuitry, control blocks and a processor placed real limits on fill factor and the available frame rates.

In looking for simplicity while correcting for FPN, some logarithmic manufacturers offered easy recalibration by allowing users to individually correct for offset variations as and when a significant change in environment occurred. In the FUGA 15RGB camera [53] for example, by imaging a uniform scene, the values of the responses of all the pixels would be stored and all subsequent responses subtracted from these values to remove offset variations. In figure 1.8, the uniform white paper, previously imaged in figure 1.7, is shown after subtracting the pixel values at one illumination (f5.6-stop) from all the other images and scaling the result.

The results are very vivid as most of the vertical lines that were present in the uncorrected image have ceased but random speckles still exist which might obscure detail if it were a detailed image. Despite the fact that this calibration method requires no skill, it is inconvenient and cannot be used in commercial or industrial and professional applications. It is prone to human error, relies a lot on uniformly imaging a

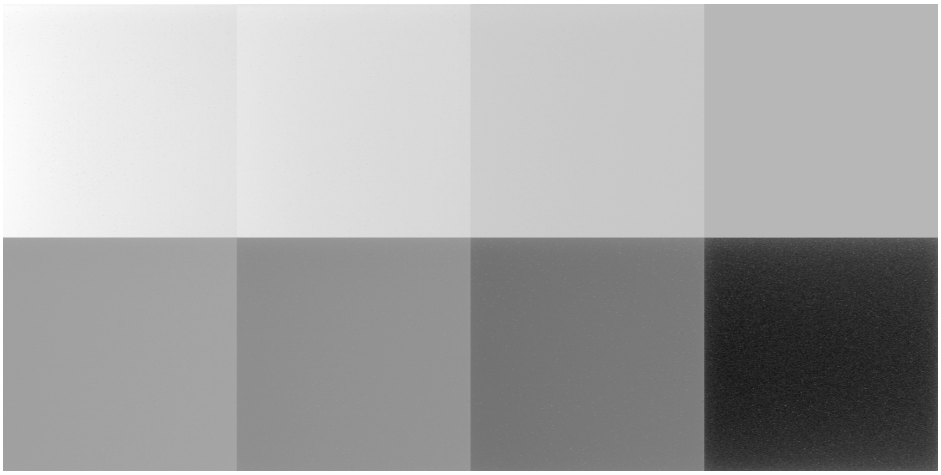


Figure 1.8: FPN corrected images of a uniform sheet of paper taken at the same f-stops as the logarithmic camera in figure 1.7. This is obtained by a subtraction of all the images from that at the $f5.6$ stop and scaling the images.

constant scene and has not been tested in high dynamic range applications.

A similar approach was also tried by Marshall and Collins [62], where FPN correction involved eliminating offsets by subtracting an out of focus image of a scene from a focussed one. However, this approach was not seen as very viable because the quality of the corrected images dropped as the scene changes from the scene used for correction. It also created a contrast encoded, edge enhanced image that is not nice to view. A logarithmic circuit with a p-type floating gate load transistor that employed programmable threshold voltages to reduce fixed pattern noise was also reported by Marshall and Collins [63]. This circuit also provided the pixel output thereby maintaining a high gain but also preventing temperature variations from affecting this output. In spite of this, the mechanical shuttering of the auto focus mechanism and the sizes of the programmable threshold voltage circuits proved prohibitive.

Loose *et al.* [14, 15] stored pixel offsets in a capacitor for each pixel based on a specific reference current determined by a voltage feedback circuit and later subtracted this offset as part of a calibration process. This yielded a reduced FPN noise level of 3.8% per decade at the expense of fairly large pixels in a 0.6μm technology. Instead of storing the individual offsets, Kavadias *et al.* [55, 64] developed a double sampling procedure where a reference current was used to acquire an initial reading before a pixel reading was taken, yielding FPN levels of 2.5% of the output signal range. Although the speed implications of reading each pixel twice in each frame, were noticeable, the size of the five transistor circuit created more variations that the very circuit it was trying to solve.

Until recently, the removal of FPN was based on the fact that its additive components dominated most

of the noise [35,61,65] and although other forms of FPN were thought to exist [55,64], little was done to reduce their effect let alone acknowledge their contribution. The removal of offset components alone did not seem to create satisfactory quality and for this reason recent researchers have endeavoured to study and remove other components of fixed pattern noise. In an X-ray imaging application, Ponchut [24] used logarithmic pixels where the offset, gain and leakage currents were extracted over a 4 decade range using a numerical non-linear fitting algorithm within a software application. Although no residual FPN results were fully discussed, an image with noise $5-10$ times that of comparative CCDs was reported. Extra costs of the computations in separate software would make this solution hard to implement in portable low cost applications such as digital cameras. An on-chip adaptive algorithm was also suggested and developed in the removal of offset and gain effects [22] but similar concerns still arise from the cost and size despite the added advantages of temperature independence and a lack of frame delay.

Even simpler FPN correction was demonstrated when Joseph *et al.* [12,13] devised a method which extracted all FPN components at each pixel location and using these values, reduced illumination independent non-uniformities. Joseph started by modelling the response of a pixel to three components, an offset, gain and leakage current per pixel, using a uniform illumination of a sheet of paper at five different light levels. Using a non-linear optimisation technique, the three components were then removed from all subsequent pixel values to yield significantly reduced FPN to levels of about a least signifncant bit (LSB) over two decades.

1.3 Pixel Comparison

Reviews of the various ways by which current electronic imagers capture and represent widely varied scenes have shown that there is no imager that can be considered ideal. The different capture methods have both merits and demerits that are suited to different, but not all applications. In table 1.1, a comparison of the major features considered desirable in contemporary imaging solutions are shown for all the imagers discussed in section 1.2.

Technology	Image Quality	Cost	Integration	Power	Dynamic Range
CCD	Very High	High	Low	High	Moderate
PPS	Low	Low	High	Low	Low
APS	High	Low	High	Very Low	Low
LOGARITHMIC	Moderate	Low	High	Very Low	Very High

Table 1.1: Relative technology and typical performance comparison. Note that performance speed is generally an application dependent measure and has been excluded.

While CCDs have remained the sensor of choice for the high-end professional applications that emphasise high image quality over cost and power, APS CMOS sensors have become the next best alternative and dominate many of the low cost applications with steadily improving quality. PPS CMOS sensors, on the other hand, do not form a viable option with poor scaling and even lower image quality due to low SNR. However, because CCDs and APS still have limited dynamic range, and increasing this range is costly, improving the image quality from logarithmic sensors does offer the possibility of making them the best-suited sensor. For the reasons below, I propose that logarithmic sensors, are indeed the best approach of easily achieving high dynamic range images in fast applications and at low cost.

1.3.1 Replicating Biology

The Human Visual system (HVS) has been known to have special features concerning its interaction with visual stimuli as well as its range, speed and ease of the interpretation of these stimuli [3, 6, 9]. Amazingly, the eye boasts a wide colour range, adapts to different lighting conditions, responds quickly to most changes and has very fine resolution and acuity. The human visual system is a high dynamic range organ capable of adaptive viewing, covering about 10 decades of scene illumination [3, 66], although not simultaneously. Figure 1.9 shows a response of retinal cone which cover 6 decades overall with up to 4 decades of illumination without adapting to the background illumination [2, 67]. Mindful of the fact that these response curves do not include the response of rods at lower illuminations, they already prove that high dynamic range vision is indeed possible and necessary for better reproduction of natural and artificially captured high dynamic range scenes.

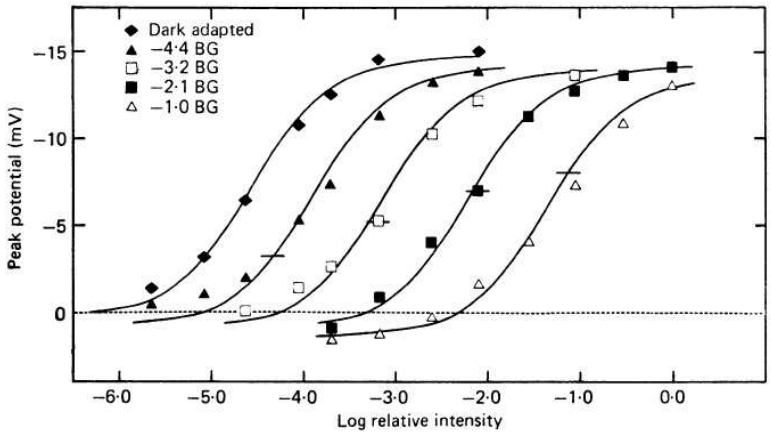


Figure 1.9: The retinal response of green and red cones; adapted at different background illuminations [2].

The figure also shows an adaptation of the response to the average illumination around a visual angle,

with lower responses to more steady light levels while steep changes to the illumination result in higher responses. These responses also reveal an unchanging contrast, which is independent of the illumination level of the image. Contrast encoding is therefore performed during the image capturing process. This means on a logarithmic scale, a high dynamic range is covered, over which responses are only to illumination changes and not at the illumination level itself. Thus implying that the visual system is sensitive to relative rather than absolute changes in illumination.

It is mechanisms such as these that first attracted researchers in the neural science community. In their seminal work on silicon retinas, T. Delbruck and C. Mead [51, 52] attempted to mimic this functionality in designing a pixel with an adaptive logarithmic pixel. Even photographic film, that has been accepted for over a century, has been found to exhibit logarithmic behaviour with the density of the developed film being proportional to the logarithm of the total amount of light to which the film was exposed [66]. Today, as research continues into the workings of the human visual system, logarithmic pixel operation is still seen as the natural way capturing images.

1.3.2 Dynamic Range Compression

In a complete system that needs to capture and display high dynamic range images, logarithmic image capture provides a direct benefit in terms of computational ease. This stems from the fact that the dynamic range of most standard display devices is just above 2 decades of illumination. Since this is far from the range of most natural and industrial scenes that can be up to 120dB [9, 33], displaying these images is potentially as large a challenge as capturing them.

Displaying such wide dynamic range images on traditional displays always results in over or under saturation with entire decades of illumination being unrepresented with unacceptable loss of detail and perceptual accuracy. To avoid this mis-representation, dynamic range compression techniques are used to maintain the relative brightness of the scenes without actually losing the detail and features in the original image. These techniques, called tone-mapping algorithms, are therefore vital in all wide range applications and need to be easy and optimised routines especially if the applications are to be utilised are real-time or video applications.

Many variations of these tone mapping algorithms exist and some are discussed in Chapter 7. However, a common feature in most of these routines, is that in trying to maintain the perceived brightness, the routines all use a logarithmic operation in compressing the dynamic range of the original scene. Therefore images captured with a logarithmic camera readily provide these values without the need for extra computation hence minimising costs through potential power and time savings. For as long as displays

are still low range, these routines will always be needed in applications that are fast becoming high dynamic range oriented, thereby giving logarithmic imagers a clear advantage over linear ones in terms of simplicity and convenience for display.

1.3.3 Colour Constancy

Another feature of the human visual system which makes logarithmic image capture attractive is colour constancy. Light entering the eye or camera is dependent on the magnitude and spectral properties of the incident and reflected light from the object. When an image is viewed under different lighting conditions, they are still perceived as exactly the same object. The visual system therefore tends to disregard the lighting characteristics from the incident source but preserves that which is reflected from the surroundings to maintain the same perception. This property, called colour constancy, is maintained when details, hue, spectral shifts and effects of the light source remain unchanged, irrespective of the spectral characteristics of the illuminant.

Although an involuntary action in the human visual system's response to changing surroundings, colour constancy is very important in applications such as machine vision and robotics where decisions and actions are based on the interpretation of the colour images captured by the sensor - colour based object recognition. Colour constancy allows captured images to appear the same irrespective of the illumination at capture. In other words, independent spectral components of a scene, could be pre-determined before being re-applied to a different environment if image appearance is to remain unchanged.

The goal of most colour constancy algorithms is to transform the pixel responses so that they become independent of the illumination and hence correlate with the reflectance properties of an object. A common feature in most of these colour constancy algorithms [68–71] is the application of the logarithmic operation on the pixel responses in the routines. As an example, in calculating the "lightness path" for each receptor class as a first step in transforming a pixel's response $I_k(x_1,y_1)$ for colour constancy, the difference between its logarithmic response and that of a neighbouring pixel $I_k(x_2,y_2)$ is determined by [72]

$$R_k(x_1,y_1) = \log I_k(x_1,y_1) - \log I_k(x_2,y_2) \tag{1.3}$$

where k is a colour channel, $R_k(x_1,y_1)$ is the new pixel value and (x,y) corresponds to the spatial location of the pixel on an array. This way, lightness paths at a pixel location with respect to all its neighbours can be determined for each colour band to give characteristics of the image that are dependent on the spectral

properties of the image but not the viewing environment. Later, other researchers such as Finlayson *et al.* [73] proposed a gamma and illumination independent hue H, defined as;

$$H = \tan^{-1}\left[\frac{\log R - \log G}{\log R + \log G - 2\log B}\right] \tag{1.4}$$

used to maintain the inherent colour characteristics of a scene. By applying the fact that colour hue is the part of the image signal that is invariant to brightness changes, its pre-determination in an image forms the unchanging component of the scene when viewed under any illumination. Logarithmic sensors therefore present a unique opportunity for easier and more convenient application of colour constancy algorithms than linear sensors, usable in special applications such as object recognition software.

1.4 Thesis Objectives

With a favourable dynamic range and the other post capture merits of logarithmic sensors, mentioned in section 1.3, the main task remains mitigating the major drawbacks of these sensors for the variety of applications that are demanding more performance from current imagers.

This thesis is primarily concerned with improving image quality from logarithmic sensors by alleviating the most dominant form of noise. Despite the plethora of attempts to correct FPN, no approach has successfully quantified FPN in a high dynamic range setting and given perceptual image quality assessment of the resulting images. Reducing illumination dependent fixed pattern noise (FPN) in a simple and convenient calibration scheme and to levels comparable to the human visual system's ability to discriminate contrast, will be the main goal of this work. A systematic approach will be taken to identify the sources of FPN before reducing pixel mismatch from both monochrome and colour images over a wide dynamic range.

Another area this thesis intends to address, is that of the manipulation and display of high dynamic range images on standard screens. This is because retaining perceptual image and colour accuracy remain of prime importance during display, to ensure fidelity and avoid image distortion. With the lack of high dynamic range standard displays, the author will propose a logarithmic colour representation that can be conveniently used in most tone mapping operators while avoiding the computational burden incurred when using linear sensors. By compressing the dynamic range of the green channel, it will be shown that image quality is generally maintained, both from a visual and quantitative perspective, continuing the general theme of simplicity and convenience.

1.5 Thesis Structure & Layout

1.5.1 Organisation

The rest of the thesis is arranged as follows. In chapter 2, a model for logarithmic pixel response that stipulates the sources of fixed pattern noise, is first reviewed. The need for an easy self calibrating scheme is discussed in view of the effect of temperature and viewing conditions on image performance. An optimised electronic calibration circuit is therefore described before a new biologically inspired contrast error measure is proposed. Using simulations, typical pixel response behaviour is then studied. Previous offset correction methods are compared using the new contrast method before a simpler offset and gain correction, and corresponding calibration strategy, suggested.

In chapter 3, logarithmic pixel behaviour is investigated in high dynamic range terms after highlighting its failure in very low and very high illuminations. The focus is on understanding, from a physical perspective, the reasons behind the changing sensitivities of logarithmic pixels in different ranges of light levels. A complete HDR model for pixel response is then developed which is tested using a complex non-linear optimisation and a simpler parameter extraction technique. This is followed by the development of two practical FPN correction methods, usable in most logarithmic pixels.

Practical FPN correction in monochrome logarithmic pixels is demonstrated in chapter 4. First, the realisation that contrast error varies minimally and that the HVS is more sensitive to relative changes than absolute ones, leads to the adoption of relative error as the appropriate metric of choice. Using real logarithmic pixels from a 200×100 array, response behaviour is typified and the nature of fixed pattern noise characterised along the rows and columns of the array. The effect of uncorrected FPN is confirmed before its correction is demonstrated as well as the added effect of quantisation error.

In chapter 5, the study of fixed pattern noise reduction in colour images commences with background colour theory including colour capture, representation, colour spaces and perceptual error measures used in their reproduction. The effect of contrast error on colour quality is also investigated. Using spectral colours from the standard Munsell colour set, acceptable image colour error is explored using these colours to determine visually distinguishable colour error and the perceptual metric to match this error.

A comprehensive MATLAB simulation, modelling FPN correction is described in chapter 6, using the simple correction method suggested from Chapter 4. While accounting for typical logarithmic sensor features such as analogue to digital conversion, the perceptual colour errors are determined over a high dynamic range in order to validate the effect of the FPN correction scheme on colour images.

A discussion on the possibility of displaying high dynamic range colour output from HDR images

is included in chapter 7. With a strong emphasis on ease and simplicity, tone mapping is reviewed as a solution for low dynamic range problem before a new colour representation is suggested that is favourable for logarithmic sensors and makes light work of dynamic range compression. Using HDR images, the proposed approach was shown to preserve perceptual detail with reduced complexity. Colour quantisation is also investigated and as a result, a low cost alternative to displaying already tone mapped images without losing significant colour fidelity, is suggested.

1.5.2 Methods

The investigations in this thesis were conducted using a variety of tools and involved the use of both theoretical and practical data sources. The initial study of fixed pattern noise, particularly the nature of uncorrected fixed pattern noise was undertaken using the FUGA 15RGB commercial camera mentioned earlier. Manufactured by C-Cam Technologies and using a sensor developed at the Interuniversity Microelectronics Centre (IMEC) [53], it was a first generation logarithmic colour camera embedded with the functionality to increase bit length from 8 to 10 bits for better precision and dynamic range. Its calibration procedure, involved manually focussing a uniform image at the illumination of interest, before using it for image acquisition. The FUGA 15RGB camera, however experienced high levels of noise and had a typically slow response hence limiting its use to still rather than video applications. Nevertheless, raw data obtained from imaging typical indoor scenes similar to the ones in figure 2.7, offered insight as well early ideas about colour work in later chapters.

A significant part of this work was also performed using data from the CADENCE circuit simulator in which pixel designs were tested and numerous simulations run. Several Monte Carlo runs of the logarithmic pixel design, were performed to emulate fixed pattern noise thus generating data that could be used to analyse typical pixel responses as well as device and parameter variability. By far the majority of the results in this work are from MATLAB simulations, primarily using functions in the image processing toolbox in addition to custom written code. By taking a system approach to the cause and effects of FPN on image and colour quality, special functions were written to gain better insight and analysis.

Lastly, readings from a 200×100 array of logarithmic pixels designed in a 0.35μm AMS process were also used to characterise real pixel response. With fairly large 10μm $\times 10\mu$m pixels and a 49% fill factor, the chip also provided typical sensor noise levels but was incapable of capturing images or pictures since it was not built as part of a camera and lens system.

Chapter 2

Fixed Pattern Noise

2.1 Introduction

Logarithmic sensors have been suggested as an alternative to linear sensors for imaging high dynamic range scenes; an option that does not involve the complexities of significantly altering pixel circuit operation or slow multiple sampling techniques. With the low number of bits required to image these scenes, they are considered a very attractive alternative except for their susceptibility to fixed pattern noise. The understanding of the nature and impact of FPN lies at the heart of developing a simple and convenient way of alleviating this drawback.

In this chapter, an independent assessment of previous efforts of correcting fixed pattern noise, forms the focus of the investigation. This is based on a new biologically inspired perceptual error measure for the quality of correction. It will be used to benchmark earlier attempts before suggesting a simple approach to alleviate fixed pattern noise over a considerably wide range of illuminations.

First, the model for the response of a pixel is shown in section 2.2, highlighting the constituents of fixed pattern noise. It is the variations stemming from any of these components across pixel devices that need to be corrected in the calibration procedure. The effect of temperature on pixel mismatch is then discussed together with the need for a self-calibrating procedure. This is in response to constantly varying imaging conditions as well as the manual calibration scheme proposed by some commercial imagers. By physically acquiring or focusing a uniformly illuminated scene (like white paper) onto the logarithmic camera for every calibration, the FUGA 15RGB [53] calibration is hugely inconvenient and error prone. In section 2.3, a circuit usable for electronically self-calibrating a camera is presented and its

operation discussed. With this circuit, advantages like faster sensor calibration that can be programmed or automated within the camera operation sequence, is possible. Typical logarithmic pixel behaviour is then characterised, followed by an assessment of the impact of uncorrected FPN, demonstrated using a set of simulated pixel responses.

As mentioned in section 1.2.3, numerous calibration schemes have been suggested to reduce fixed pattern noise levels. These include injecting hot carriers [60] and feedback voltages from a comparator [14, 15] to compensate for threshold variations. Others have involved using reference currents [55, 64] or high currents [4], electronically stored uniform images [53] and modelling techniques [12, 13, 24]. However, these methods have all reported the residual FPN in the calibrated images differently and using separate error measures. Besides, no statistical or perceptual comparison of these results has been found in relevant literature.

Due to the non-perceptual nature of the above FPN error measures, a perceptually based contrast error measure is suggested in section 2.4, from where the results of previous offset correction methods are benchmarked - section 2.5. The results are unsatisfactory but a theoretical explanation follows thereby revealing the reasons behind this approach. Consequently, a simpler offset and gain procedure is proposed in section 2.6, with strong emphasis on simplicity and accuracy over the same wide range. In this routine, it is proposed that only two carefully selected data points are required to effectively minimise offset and gain mismatches - section 2.6.1. By comparing the results of this offset and gain correction method to Joseph's regressive double correction algorithm [12, 13] using the same set of simulated pixels, it is shown that similar accuracy can be obtained but much more conveniently than his computationally heavy approach.

2.2 Calibration & Temperature

Logarithmic sensors, like other types of imagers, are intended for use in a number of environments, some of which have varying temperatures. Industrial and automotive applications are a particular example where in the summer, temperatures could be above 30°C and drop to below 0°C in the winter. The way in which temperature affects the performance of silicon circuits and devices therefore becomes an important factor if it significantly alters typical operation. However, because different circuits are made of several devices whose relationship with temperature vary independently, the overall effects [74] are unpredictable. This means either circuits that are temperature independent have to be designed or the calibration process has to made as easy as possible to allow for recalibration whenever required. In the

former, Collins *et al.* [63] designed a programmable logarithmic pixel with a floating gate that was able to achieve a high pixel gain resulting in a large voltage swing hence reducing the temperature effects in the readout circuit. However, this solution did not cater for the relatively significant variations in temperature without significantly increasing design complexity and increasing costs. It is perhaps easier to devise a simple calibration system that can be conveniently repeated when a significant change in temperature or illumination condition occurs. The sensor must also be able to re-calibrate itself and avoid any inconvenient manual intervention.

However, before executing such a strategy, it is important to investigate precisely how changes in temperature affect fixed pattern noise in a logarithmic pixel and understand how each of the constituent sources of fixed pattern noise are related. This can be undertaken using the three parameter model for logarithmic pixel response y_{ij}, proposed by Joseph [12, 13], and re-written as;

$$y_{ij} = \acute{a}_j + \acute{b}_j \ln(x_i + \acute{c}_j) \tag{2.1}$$

where x_i represents the scaled photocurrent at an illumination i and $\acute{a}_j$, $\acute{b}_j$ and $\acute{c}_j$ represent the offset, sub-threshold slope and bias parameters respectively for pixel j. It is the variation of any of these parameters or any combination thereof, which results in fixed pattern noise. Offset mismatch is thought to arise mainly from the variations in the threshold voltages among pixel devices while the sub threshold slope mismatch primarily originates from readout gain variations.

The effect of temperature is also known to cause the threshold voltage to alter by $1-2$mV per degree [74] while a 10% change in thermal voltage, which is a component of the sub-threshold slope [12, 13], could result for a 30° C degree temperature variation at room temperature. This means a calibration performed at a particular temperature might not be effective at another. Consequently, the offset and gain parameters experience a proportional increase in temperature [75, 76]. On the contrary, the relationship between the dark leakage current and temperature is not a linear one since the leakage current doubles for every $10^{\circ}C$ rise in temperature [74]. For typical logarithmic sensors with a 50mV/decade sensitivity, the offset, gain and bias parameters generally show significant temperature dependencies, making the pixel responses equally dependent.

In his investigation, Joseph [75] confirmed the significance of this temperature influence on logarithmic sensors using the FUGA 15RGB sensor, recalibrated for every 5°C over a 20°c to 50°C temperature range, by manually focusing seven uniform scenes onto the imager at several illuminations. He obtained residual errors of 4.2 LSB and 2.0 LSB after employing two modelling methods - unconstrained offset

cancellation and temperature proxy models respectively [75]. Despite correcting FPN in this temperature range and over less than a three-decade range, the potential burden and inconvenience of the numerous readings and measurements, highlighted the need for any calibration scheme to be as simple as possible, as re-calibration might occur during sensor operation due to changes in operating temperature. An electronic self-calibrating camera is therefore seen as ideal in this approach. However, temperature dependent logarithmic calibration is not covered in the scope of this thesis although the clear advantage of an electronic calibration means over manual calibration, cannot be over emphasised.

2.3 Electronic Calibration Circuit

In any calibration, meeting the objective of getting a predictable response from a pre-defined input, relies on the ability to represent the input accurately. Photo illumination captured by an imager can be represented by a photocurrent generated electronically rather than from a photodiode. This was made possible using a circuit designed in the CADENCE software circuit simulator - figure 2.1 [77, 78]. It includes the basic logarithmic 3 transistor circuit structure adjacent to a differential readout and extra circuitry for the sole purpose of making calibration an easy, electronic procedure.

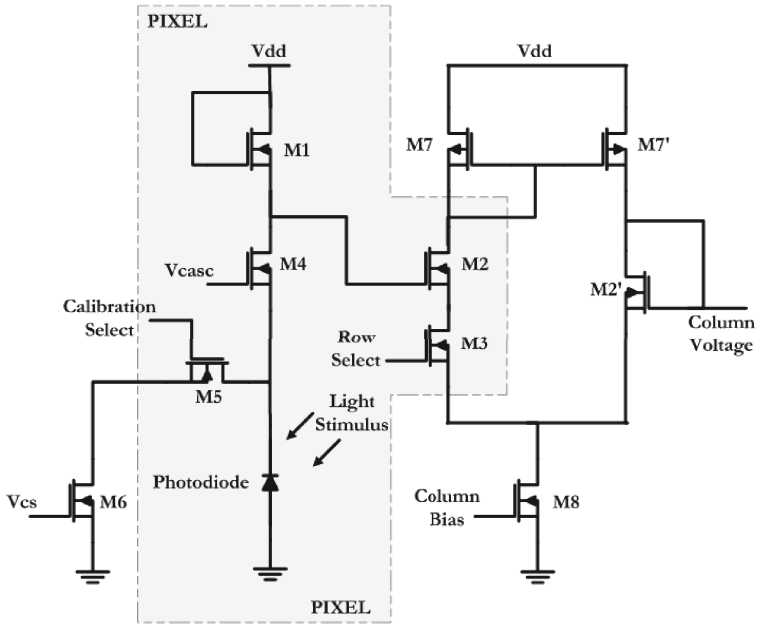


Figure 2.1: A logarithmic pixel circuit with a single stage differential amplifier readout as opposed to the traditional source follower. The extra circuitry $M4$ and $M6$ are to reduce leakages and for electronic calibration respectively.

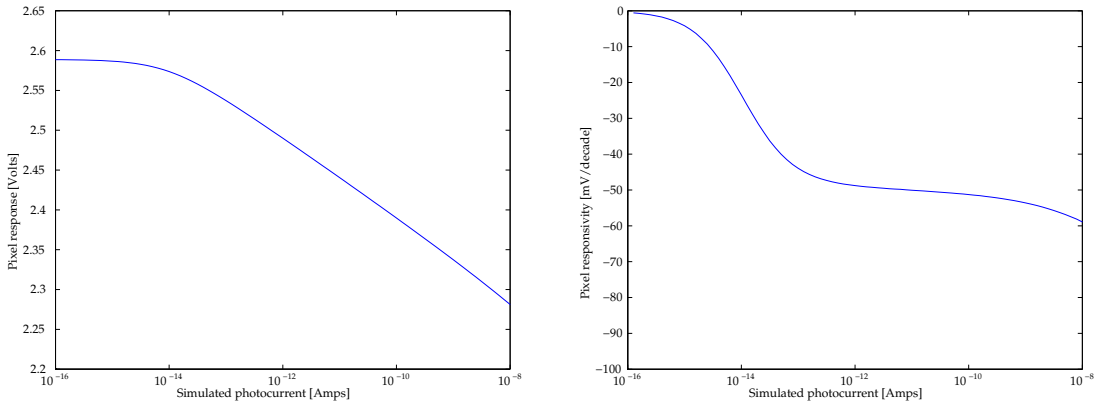
In this figure, transistors $M1$, $M2$ and $M3$, which are part of the conventional logarithmic pixel, are still in the circuit but with $M2$ and $M3$ becoming part of a new differential readout circuit, while the load transistor $M1$ stays unaltered. The reason for the change in readout circuit from the conventional source follower to the differential readout circuit, is that the latter has a higher gain[1] and hence the effect of the same distribution of gain variations in both readouts are lower in this case as compared to the former [77]. Considering that for every additional amplification stage, the gain of the source follower falls further from unity, this differential readout reveals a particular advantage when there is a second stage readout circuit containing a column select.

The calibration circuit also includes transistors $M4$, $M5$ and $M6$ adjacent to the photodiode. The logic of the circuit is that the current flowing through the photodiode can be emulated by the voltage controlled current source $M6$ when the aperture is closed. Any photocurrent can therefore be driven through the load transistor by carefully choosing the voltage V_{cs}, at the gate of the transistor $M6$ that is acting as a selectable calibration switch while keeping $M5$ pulsed high. Note that this configuration has five transistors since the voltage controlled current source is situated outside the pixel and along each column to reduce pixel-to-pixel variations. Thus disadvantages like high power consumption and reduced fill factor that some calibration schemes suffer are not introduced. Transistor $M4$ is expected to limit the voltage across the photodiode and therefore reduce the leakage effects especially in low illumination conditions. Dark pixel readings however still require the closure of the shutter to simulate a response when no stimulus is detectable.

The process of calibration begins with acquiring the responses of uniform scenes, from which the factors that lead to variations can be eliminated. Using the circuit in figure 2.1, these uniform scenes can now be easily simulated hence emulating a diffuse white scene, by setting all the voltage controlled current sources to the same gate voltage and enabling $M4$ and $M5$ before the pixel readings are taken. This then becomes a fast, reliable and very convenient way of uniform image acquisition - the first step of the calibration process.

The resulting pixel readings from the uniform electronic input contain offset and gain mismatch, however they do not incorporate low current leakage variations. This is because the photo detector, the photodiode that is the source of leakage variations is shorted out. On the other hand, this circuit has a big advantage in that it can be used to achieve the effect of different and widely varying photo illuminations spanning higher dynamic ranges than previously achieved. Figure 2.2(A) shows a typical logarithmic pixel response curve over a wide dynamic range obtained after the circuit of figure 2.1 had

[1]The gain of a differential current mirror is closer to unity compared to that of a source follower.



(A) (B)

Figure 2.2: The response of a single pixel simulated over a wide range of photocurrents shown in (A). The response was taken after two stages of the row and column differential readouts while (B) is the variation of the sensitivity of a single pixel over a wide range of photocurrents.

been simulated in CADENCE software. The horizontal axis represents source-drain photocurrents that were simulated when the voltage controlled current source at V_{cs} was varied in such a way that 10 data points were achieved per decade change in current. The vertical axis shows the pixel responses at these simulated photocurrents.

By taking the gradients between the data points, the sensitivity of the pixel can be obtained as in figure 2.2(B). It shows low absolute pixel responsivity at low illuminations which steadily rises to a fairly uniform value of approximately 50mV/decade. This behaviour of the pixel is expected and can be explained using the three-parameter model of equation (2.1). At the very low photocurrents of the photocurrent range, synonymous with low illumination, the leakage currents dominate the photo current and hence $\acute{c}_j \gg x_i$. The pixel response will therefore be dominated by leakages yielding very low dependence on simulated photocurrent and tending towards zero gain or sensitivity in the dark. This region is therefore non-ideal for logarithmic pixel operation since not only is the sensitivity close to zero or null but leakage current overshadows the influence of photo illumination. Nevertheless, responsivity grows gradually as the photocurrent values become comparable to the leakage current seen in the increasing absolute gain values. As the currents increase, the photocurrents slowly become a lot greater than these leakages therefore $x_i \gg \acute{c}_j$, leading to a gentle plateau; $\frac{\Delta y_{ij}}{\Delta \ln x_i} \approx \acute{b}_j$, covering about 4 decades. This is the region where the pixel transfer function is purely logarithmic and ideal device operation is desired for the full benefits of this type of pixel to be tapped.

2.3.1 Impact of FPN

The simulation of a logarithmic pixel with the accompanying electronic calibration circuit has shown a typical pixel response over a wide dynamic range. Using the model equation (2.1), valuable insight into the way pixel response changes with illumination was also gained. However, an understanding of the impact of the mismatch created by this logarithmic pixel configuration would be beneficial to understanding and therefore benchmarking fixed pattern noise over a wide dynamic range.

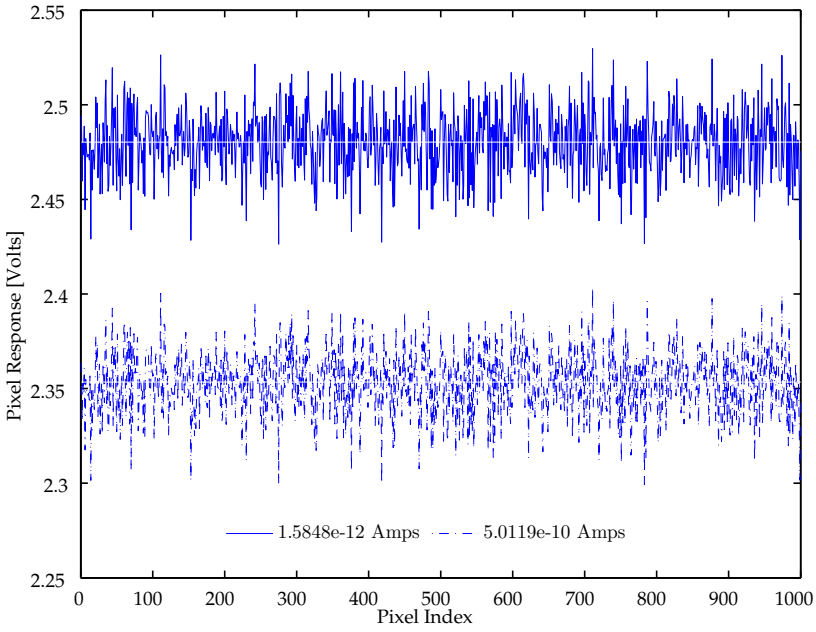


Figure 2.3: The readings of $1,000$ identical pixels (from a design perspective) at two photocurrents from a simulation, showing fixed pattern noise variations.

Using the Monte Carlo simulation[2] option in the CADENCE circuit simulator, the response of $1,000$ pixels was obtained by sweeping the current across the photodiode over 8 decades, in 0.1 decade intervals. In this scenario, process parameters like the transistor aspect ratio, gate capacitance per unit area and transistor threshold voltages were inherently altered in a controlled manner while determining the pixel response to a range of photo illuminations. Figure 2.3 shows the readings of the $1,000$ pixels at two simulated photocurrents which are 2.5 decades apart in the purely logarithmic current region. It is clear that the response of these pixels to a uniform current varies by about 100mV. This means that in the worst case, the fixed pattern noise in this region where the gain is approximately 50mV/decade - figure 2.2(B) - is equivalent to a photocurrent being mis-represented by a factor of about 2 decades. In most applications,

[2]Monte Carlo simulations are a statistical sampling technique that involve running numerous instances of a process, calculation or method under investigation, in order to assess the impact of different combinations of dependent parameters.

this is unacceptable considering that an image would be 100 times brighter or darker, if uncorrected, making captured images radically different from the ideal image. There is therefore little doubt that fixed pattern noise needs to be eliminated if not minimised to bearable levels. However, what constitutes the minimum allowable residual FPN level remains non-standard, as the varying results from different studies have shown.

2.4 New FPN Measure

Discussions in section 1.2.3 showed numerous ways of reducing fixed pattern noise from the response of logarithmic sensors. Unfortunately, these techniques have all used different statistical ways of reporting their level of success. Some studies have quantified residual FPN as a fraction of the entire voltage swing [55, 64] or as the root mean square of the residual FPN per decade voltage change [14, 15] while other have used the standard deviation in LSB units [12, 13].

Although all the above measures have a tangible meaning on a physical system level, they have no perceptual correlation to the quality of an image after FPN correction. Despite the fact that FPN correction to 2.5% of the entire voltage swing is better than 3.8% per decade, it is impossible to tell from a visual perspective, if this level of correction is merely adequate or an excellent result. In addition, there is no standard way of comparing the various approaches of dealing with FPN. To address this problem it is therefore necessary to devise a more perceptually meaningful error measure that has both physical and visual meaning. One way to approach this delicate subject is the consideration that most of the images we capture, store and manipulate are ultimately meant for use by human viewers. In most applications, these images are displayed to humans for decision-making, aesthetics, surveillance and control. For this reason, it seems appropriate to use the human visual system as the image quality standard.

Another reason to apply the HVS as an independent reference, is because it is a hugely complex imaging tool with a multitude of biological and operational functions scientists can leverage [3, 6, 9]. Among the human visual system's vast range of abilities, its capability to resolve contrast, termed just noticeable difference (JND), becomes of particular interest. Experiments have shown that the human visual system has the ability to resolve differences in illumination ΔI, against the background illumination I, to an accuracy of about 1% in high illumination [6, 9]. For instance *Weber's ratios of just noticeable difference*, confirms the HVS ability to discriminate an illumination change over a wide range, with the JND dropping from about 10% in low light levels to nearly 1% in high illumination [3] - figure 2.4. An absolute contrast, based on the just noticeable difference, can therefore be defined as a new and fair

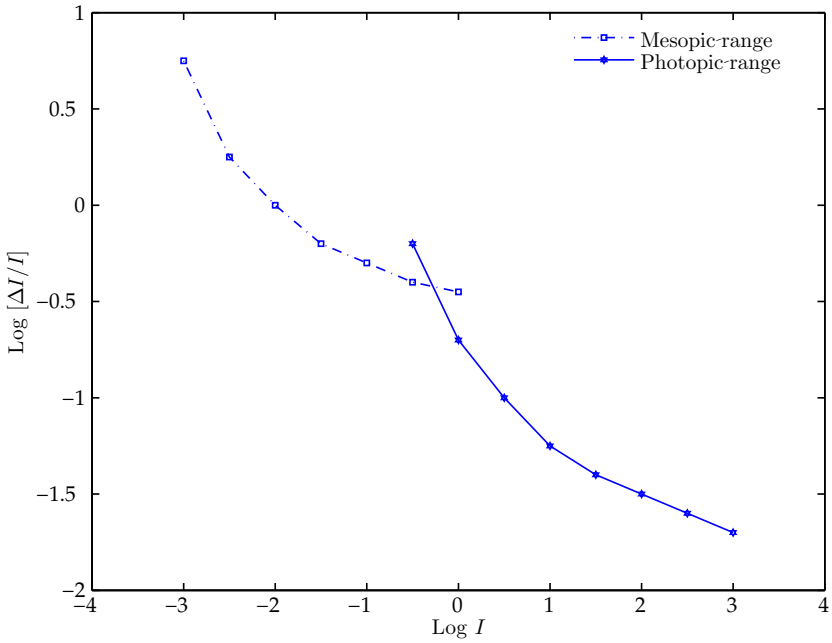


Figure 2.4: A figure showing the just noticeable difference as a function of intensity maginitude [3].

measure for the accurate reproduction of an illumination using,

$$\text{Absolute contrast} = \frac{\Delta I}{I} = \frac{I - I_{est}}{I} \tag{2.2}$$

where ΔI is the just noticeable difference in illumination and I_{est} is the estimate of the illumination after correction. This means that at high illumination, the eye is incapable of perceiving errors in misrepresented illuminations that are less 1% of the original illumination because the limit of difference perception has been reached. Although, this absolute contrast value increases as the illumination level drops, the use of a 1% contrast target as an error measure provides the most stringent or worst case metric usable for measuring the success of FPN correction but also ensures that its use can be a visually meaningful standard.

Applying this theory to the FPN problem and assuming each corrected pixel is trying to accurately reproduce the photo-illumination at a scene, the 1% absolute contrast error value becomes the best attempt at matching FPN error to the sensitivity of the human visual system. This appears to be a better target error measure than those suggested in section 1.2.3 considering that contrast error is not only inspired from an end-user perspective but can also be a common independent measure of performance in a field where many metrics have prevailed.

Using this contrast measure, a perceptually relevant way of quantifying the effect of uncorrected FPN

can be reached by calculating the ratio of the standard deviation of the simulated responses of the pixels at the different currents to the change in the output voltage caused by a 1% change in photo stimulus under the same input conditions. This measure of the visual rather than quantitative impact of fixed pattern noise is shown in figure 2.5. The results show that for most of the operating range of the pixels the fixed pattern noise is approximately 80 times larger than the signal created by a one percent contrast change in the photocurrent. This value increases at the low illuminations because of the lower sensitivity in this region - figure 2.2(B) - as well the increased significance of leakage variations which are negligible at higher currents. Again, this noise level is unacceptable for most domestic and commercial imaging applications.

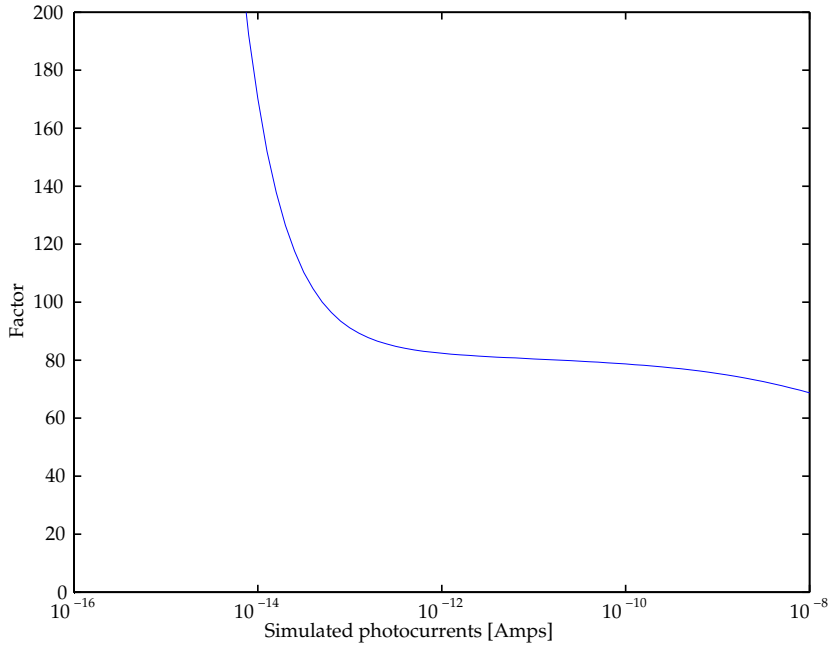


Figure 2.5: Impact of uncorrected fixed pattern noise. The y-axis is obtained by taking the ratio of the standard deviation of the $1,000$ simulated pixel responses to the voltage corresponding to a 1% contrast change at that photocurrent.

2.5 Previous Offset Correction

With this new error measure, it is important to compare the successes of the previous correction techniques against this metric. Table 2.1, shows the several methods used previously and their residual FPN values transformed into absolute contrast in the last column. These values were obtained by calculating the percentage contrast error the residual FPN voltage would represent in the purely logarithmic region of

the pixel in question. It is clear that, despite increasingly greater reduction of residual FPN from currently existing techniques, none produces FPN reduction to levels that are undetectable to the human eye. With an independent error measure in place, the challenge then remains to devise a correction method, that is simple, easy to operate and perceptually accurate over a wide dynamic range.

Authors *et. al*	Dynamic Range	Average Sensitivity	Residual FPN	Equivalent Contrast
TARGET	120dB	50mV/decade	0.43% of decade	1.00%
Loose [14, 15]	120dB	250mV/decade	3.80% of decade	$> 9.14\%$
Kavadias [55, 64]	120dB	50mV/decade	2.50% of swing	41.25%
Storm [79]	80dB	77mV/decade	4.00% of decade	9.65%
Joseph [12, 13]	40dB	$45 LSB/decade$	3.00 *LSB*	16.59%
	40dB	$45 LSB/decade$	2.00 *LSB*	10.78%

Table 2.1: Comparison of FPN correction results in different studies, shown in terms of absolute contrast. In the work performed by Joseph *et al.* [12, 13], offset and gain FPN correction was performed, with the better result obtained when both variations were minimised. Also, Loose *et al.* [14, 15] reported results as rms values rather than the peak to peak values hence a higher equivalent absolute contrast is expected.

However, one feature stands out from the discussion in section 1.2.3 and table 2.1. The techniques that have the highest contrast values are offset only correction schemes. Despite FPN constituting an offset, gain and leakage component, the offset coefficient is known to be the major contributor of FPN mismatch [18]. Consequently, the earlier attempts including others like the FUGA 15RGB [53], and even Lai *et al.*[3] [4], only concentrated on correcting this form of mismatch.

For instance, Kavadias *et al.* [55, 64] minimised threshold voltage variations by comparing pixel responses to those from a reference dark current after a reset operation. This allowed for subtraction after double sampling in between the frames. In contrast, Lai *et al.* [4] used a high current as a reference and correlated double sampling to obtain a similar result in his bipolar detector imager. On the other hand, the offset calibration procedures of the FUGA 15RGB logarithmic camera do not recommend any particular reference current since calibration is advised whenever the illumination of the scene is perceived to change substantially.

Figure 2.6 shows a graphical comparison of the results of some offset correction techniques against percentage contrast. By taking the ratio of the residual FPN after correction to the voltage change caused by a 1% contrast at a specific current, the error profiles suggested in dark current [4], high current [55, 64] and mid current FUGA 15RGB subtraction are highlighted. The errors heavily depend on the choice of calibration current with the contrast error increasing as the photocurrent deviates from the current at

[3]The results from this work were based on the a pixel that only partially operated logarithmically, so a fair assessment is difficult to make.

calibration. For instance when imaging a wide range scene, the dark and high current offset subtraction create a higher likelihood of error compared to when an offset calibration is performed at a photocurrent midway the dynamic range. In this case, approximately 4 decades are covered at below 10% contrast. Nevertheless these results and the converted error values of the last column of table 2.1, show that offset correction procedures are not sufficient to achieve the target accuracy over a high dynamic range.

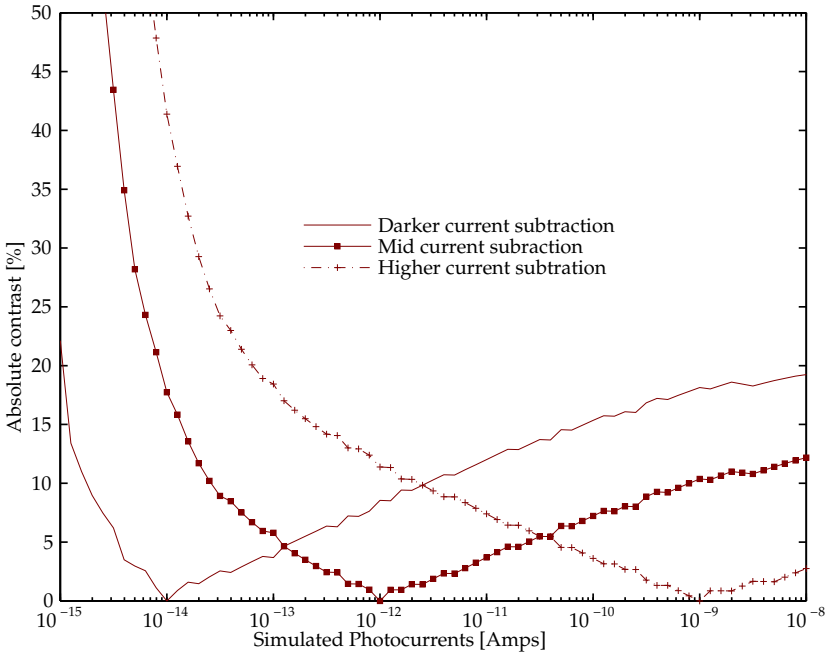


Figure 2.6: An illustration of the effect of the offset subtraction techniques on image quality over a wide dynamic range.

An explanation for the poor results from the correction of only offset variations can be derived from equation (2.1). Taking a logarithmic pixel's response y_{ij} to an illumination represented by a photocurrent x_i, j and i denoting the pixel index and photo illumination respectively, it is possible to understand the cause of the residual fixed pattern noise. From the original equation (2.1)

$$y_{ij} = \acute{a}_j + \acute{b}_j \ln(x_i + \acute{c}_j)$$

and considering the reduced sensitivity at low light conditions coupled with the fact that most photo illuminations result in photocurrents where the leakage effects are negligible, $x_i \gg \acute{c}_j$. The pixels at the scaled calibration current x_c and any other scaled current x_i will then have the responses y_{cj} and y_{ij},

respectively;

$$y_{ij} = \acute{a}_j + \acute{b}_j \ln(x_i) \tag{2.3}$$

$$y_{cj} = \acute{a}_j + \acute{b}_j \ln(x_c) \tag{2.4}$$

The result of any of the offset subtraction procedures described above is equivalent to taking the difference between these two equations to get a corrected voltage, y_{corr}, whichever calibration current x_c is taken, thus

$$y_{corr} = y_{ij} - y_{cj} = \acute{b}_j \ln(x_i/x_c) \tag{2.5}$$

From this equation, there will be no fixed pattern noise error in the offset corrected pixels at the current that is the same as the calibration current x_c. On the contrary, at any other photocurrent, the errors increase in proportion to the logarithm of the ratio of these currents provided $\acute{b}_j$ is constant for all pixels. This explains the rise in error. It is therefore conceivable that this rise will continue as the current varies unless a recalibration is done closer to the photocurrent of interest. In addition, equation (2.5) shows that the residual error is dependent on the choice of calibration current and the variations of the gain parameters of the pixels in the array. The amount of variations in $\acute{b}_j$ therefore also contribute to the amount of residual FPN. The total error in y_{corr} is a combination of the spread of the gain coefficient and the ratio x_i/x_c. This dependence on the spread of gain coefficients explains why, despite their other drawbacks, the other offset correction methods like hot carrier injection used by Ricquier *et al.* [60] and the voltage feedback control employed by Loose *et al.* [14,15] are still inadequate in meeting the 1% contrast target especially over a wide dynamic range.

Over a small dynamic range, some logarithmic sensors like the FUGA 15RGB [53] have endeavoured to reduce the effects of residual pixel gain mismatch by performing digital imaging techniques like median filtering. Figure 2.7 illustrates how the remnant outliers are removed more clearly. The picture on the left shows a fixed pattern noise infested indoor image that was captured using the FUGA 15RGB logarithmic camera and then offset corrected by subtraction. The image still has random artefacts attributed to the arbitrary residual gain variations within the pixels. However after median filtering using a 3×3 kernel, the image on the right is cleaner and lacks the black dots that stood out in the image on the left. This off chip post processing approach appears a viable solution for manufacturing processes with limited gain variations but produces far from ideal solution because of other limitations. For example, in

Figure 2.7: A low dynamic range indoor scene captured by the FUGA 15RGB camera and then offset corrected by subtraction of a stored image. On the right, the image is median filtered using a 3×3 kernel to spatially filter off outliers.

real-time or fast frame applications the computational effort for megapixel size resolutions (often exceeding 10 megapixels) is demanding and expensive to be embedded in typical low cost, portable and power sensitive applications, notwithstanding the slowing down of sensor performance. In addition, median filtering will not remove any FPN noise resulting from significant changes in the scene illumination from those of the reference calibration current x_c as experienced in HDR imaging. Above all, artefacts may result from the spatial filtering that are often insensitive to different scene features and tend to obscure clarity [3].

2.6 Offset & Gain Correction

At best, offset subtraction not only results in absolute contrast errors that are higher than the human visual system can detect over many decades, but also increases as the illumination at capture changes from the calibration current. Since correcting for only additive mismatch is inadequate, it is possible that if other forms of variations are minimised, the residual contrast error will be reduced.

Aware of the fact that the leakage currents only affect the success of fixed pattern noise correction in very low illumination, coupled with the inadequacy of equation (2.5), the possibility of improving image correction is investigated by performing gain as well as offset correction. It is expected that this will increase accuracy levels over a wide range.

Considering the j^{th} pixel in an array, whose response at scaled photocurrents x_1 and x_2 are y_{1j} and y_{2j}, the following equations can be written assuming leakage currents are negligible.

$$\begin{aligned} y_{1j} &= \acute{a}_j + \acute{b}_j \ln(x_1) \\ y_{2j} &= \acute{a}_j + \acute{b}_j \ln(x_2) \end{aligned}$$

The gain and offset pixel parameters $\acute{b}_j$ and $\acute{a}_j$ can then be obtained as

$$\acute{b}_j = \frac{y_{1j} - y_{2j}}{\ln(x_1/x_2)} \tag{2.6}$$

$$\acute{a}_j = y_{1j} - \acute{b}_j \ln(x_1) \tag{2.7}$$

Hence, at any current x_m, the pixel response y_{mj} would be corrected to y_{corr} by taking the ratio of the difference between the response and the offset parameter $\acute{a}_j$ to the gain parameter $\acute{b}_j$;

$$y_{corr} = \frac{y_{mj} - \acute{a}_j}{\acute{b}_j} = \ln(x_m) \tag{2.8}$$

Therefore by determining the offset and gain parameters of all the pixels in an array, using equations (2.6) - (2.7) and then storing them, any other image that induces photocurrents significantly higher than the dark leakages can be corrected for these two forms of fixed pattern noise. Besides, implementing this correction outside the pixel and leaving the fill-factor unaltered, this approach offers the possibility of simplicity and the convenience desired in a calibration scheme.

2.6.1 Point Selection Strategy

In the selection of data points from which the extraction of the offset and gain parameters is to occur, there are numerous photocurrents from which to choose along the illumination range. It is vital to acknowledge that among the multitude of possible points, the two data points should be chosen such that the possibility of achieving the 1% target contrast is maximised over a wide dynamic range. The two photocurrents should also be chosen to ensure that the two pixel responses are as different and far apart as possible so as to reduce the effect of numerical error. It is therefore logical to set the smaller of the two photocurrents at a current which is at least 2 decades larger than the value of the dark leakage current of the pixel i.e. at least 100 times larger than parameter $\acute{c}_j$ since our target contrast is 1%.

Another consideration is to reduce the effects of rounding error and random noise when estimating

the gain parameter from equation (2.6). A close inspection of the pixel response curve of figure 2.2(A) and the corresponding sensitivity curve in figure 2.2(B), reveals that the first point should be located just above 1×10^{-12}Amps, two decades from the bias current 1×10^{-14}Amps. The second photocurrent however, is not as clear since it is impossible to determine a simple relationship between the currents at which the pixel is characterised and the accuracy of the final model. This relationship has therefore been investigated empirically. To investigate the effect of the choice of the second data point, a single pixel from the $1,000$ simulated pixels was randomly chosen to determine the best pair of data points for calibration.

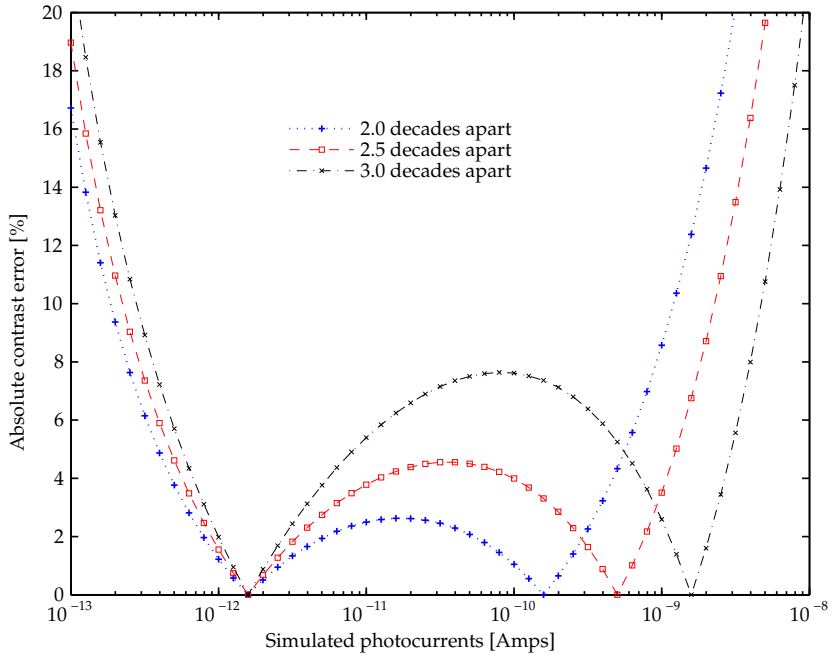


Figure 2.8: An illustration of the choice of the larger calibration photocurrent in relation to the smaller photocurrent thus highlighting the trade off between dynamic range for a specified accuracy.

Figure 2.8, illustrates the search for an ideal photocurrent to be used as a second calibration point for offset and gain calibration. For this pixel design, the second photocurrents were positioned 2.0, 2.5 and 3.0 decades from the smaller photocurrent point and the absolute contrast determined. The first significant point is that the contrast values are lower than the ones obtained from the offset correction procedures. Not only are these contrasts low, but also stay this way for a significant dynamic range. In fact, from about 4×10^{-13}Amps to approximately 1×10^{-9}Amps range, errors have been reduced to below 5% contrast when the two calibration points are 2.5 decades apart. Admittedly, this is still above the levels of the human eyes contrast discrimination but they represent a twofold decrease from the result obtained with offset correction only, albeit for an extra decade. Secondly, the figure has very distinct zero error

points at the photocurrents that were used for the calibration since these points satisfy the equations (2.7) and (2.6). It also shows very high and steep curves at the low current region, possibly due to the dominant effects of leakage current at those illuminations. Similarly, at the very high current end, the errors soar as the photocurrent increases.

More importantly, as the larger calibration point is driven further from the smaller photocurrent, there is an increase in absolute contrast. There is a trade-off between contrast and dynamic range while making the choice of the ideal photocurrents. This becomes even more complicated when you consider the fact that every fabrication and design process is characterised by unique variations (leakage currents and others). Hence, no universal calibration photocurrents can be defined. It then becomes easier for every manufacturing process to characterise its own pixels, taking into consideration the need to balance the desired image quality and required dynamic range. In applications where dynamic range is more important than accuracy, the two photocurrents can be chosen so that they are as far apart as possible, whereas if accuracy is a more vital criteria it might be desirable for them to be closer to achieve higher accuracy.

2.6.2 Results

With an optimal data point selection strategy, offset and gain correction presents an opportunity to achieve better accuracy and dynamic range in an easy manner using the calibration circuit in figure 2.1. Using the same data simulated by this circuit, it is also possible to compare the results from the suggested offset and gain correction procedure using simple parameter extraction to other results previously obtained. Although the correction of additive and gain variations has been attempted by Ponchut *et al.* [24] and Joseph *et al.* [12, 13], Ponchut unlike the latter, performed this correction within a software application and only mentioned this fact without offering any results.

However, in Joseph's mean sum square optimisation procedure, five uniform images were used to correct offset and gain mismatch by calculating the sum square error between the pixel responses to those uniform images and their estimated responses. He was therefore correcting all the responses to a mean response by minimising this difference using multiple images. This routine is heavily computational, the strain increasing with the number of uniform calibration images used, the resolution of the sensor to be calibrated and in real time video applications where high frame rates are required.

This optimisation routine was replicated as well as the simple offset and gain correction method on the same set of $1,000$ simulated pixels over the same current range. Since there were 10 data points per decade, there was a significantly higher number of computations in the optimisation than those used by

Joseph. To avoid excessive errors arising from the low illumination, only photocurrents higher than at least 10 times the leakage were used in the optimisation.

For comparison, the calibration photocurrents 1.26×10^{-12}Amps and 1.26×10^{-9}Amps selected for the simple offset and gain correction, were those that offered the widest dynamic range as opposed to the best accuracy. The extracted offset and gain parameters were then compared with those obtained from the simple parameter extraction. Table 2.2 shows a very strong similarity between the values of both parameters as well as their variation. Note that the gain co-efficient is shown as a voltage change on a natural logarithm scale or Neper[4]

Parameter (Units)	Offset a_j (Volts)		Gain, b_j (Volts/Neper)	
Measure	Mean	Std	Mean	Std
MSE optimisation	1.8779	0.017814	-2.2155×10^{-2}	4.7557×10^{-5}
Simple Extraction	1.8778	0.017806	-2.2164×10^{-2}	4.7375×10^{-5}

Table 2.2: A comparison of the extracted values of the offset and gain parameters determined from the optimisation technique and the simple parameter extraction respectively.

For the contrast errors to be determined, the estimated photocurrents were also calculated using these extracted parameters. Figure 2.9 shows the result from both methods; the curves exhibiting a significant amount of similarity in the shape of the curves, dynamic range covered and contrast accuracy. Both approaches neglect leakage currents so high errors are expected in the low regions while the reason for the error in the higher currents forms the main discussion in the proceeding chapter. In spite of this, the MSE optimisation yields better results than the 10.78% obtained in the last column of table 2.1, possibly because the author's results were compromised by quantisation noise and overall data quality [12, 13].

Nevertheless, the simple extraction procedure produces remarkable results considering that only two data points are used compared to the multitude used in the optimisation procedure to obtain the same level of accuracy. Thus, despite achieving similar dynamic range and contrast error results, the simple extraction and correction procedure offers the major advantages of simplicity, time saving and the ease of calibration with the proposed electronic circuit.

2.7 Summary

In this chapter, the model for the response of a logarithmic pixel was used to show the constituents of fixed pattern noise variations in the pixels of a sensor. Apart from a spatial variation, these FPN components

[4] A Neper is a dimensionless unit of ratios, named after John Napier (who invented logarithms), and is used to represent a gain, loss or a change on a natural logarithm scale.

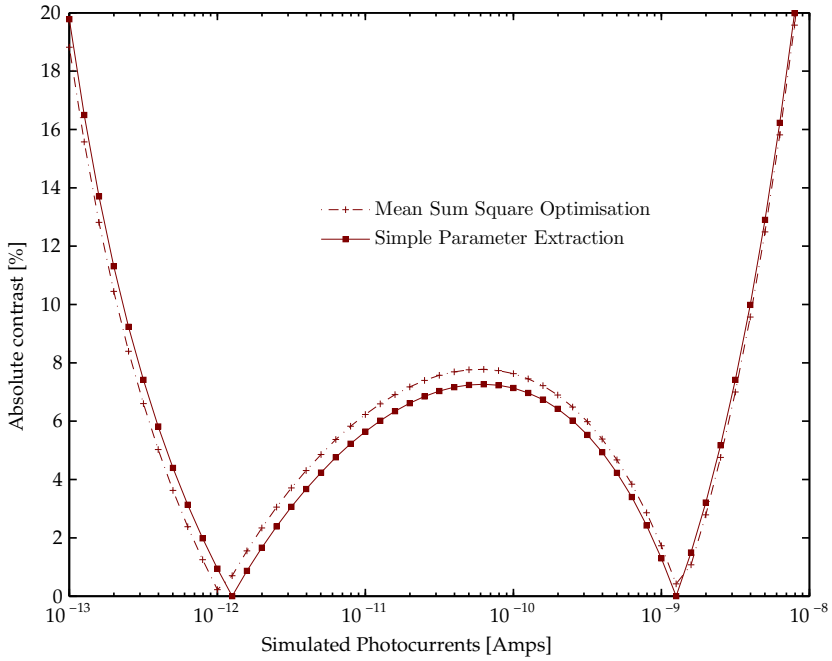


Figure 2.9: A comparison of the accuracy of the offset and gain correction using the mean sum square optimisation technique as opposed to the simple parameter extraction method.

showed a significant temperature dependence hence requiring a simple and repeatable calibration process. Unlike some earlier calibration schemes that were manual or had other drawbacks of large size (reduced fill factor) or high power consumption, a self-calibrating procedure was deemed necessary.

The structure of an pixel circuit that can be electronically calibrated was then presented and its operation discussed. Using data from the simulated pixel circuit designed in the CADENCE circuit simulator, properties of a typical pixel from a set of simulated $1,000$ pixels were characterised as well as the impact of uncorrected fixed pattern noise.

However, previous fixed pattern noise correction routines reported their residual FPN errors in different and varying measures, none of which gave an indication of the perceptual quality of the resulting image. A new perceptual error measure was introduced in section 2.4, that offered an independent way of establishing residual image quality after FPN correction. It was a biologically inspired metric based on the performance of the human visual system's ability to discern an illuminance.

With the set of simulated pixels, previous offset correction methods were then benchmarked to this target contrast error measure - section 2.5. The disappointing results were explained using the model equation introduced in section 2.2 and residual gain variations. Although, image-processing techniques like median filtering had been suggested to reduce the remaining noise, this is deemed inadequate for most

real-time, high dynamic range applications since it was very computational and introduced undesirable artefacts.

Simple offset and gain correction was then proposed to alleviate gain mismatch on the assumption that leakage currents were negligible since they only significantly affected image quality in low illumination. Using only two data points, the extraction of the offset and gain parameters was shown and the correction for these forms of mismatch demonstrated at any photocurrent. However, because of changing sensitivity along the photocurrent range and the need to attain the target contrast, the choice of these two data points had to be carefully selected. At the same time, numerical errors had to be minimised while covering a wide dynamic range. Consequently, the photocurrents had to be as different as possible with the smaller current being at least 100 times larger than the c_j parameter. In empirically determining the larger photocurrent, it was found that there was a trade-off between dynamic range and contrast accuracy. Although this made the choice of a definite point impossible, this trade-off was thought suitable to the range of applications whose prime specification varied between image quality and dynamic range.

Finally, the simple offset and gain correction was then benchmarked along the best offset and gain correction method - the mean sum square optimisation method. It was found that the two methods were similar over the same range and despite the model error not improving at high currents, the simple offset and gain method was simpler, more convenient, and less computational making it a more viable choice for fixed pattern noise correction.

Chapter 3

Model Derivation & Parameter Extraction

3.1 Introduction

It has been shown in the previous chapter that both offset and gain mismatch need to be corrected from logarithmic pixel responses. Not only can this occur in a simpler routine than previously employed, but also electronically to compensate for temperature effects. By employing only two carefully selected data points, the results from simulated pixel responses in figure 2.8 show that on average the incident photocurrent can be estimated to an accuracy better than 8% for approximately 4 decades. In fact, the contrast accuracy is marginally better than the non-linear optimisation technique used by Joseph *et al.* [12, 13], over the same range.

However, both methods result in gradually decreased contrast accuracy as the illumination varies from the 4 decade mid-current range. Expectedly, the region of reduced accuracy at the very low photocurrents corresponds to the steep change in sensitivity in figure 2.2(A). The explanation for this trend is that the simplified two parameter form of the model used in the reduction of the offset and gain variations is no longer valid. This is attributed to the fact that, unlike the simplified form, the three parameter model predicts pixel response in regions where leakage currents compare to the actual photocurrents. In other words, by neglecting leakage currents in the simplified form of the model, not only does the accuracy fail at low illumination, but also a limit is placed on the usable dynamic range. None the less, the change in contrast at the high illumination cannot be explained this way.

The validity of the three parameter model is based on the assumption that the load transistor is biased in sub-threshold. Although, this assumption enables accurate predictions of logarithmic pixel response over a dynamic range greater than those in conventional cameras, there seems to be a vulnerability over wide dynamic range scenarios especially at high currents. In this chapter, the higher contrast values at the high end of the photocurrent range are investigated. The study is based on the assumption that the simpler form of the model used in chapter 2, and indeed the three parameter model are, all part of a more generic model that completely predicts the pixel response over all stimuli. If this model exists then its characterisation should lead to a better understanding of high dynamic range logarithmic pixel behaviour and provide an invaluable means of correcting pixel response mismatch. Undoubtedly, this will explain the dynamic range limitations of the present model and, quite importantly, reduce the degradations in accuracy over the entire range.

Using transistor theory, a more complex four parameter model is developed in section 3.3 thereby allowing complete characterisation of a logarithmic pixel's response. To utilise this model, its coefficients need to be extracted for each pixel. By using a regressive non-linear optimisation technique, and the response of a simulated logarithmic pixel at numerous illuminations, the four parameters are obtained in section 3.3.1. Alternatively, using only five data points, a model fit to the same simulated pixel response is made in section 3.3.2, resulting in an accurate fit over the entire dynamic range. With an improvement in the amount of computational time and effort over the optimisation routine, the latter approach is favoured in the development of an FPN correction procedure. However, with the difficulties of obtaining in-pixel currents coupled with a greater understanding of pixel behaviour, the two parameter model is chosen for its simplicity and dynamic range coverage. Finally, practical parameter extraction procedures, usable in sensor calibration, are developed in sections 3.5.1 and 3.5.2 along with a proposed calibration sequence for practical correction in section 3.5.3.

3.2 Three Parameter Model Failure

In neglecting the effects of dark currents during offset and gain correction in section 2.6, the two parameter model was sufficient to accurately reproduce pixel functionality for currents far greater than the dark currents, $x_i \gg \acute{c}_j$. It however failed when the photo current x_i was comparable to $\acute{c}_j$. In fact, in very low illumination, the dark currents dominate, leading to negligible photo sensitivity. The two parameter model is therefore incomplete and gives rise to the high contrast values at the low current end of the photo illumination range.

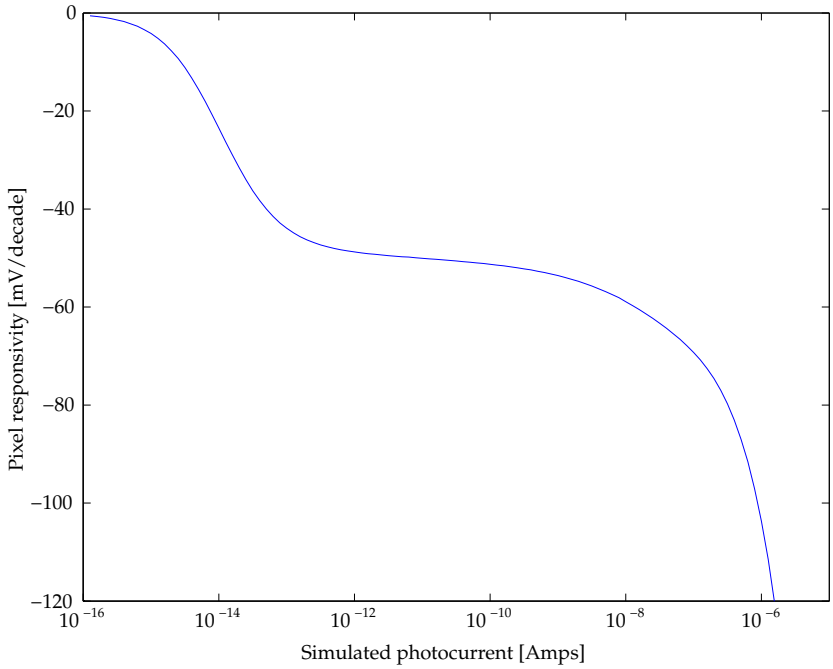


Figure 3.1: Curve showing how the sensitivity of a single pixel varies over a wide range of photocurrents.

Nevertheless, in the high current region, there seems no valid explanation for the rising error, despite the photocurrents still being far greater than the leakage effects. Since the three parameter model always assumes weak inversion operation of the load transistor, it is possible that as the photo illumination is increased, the load transistor is driven out of weak inversion [80].

By simulating photocurrents to values higher than 10nAmp, it is possible to study the corresponding change in logarithmic pixel behaviour. Figure 3.1 shows the pixel sensitivity as higher currents are passed through the device. There is a steep rise in absolute sensitivity associated with a possible change of device operational behaviour. It is probable that with this rise in current, the transfer function is no longer logarithmic and hence falls short of predicting the accurate pixel response in these conditions. The use of the simple two parameter model therefore leads to significant error as deriving the estimated photo illumination would be based on a false premise.

3.3 HDR Modelling

To completely model a logarithmic pixel over a high dynamic range, one requires a model that avoids the assumption of the operating regime associated with the level of illumination captured by the sensor. In other words, the model should predict the pixel output as the load transistor smoothly moves from weak

to moderate inversion.

Transistor device theory presents a transfer function offering an expression for the current-voltage characteristics of a MOS transistor that represents all the operating regions of the transistor from weak to strong inversion - the Enz-Krummenacher-Vittoz (EKV) MOS transistor model [81]. In this model, and referring back to figure 2.1, the drain-source current, I_{ds}, at a gate-source bias voltage V_{gs}, flowing in transistor $M1$ of width W and length L with an effective capacitance per unit area C'_{ox} can be written in the form [57]

$$I_{ds} = \frac{W}{L}\mu C'_{ox} 2n\phi_t^2 \left[\ln\left(1+\exp((V_{gs}-V'_t)/2n\phi_t)\right)\right]^2 \tag{3.1}$$

where μ is the effective carrier surface mobility, V'_t is the transistor's effective threshold voltage, ϕ_t is the thermal voltage, and n is the sub-threshold slope parameter. Within the pixel, the drain-source current of the load transistor $M1$ is the sum of the photocurrent I_p and the dark current I_s. Thus,

$$I_{ds} = I_p + I_s \tag{3.2}$$

The pixel output is measured at the source of the load transistor and since the gate of $M1$ is connected to V_{dd} - figure 2.1 - the predicted pixel response becomes

$$V_{out} = V_{dd} - V'_t + 2n\phi_t \ln\left[\exp\left(L(I_p+I_s)/(W\mu C'_{ox}2n\phi_t^2)\right)^{1/2} - 1\right] \tag{3.3}$$

This equation can be re-written with only four parameters in the form below [58],

$$y_{ij} = a_j + b_j \ln(\exp(\sqrt{c_j + d_j x_i}) - 1) \tag{3.4}$$

with the terms;

$$a_j = V_{dd} - V'_t \tag{3.5}$$

$$b_j = 2n\phi_t \tag{3.6}$$

$$c_j = I_s \frac{L}{W\mu C'_{ox} 2n\phi_t^2} \tag{3.7}$$

$$d_j = \frac{L}{W\mu C'_{ox} 2n\phi_t^2} \tag{3.8}$$

employed for simplicity. The coefficient a_j represents an additive contribution or offset in the output

voltage, b_j represents the pixel gain, c_j represents the effects of the dark current in the pixel and x_i is the photocurrent I_p. Comparing the equation for d_j to the simple model of a transistor operating in saturation, with a source-gate voltage greater than the threshold voltage,

$$I_{ds} = \frac{W}{2\,L}\mu C'_{ox}(V_{gs} - V'_t)^2 \tag{3.9}$$

suggests that d_j is the inverse of the current that will flow through the load transistor when it is biased just above its threshold voltage by a gate-source voltage of

$$V_{gs} = V'_t + 2\phi_t n^{1/2} \tag{3.10}$$

Since n is approximately unity [57] and at room-temperature $\phi_t \approx 26$mV [56], this is a gate-source voltage of approximately 52mV larger than the threshold voltage of the device, clearly surpassing the condition for weak inversion ($V_{gs} < V'_t$).

The first test to check the validity of this model over the high dynamic range is to confirm that, in the low current regime where the load transistor is in sub-threshold, the four parameter model reduces to the three parameter model. In this region, the current flowing through the load transistor will be significantly less than the current that will flow through this device when it is biased just above its threshold voltage. Here, the $d_j x_i$ term in $y_{ij} = a_j + b_j \ln(\exp(\sqrt{c_j + d_j x_i}) - 1)$, is less than unity. Similarly, the current flowing through a device biased near its threshold voltage is much larger than the dark current through the pixel and hence c_j is much less than unity. Under these conditions and taking the Taylor expansion [58],

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \tag{3.11}$$

$$\exp(\sqrt{c_j + d_j x_i}) \approx 1 + (\sqrt{c_j + d_j x_i}) \tag{3.12}$$

and therefore the four parameter model equation 3.4 can be re-written as

$$\begin{aligned} y_{ij} &= a_j + \frac{b_j}{2}\ln(c_j + d_j x_i) && (3.13)\\ &= \left(a_j + \frac{b_j}{2}\ln(d_j)\right) + \frac{b_j}{2}\ln\left(x_i + \frac{c_j}{d_j}\right) && (3.14)\\ &= \acute{a}_j + \acute{b}_j \ln(x_i + \acute{c}_j) && (3.15) \end{aligned}$$

As expected, this is the same form as the three parameter equation proposed by Joseph and Collins [12,

13], making the four parameter equation a valid model for this region.

However with the onset of moderate inversion, a square law relationship between the drain source current and the gate source voltage is realised, equation (3.9). Figure 3.2 shows the inversion current at which the transition occurs in relation to the rest of the load transistor operating regimes. This current is important in signifying the end of the logarithmic response and can be determined using the reciprocal of equation (3.8)

$$\frac{1}{d} = \frac{W}{L}\mu C'_{ox} 2n\phi_t^2 \tag{3.16}$$

where all the terms are design or process dependent parameters, and the sub-threshold slope n is acquired from the sub-threshold current-voltage equation

$$I_{ds} = I_o \exp\left(\frac{V_{gs}}{nkT/q}\right) \tag{3.17}$$

$$\ln\left(\frac{I_{ds}}{I_o}\right) = \frac{V_{gs}}{nkT/q} \tag{3.18}$$

$$\frac{d\ln I'_{ds}}{dV_{gs}} = \frac{1}{nkT/q} \tag{3.19}$$

$$\therefore \quad n = \frac{dV_{gs}}{d\ln I'_{ds}}\left(\frac{1}{kT/q}\right) \tag{3.20}$$

I_o being the current that flows when the transistor gate source voltage is null and I'_{ds} is the scaled drain source current through the load transistor $M1$.

For pixels designed using the CADENCE software simulator described in figure 3.1, a pixel gain of 50mV/decade would be equivalent to 21.5mV/Neper. Considering that at room temperature the thermal voltage, kT/q is approximately 26mV, equation (3.20) gives a sub-threshold slope of 0.87. With the design load transistor length and width of 0.6μm and 1.0μm respectively, a process capacitance per unit area of 4.6×10^{-3}F/m^2, a carrier surface mobility μ of 3.85×10^{-2}m^2/Vs, the current at which the pure logarithmic model should fail is 2.45×10^{-7}Amps, equation (3.16) and shown on figure 3.2. This current corresponds to the knee point of the sensitivity curve of figure 3.1, at which a sharp increase in gain is noted. Therefore it is most probable that in the consideration of the purely logarithmic two parameter model, the moderate inversion current places a limitation on the choice of the second calibration point that was empirically determined in section 2.6.1.

In light of the above, the four parameter model confirms the loss of the logarithmic function at high photocurrents and explains the high contrast errors at high illumination when the simpler models are used. Consequently, the four parameter model shows immense promise in fitting the response of a logarithmic

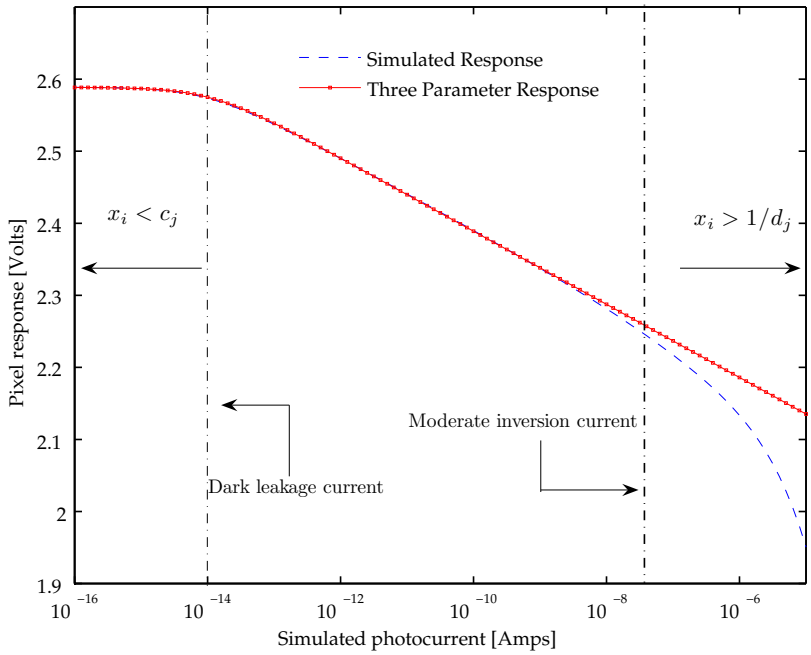


Figure 3.2: Illustration of the logarithmic pixel operating regions. The dark leakage current is at 1.24×10^{-14}Amps while the departure from weak inversion current occurs at 2.45×10^{-7}Amps.

pixel at all currents and the subsequent correction of FPN mismatch.

3.3.1 Non-Linear Optimisation

The four parameter model has been proposed as a model that will hold at both low and high photocurrents. To test its validity, the response of a typical pixel has to be compared with the model response at the same illumination. However to make the comparison the parameters of the new model have to be extracted before the model response is determined.

Since the relationship between the response y_{ij} of a logarithmic pixel j, at an illumination i, expressed in the four parameter model, is a non-linear one, the extraction of the co-efficients a_j, b_j, c_j and d_j can be undertaken using non-linear optimisation and/or regressive techniques. Typically, these methods require many data points to reduce the mean sum square and are always iterative, only terminating when a preset condition is met. These conditions may vary from a desirable tolerance level, a set maximum number of iterations, a search range to even a particular number of function evaluations.

By employing the *fminbnd* function minimisation routine from the MATLAB optimisation toolbox, the four parameters of the model could be calculated. First, by iteratively searching for the non-linear co-efficients within a user specified range, d_j and c_j were determined before linearly regressing for a_j

and b_j. During this optimised search, the iterated values of d_j and c_j that resulted in the minimum sum square error between the simulated and modelled response were used in determining the a and b values. Thus, using the CADENCE simulated pixel response, this optimisation procedure resulted in the extracted values shown in table 3.1.

Parameter	Offset, a_j	Gain, b_j	Leakage factor, c_j	Inverse current, d_j
Units	$Volts$	$Volts/Neper$	$--$	$Amps^{-1}$
Value	2.2089	-0.0447	4.3552×10^{-8}	3.3370×10^{6}

Table 3.1: Extracted parameters from a pixel using the function minimisation routine. In this procedure 111 data points were used ranging from 1×10^{-16}Amps to 1×10^{-5}Amps.

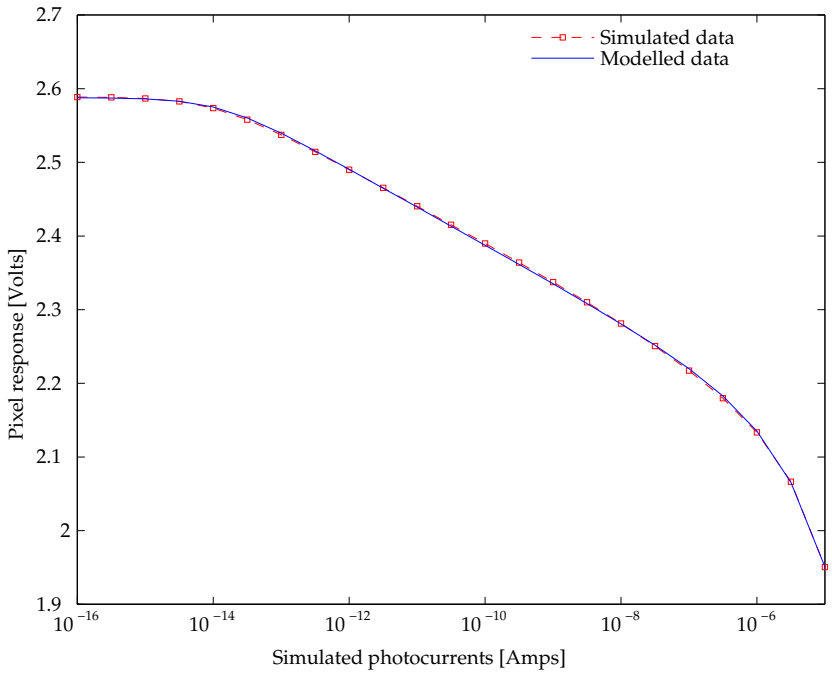


Figure 3.3: A fit of a four parameter modelled pixel response after the parameters were determined using a regressive optimisation technique.

With these coefficients, the modelled response was compared to the simulated response in figure 3.3. The figure shows a close fit of the of the four parameter model response and the simulated response over the entire 11 decade photocurrent range. From figure 3.3, it is possible to say that all the regions of transistor operation, are closely modelled with all error between the two responses falling below 3.5mV over the entire region - figure 3.4. Considering that the precision of the simulated data was 0.1mV, the residual error called model error, arises from the inability of the simulator to accurately model leakage current, which is actually the ratio of the c_j and d_j parameters being extracted. The close fit however, confirms the versatile nature of the four parameter model and makes it ideal for pixel modelling over

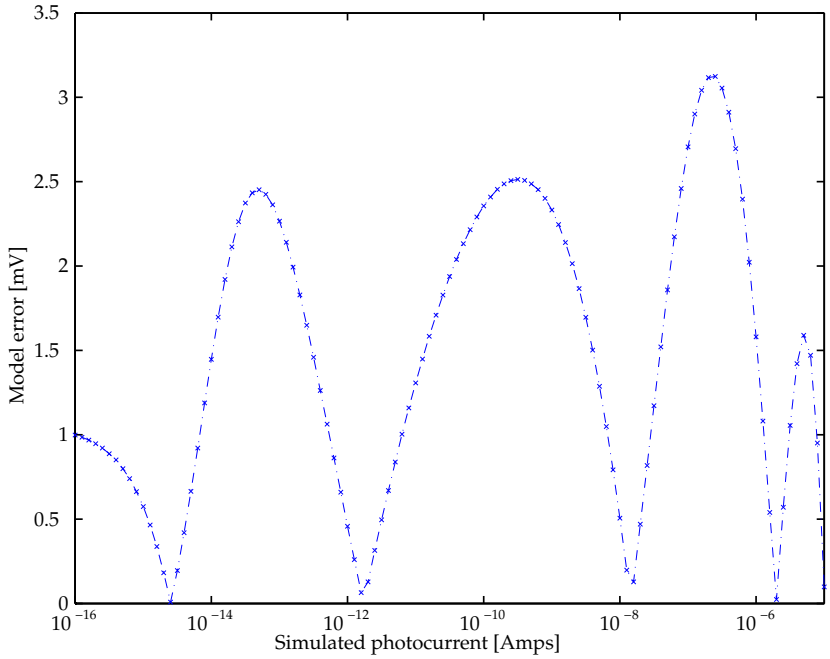


Figure 3.4: The model error plot, obtained as a result of using the parameter optimisation technique.

large photo illumination ranges.

However, a potential drawback of the four parameter model is the complexity of the parameter extraction. In fact, despite the non-linear procedure achieving a very good fit over the entire photocurrent range, the parameter extraction procedure is very computationally demanding. If this method were to be applied in an image correction or calibration scheme for real-time applications, it would be costly since it would impart a heavy strain on the computational resources and be immensely time consuming for large arrays. For instance, when this procedure was applied to $1,000$ pixel responses, which were simulated at currents 0.1 decades apart over 11 decades, it took an astonishing 7.08 minutes. This was performed on a PENTIUM IV Intel processor based PC, running no other applications. Like all numerical analysis methods, the processing of time can be reduced by limiting the number of iterations and permitted tolerance level but at a balance against achievable accuracy and desired cost. Nevertheless, it is still necessary that a simpler method of extracting the coefficients from this robust four parameter model be developed.

3.3.2 Simpler Parameter Extraction

Parameter extraction can be made easier if the four parameter model is split into specific regions of transistor operation, rather than as a whole. In these regions, appropriate assumptions can be made about

the specific region of operation for which the four parameter model reduces to a simpler form but only for that region. It is then, that the parameters can more easily be determined.

In the first instance, the four parameter model can be simplified by assuming that, although the current in the load transistor is small enough to ensure that the load transistor is operating in weak inversion, the contribution of the dark current is still negligible. The model for the response of the pixel then reduces from equation (3.13) to;

$$y_{ij} = a_j + \frac{b_j}{2}\ln(d_j x_i) \tag{3.21}$$

Under these conditions, the pixel output voltage is proportional to the logarithm of the current in the load transistor. If the responses y_1 and y_2 of a pixel at two photocurrents x_1 and x_2 in this operating region are known, then the gain parameter can be determined by taking the differences of the two responses giving;

$$b_j = \frac{2(y_{1j} - y_{2j})}{\ln(x_1/x_2)} \tag{3.22}$$

To determine the next parameter, consider the situation when the current through the load transistor is large enough so that $d_j x_i$ in equation (3.4), is significantly greater than one; then the effects of the dark current will be negligible leading to;

$$\exp(\sqrt{c_j + d_j x_i}) - 1 \approx \exp(\sqrt{c_j + d_j x_i}) \tag{3.23}$$

$$\approx \exp(\sqrt{d_j x_i}) \tag{3.24}$$

this means that the pixel output will be

$$y_{ij} = a_j + b_j\sqrt{d_j x_i} \tag{3.25}$$

In this high current condition, the output voltage is proportional to the square root of the current flowing through the pixel, equation (3.9). More importantly, it is then possible to determine the value of parameter d_j from the value of b_j, equation (3.21) and the responses of the pixel to two high photocurrents x_3 and x_4

$$d_j = \left(\frac{y_{3j} - y_{4j}}{b_j(\sqrt{x_3} - \sqrt{x_4})}\right)^2 \tag{3.26}$$

where y_{3j} and y_{4j} are the responses of the pixel to high currents x_3 and x_4 respectively. Once the values of b_j and d_j are known it is then possible to determine the value of parameter a_j from equation (3.21) using either x_1 or x_2,

$$a_j = y_{1j} - \frac{b_j}{2}\ln(d_j x_1) \tag{3.27}$$

The value of the final parameter can be determined under a special condition. When $x_i = 0$, a dark reading, which is a lot easier to achieve, unlike the other conditions, only involves closing off a shutter and taking a reading. Here, only the dark leakage current flows through the load transistor. This is the dark response and in this operating condition $c_j \gg d_j x_i$ therefore from equation (3.13)

$$y_{ij} = a_j + \frac{b_j}{2}\ln(c_j) \tag{3.28}$$

the value of parameter c_j can then be determined from the response of the pixel under these conditions y_{5j} using the equation

$$c_j = \exp\left[2(y_{5j} - a_j)/b_j\right] \tag{3.29}$$

Parameter	Offset, a_j	Gain, b_j	Leakage factor, c_j	Inverse current, d_j
Units	$Volts$	$Volts/Neper$	$--$	Amps^{-1}
Value	2.2147	-0.0441	4.2579×10^{-8}	3.6844×10^{6}

Table 3.2: Extracted parameters from the simpler parameter extraction routine using only 5 data points and over the same 11 decade photocurrent.

Thus, using the response of the pixel to five specific photocurrents, it is possible to estimate the four parameters required to model the pixel. Using the same CADENCE simulation as that in the previous section, this extraction routine was tested. The currents and responses at two mid range photocurrents, a dark response and two high photocurrents were used to obtain the values presented in table 3.2.

The parameter values are very similar to those in table 3.1. There is also hardly any noticeable difference when the modelled pixel response is compared to the simulated results, figure 3.5. In fact, the model seems to fit all the model regions as well as the optimisation procedure, except for the transition from weak to moderate inversion. This is highlighted by the higher model error values peaking at about 6mV, at the transition region of figure 3.6. Using the five data points, this procedure is as good everywhere else but only loses accuracy at the high current end with an approximate doubling of model error. It is

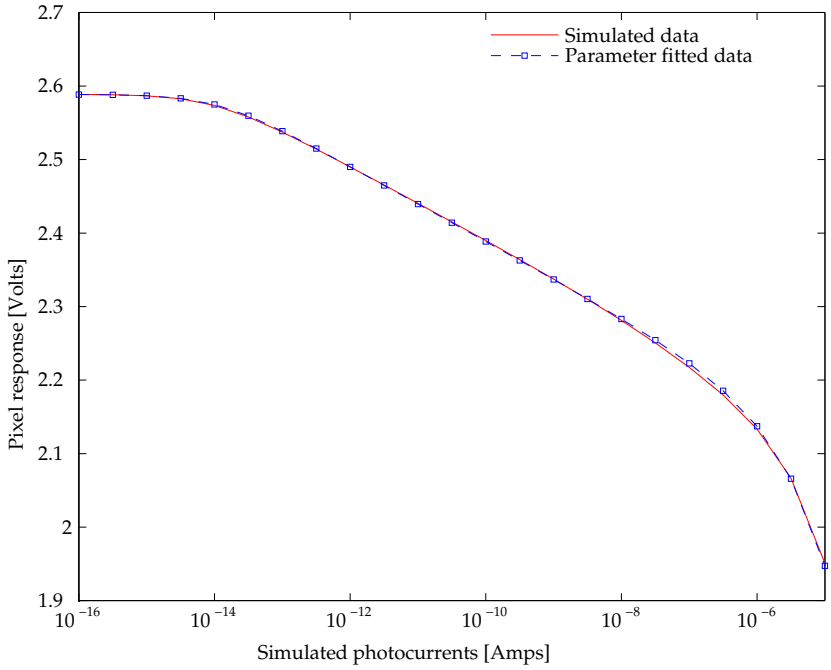


Figure 3.5: Parameter model fitting using five data sets from the response of a simulated pixel response.

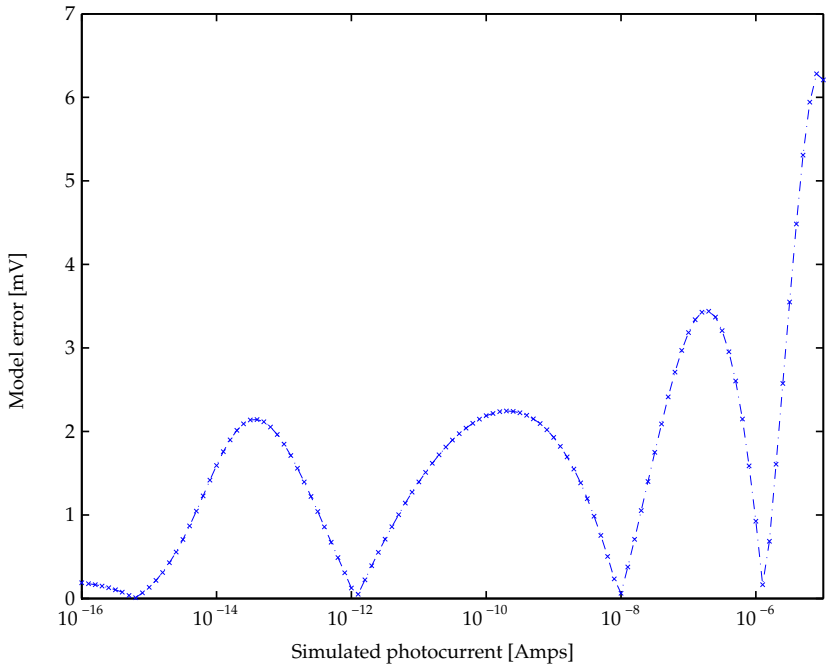


Figure 3.6: The resulting error after extracting the four parameters using five data points.

thought that, because the two high current data points are located right after the transition, they are not far enough in the square law region of operation, for the perfect fit. In terms of computational ease and effort, the coefficients of $1,000$ pixel responses, simulated over 11 decades 0.1 decades apart, were extracted in only 0.82 seconds compared to the 7.08 minutes used by the non-linear regression algorithm. Therefore, any pixel response can easily have all variations in its four parameters corrected using this extraction procedure over the same illumination range.

3.4 'Novel Logarithmic Response'

Modelling results such as those in figure 3.5 and 3.6 suggest a potential for simple image correction by classifying the logarithmic pixel response as a combination of a low current region and a high current region. In fact, while trying to increase the pixel voltage swing by using a bipolar junction transistor as a photo detector, Lai *et al.* [4] claimed a "novel logarithmic response" by using this approach and independently considering the purely logarithmic transfer function in low illumination and a square law one in high illumination. Figure 3.7 shows the responses at both illumination ranges to illustrate this strategy. Although his approach to the correction of fixed pattern noise was not model based, the cross-over between high and low illumination was not accorded adequate thought. However, unlike in the four parameter model, the transition range, 50nAmps - 1μAmps was not accurately accounted for, resulting in susceptibility to error in this two and half decade bracket.

In spite of this, employing the simpler extraction procedure does have another physical and tangible drawback. In this extraction procedure, accurate knowledge of the actual currents is required. In particular, to determine the d_j coefficient, the absolute values of the large currents, x_3 and x_4 are needed, equation (3.26) while parameter a_j requires the knowledge of x_1 - equation (3.27). This is not practically available within a sensor setup. Even if it were possible to create large current sources which were measurable, the mismatches between the current sources in the array, used in the determination of the currents x_3 and x_4, make the task more difficult. Furthermore, when huge currents drive the load transistor into moderate inversion, the current-voltage characteristic is no longer a purely logarithmic one; thus the special advantages of encoding contrast information while capturing the scene as well as the ease of applying techniques like colour constancy, are inherently lost. Overall, it is desirable that for the correction of fixed pattern noise, the model employed be one that satisfies two main conditions.

First, it should be able to adequately cover a wide dynamic range with a purely logarithmic transfer characteristic. This condition is best fulfilled by the two parameter model since the three parameter

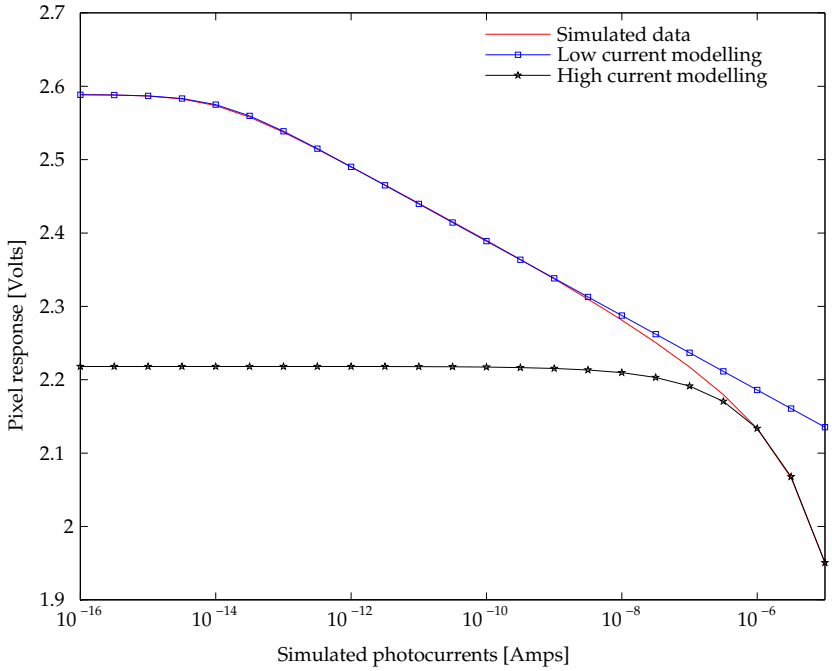


Figure 3.7: Fitting a typical simulated pixel response to high and low current model as suggested by Lai *et al.* [4]

model has very low sensitivity at low illumination while the four parameter model has a non-logarithmic operation at the high current end. To maximise the dynamic range therefore, the leakage effects need to be as low as possible while the inversion current as high as possible. For a fabrication process associated with particular levels of leakage currents, pixel designers can influence the amount of dynamic range between the dominance of dark currents and the onset of moderate inversion; from equations (3.7) and (3.8);

$$c_j = I_s d_j \tag{3.30}$$

$$\therefore \quad c_j \propto \frac{1}{I_{inv}} \tag{3.31}$$

where I_{inv} is the current flowing when V_{gs} is just above V_t. Therefore, by increasing the load transistor aspect ratio and keeping the leakage factor c_j as low as possible, equation (3.7), the dynamic range where the model is purely logarithmic, can be maximised without adversely affecting performance [54]. With this possibility, ranges adequate for imaging natural scenes can be successfully claimed [58]. This avoids the need to take a third, dark reading which would be required if the three parameter model were to be applied, although at the expense of a 2 decade loss in dynamic range.

The other condition desirable in performing image correction, is that it has to be a convenient procedure that does not involve the difficult task of measuring or acquiring the actual photocurrent values. In other words, FPN correction should be performed without the knowledge of the actual photocurrents induced in the pixel photo detectors. This is relevant for the electronic calibration procedure since the calibration current sources and other device components within the pixel are temperature dependent; therefore measuring currents for every change in imaging conditions is undesirable. These two conditions point to the two-parameter model as the best model to achieve simple fixed pattern noise correction from a logarithmic sensor over a sufficiently reasonable dynamic range. What remains, is it to develop a practical image correction procedure to realise this correction.

3.5 Practical Parameter Extraction

3.5.1 Mean Response Method

One way of developing a practical routine is to correct the pixel responses from a single stimulus to a mean response at each photocurrent - this mean response constituting a mean offset and an average gain. Considering that the assumptions in section 3.5 still hold, then at two optimally chosen data points x_1 and x_2, we have responses y_{1j} and y_{2j} given by;

$$y_{1j} = a_j + \frac{b_j}{2}\ln(d_j x_1) \qquad y_{2j} = a_j + \frac{b_j}{2}\ln(d_j x_2) \tag{3.32}$$

with the mean responses of the N pixels at those same photocurrents, $\overline{y_{1j}}$ and $\overline{y_{2j}}$ being;

$$\overline{y_{1j}} = \overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j x_1) \qquad \overline{y_{2j}} = \overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j x_2) \tag{3.33}$$

From the equations (3.32) and (3.33), it is possible to eliminate the offset a_j and mean offset $\overline{a_j}$ parameters to obtain the gain b_j and mean gain coefficients $\overline{b_j}$;

$$b_j = \frac{2(y_{1j} - y_{2j})}{\ln(x_1/x_2)} \tag{3.34}$$

$$\overline{b_j} = \frac{2(\overline{y_{1j}} - \overline{y_{2j}})}{\ln(x_1/x_2)} \tag{3.35}$$

Each pixel gain can be referenced to the mean gain parameter using pixel responses and the mean responses at the same two illuminances to get a ratio b_{jref} defined as,

$$b_{jref} = b_j/\overline{b_j} \tag{3.36}$$

$$= \frac{y_{1j}-y_{2j}}{\overline{y_{1j}}-\overline{y_{2j}}} \tag{3.37}$$

But, from the first of equations 3.33

$$\overline{y_{1j}} = \overline{a_j}+\frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)+\frac{1}{2N}\sum_{j=1}^{N} b_j \ln(x_i) \tag{3.38}$$

$$\therefore \quad \ln(x_1) = \frac{\overline{y_{1j}}-\left[\overline{a_j}+\frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)\right]}{\overline{b_j}/2} \tag{3.39}$$

Substituting this into the first of equations 3.32, then we get,

$$y_{1j} = a_j+\frac{b_j}{2}\ln(d_j)+\frac{b_j}{2}\left(\frac{\overline{y_{1j}}-\left[\overline{a_j}+\frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)\right]}{\overline{b_j}/2}\right)$$

$$= a_j+\frac{b_j}{2}\ln(d_j)+b_{jref}\left(\overline{y_{1j}}-\left[\overline{a_j}+\frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)\right]\right)$$

$$y_{1j}-b_{jref}\overline{y_{1j}} = \left(a_j+\frac{b_j}{2}\ln(d_j)\right)-b_{jref}\left(\overline{a_j}+\frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)\right) \tag{3.40}$$

Now defining a new term from equation (3.40),

$$a_{jref} = y_{1j}-b_{jref}\overline{y_{1j}} \tag{3.41}$$

$$\text{then} \quad a_{jref} = \left(a_j+\frac{b_j}{2}\ln(d_j)\right)-b_{jref}\left(\overline{a_j}+\frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)\right) \tag{3.42}$$

$$= a_j{}'-b_{jref}\overline{a_j}{}'$$

where a_{jref} is the reference offset parameter, obtained independently of any currents. This means at every pixel within a set of pixels, the terms a_{jref} and b_{jref} can be obtained using equations (3.41) and (3.37). Hence at any other photocurrent x_k, and using equation (3.42), it is possible to see that, after substituting

for a_j, an expression including a_{jref} and b_{jref} can be obtained from the equation;

$$\begin{aligned}
y_{kj} &= a_j + \frac{b_j}{2}\ln(d_j x_k) \\
&= a_{jref} - \frac{b_j}{2}\ln(d_j) + b_{jref}\left(\overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)\right) + \frac{b_j}{2}\ln(d_j x_k) \\
&= a_{jref} + b_{jref}\left(\overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j)\right) + \frac{\overline{b_j} b_{jref}}{2}\ln(x_k) \\
&= a_{jref} + b_{jref}\left(\overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j x_k)\right) \\
\frac{y_{kj} - a_{jref}}{b_{jref}} &= \overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j x_k) \\
&= \left(\overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln d_j\right) + \left(\frac{\overline{b_j}}{2}\ln(x_k)\right) \\
&= \overline{a_j}' + \overline{b_j}'\ln(x_k) \\
\therefore \quad \frac{y_{kj} - a_{jref}}{b_{jref}} &= \overline{y_{kj}}
\end{aligned} \tag{3.43}$$

pixel responses with FPN mismatches can therefore be corrected for variations from the average response by subtracting the quantities a_{jref} and b_{jref}. As shown by equations (3.37) and (3.41), these quantities are obtained without the knowledge of the actual photocurrents offering a simple, practical and convenient image correction routine. With the a_{jref} and b_{jref} calculated per pixel and stored in memory during calibration, image correction can be executed with a subtraction and division operation at any other photocurrent. Note the striking similarity in the form of equation (3.43) and (2.8), where both equations correct FPN by negating an offset parameter and dividing the result by a gain term. This correction procedure could also provide pixel designers with vital information about the nature of the spread of defects or mismatches in the manufacture of uniform devices. The maps of extracted a_{jref} and b_{jref} offer invaluable clues to the performance of rows and columns of a pixel array. For example dead or flat pixels can be easily identified using this mapping[1] and then excluded from the image processing chain to minimize the effects of non-responsive pixels.

3.5.2 Photocurrent Ratio Method

One other way by which fixed pattern noise correction can be practically performed is by using the photocurrent ratio method. Here, the subtraction of a pixel response from another to remove offsets,

[1] This will be further illustrated in chapter 4.

is improved by accounting for the variations caused by gain mismatch at every pixel location. The corrected pixel response, free from both forms of fixed pattern noise, becomes merely a transformed and then scaled version of the original response. This can be best understood analytically. By taking the same assumptions of purely logarithmic operation over the four parameter equation, the response of the j_{th} pixel at the i_{th} illumination remains

$$y_{ij} = a_j + \frac{b_j}{2}\ln(d_j x_i)$$

As before, when a pixel is stimulated at two carefully chosen calibration currents x_1 and x_2 within the purely logarithmic region, the pixel gain parameter, b_j can be obtained from the responses y_{ij} and y_{2j},

$$b_j = \frac{2(y_{1j} - y_{2j})}{\ln(x_1/x_2)} \tag{3.44}$$

At any other photocurrent x_k, of the same pixel, the additive parameter can be removed by taking the difference of the response y_{kj} and either of the calibration responses y_{1j} or y_{2j}

$$y_{kj} - y_{1j} = \frac{b_j}{2}\ln(x_k/x_1) \tag{3.45}$$

but from equation (3.44), we can replace the gain parameter, b_j;

$$y_{kj} - y_{1j} = \frac{y_{1j} - y_{2j}}{\ln(x_1/x_2)}\ln(x_k/x_1) \tag{3.46}$$

$$\therefore \quad y_{kjcorr} = \frac{y_{kj} - y_{1j}}{y_{1j} - y_{2j}} = \frac{\ln(x_k) - \ln(x_1)}{\ln(x_1/x_2)} \tag{3.47}$$

Thus, a pixel response may be corrected for its offset and gain mismatches at any illumination by using its responses at two distinct currents within the logarithmic region. The left hand side of equation (3.47) can be interpreted as an FPN correction operation that is a ratio of two offset corrected terms and is independent of the actual photocurrent values. By extension, the responses of an array containing a multitude of pixels may be corrected using the same equation. In practical terms, this image correction technique combines simplicity and convenience with computations limited to two subtractions and a division per pixel per photocurrent. Comparatively this procedure is even simpler than the mean response method which requires several operations to acquire the a_{jref} and b_{jref} terms per pixel.

3.5.3 Calibration Routine

The two approaches described above show the practical parameter extraction approaches usable in the simple calibration of logarithmic sensors. In either method, the primary goal is to conveniently acquire the pixel readings at two photocurrents, the choice of which would have followed the strategy in section 2.6.1 and been pre-determined from the pixel design along with the desired choice of accuracy and dynamic range.

Below is the proposed sequence of operations that would have to be executed in a simple logarithmic sensor calibration chain using the circuit of figure 2.1 and a suitably sized fast access external memory module.

1. Shutter the sensor to avoid any light stimulating the photodiode - figure 2.1. This should be electronically incorporated into the calibration design mechanism to avoid any manual intervention.

2. Using the calibration circuitry of the same figure, pulse transistors $M5$ and $M4$ high and vary the voltage controlled current source $M6$ to replicate the two photocurrents that would have flowed through the shunted photodiode. The precise loading of this current source would have also been determined at pixel design.

3. Take the reading of all the pixels at these two calibration photocurrents and store them in memory by sequentially selecting the row and column bias along the entire array.

4. If applying the mean response correction method, calculate the mean response of all the pixels at the two calibration photocurrents then determine b_{jref} for each pixel using equation (3.37). Thereafter calculate a_{jref} per pixel using eqn (3.41).

5. With parameters a_{jref} and b_{jref} stored for each pixel in the array, executing the operation represented in equation (3.43) at each pixel location will remove all offset and gain variations at any other photocurrent, generated during sensor operation.

6. However, if applying the photocurrent ratio correction method, obtain the difference between the pixel responses at the two calibration photocurrents and store them for each pixel.

7. Using eqn (3.47) take the ratio between the difference between the response at any photocurrent and the response at any of the calibration currents to that of the stored difference value. Similarly, the result is a pixel response deficient of any offset and gain variations.

8. Open shutter to commence sensor operation.

3.6 Summary

It has been shown in this chapter that unlike in the low illumination region, the high contrast errors at the high end of the current spectrum are due to the load transistor being driven into moderate inversion. This was investigated and, as suspected, the logarithmic transfer function becomes invalid in this region. Here, a square law relationship exists between the drain source current I_{ds} of the load transistor and its gate source voltage, hence the steep contrast errors.

Therefore, to model logarithmic pixel response over the entire photocurrent region a novel four parameter model based on the EKV MOS transistor model was developed that smoothly represents pixel behaviour over eleven decades of photo stimulus. The model is found to make an accurate fit with the simulated pixel's response when the four parameters are extracted using a regressive non-linear algorithm in MATLAB. Although the model errors from the fit were all below 3.5mV, this method of extracting the parameter was computationally demanding and could prove a burden for real-time applications.

In order to reduce the effort and time consumed in the parameter extraction procedure, a different approach was proposed. The method which focussed on specific regions of the photocurrent range, helped reduce the four parameter model to a simpler forms that could be used in the extraction of the coefficients. Five conveniently chosen data points were used in this procedure, resulting in a similar fit to the function minimisation method, only for higher error at the high end, peaking at about 6mV. Although this error was at least two times that of the optimisation result, there was a tremendous improvement in computation time and effort. Principally, the simple extraction technique provided ease, convenience and accuracy over a wide range that was not possible previously. However, because of the impracticality of measuring currents in real circuits, coupled with the need to retain the logarithmic properties of pixels, the optimised two parameter logarithmic model was considered prime in the development of a practical fixed pattern noise correction technique.

In this quest, two approaches were discussed. In the first procedure, all pixel responses were corrected to the average pixel response at the point of calibration, the assumption being that all corrected pixels constituted a mean offset and mean gain. This was done using pixel responses only. In the second method, the responses were also corrected for offset and gain variations without the actual knowledge of photocurrents but by taking the ratio of two offset corrected terms. In both methods, the significance of the location of the two calibration points was retained, resulting in easy and convenient strategies for FPN image correction. Finally, a practical calibration routine was outlined as a series of operations that can be implemented in conjunction with the electronic calibration circuit of chapter 2.

Chapter 4

Monochrome FPN Correction

4.1 Introduction

The complete understanding of the behaviour of a typical logarithmic pixel under high dynamic range illumination led to the confirmation of the two parameter model as the most appropriate model for convenient fixed pattern noise correction. Using this model, only two forms of mismatch are to be corrected in a simple, convenient and practical way, while maintaining sufficient dynamic range for HDR imaging of typical natural scenes.

This essentially means any pixel in an array can be accurately characterised to allow for practical offset and gain fixed pattern noise correction, at any illumination where the load current is below moderate inversion but significantly higher than the dark leakage current. To employ the practical image correction procedures described in section 3.5, the responses at the two carefully chosen calibration currents are all that are required to use the electronic calibration circuit of figure 2.1.

Previously, the results of earlier methods of correcting variations in fixed pattern noise were recalculated in terms of contrast error. This was an error measure derived from the properties of the complex human visual system. Nevertheless, the Weber ratio of just noticeable difference [3, 82], used in the development of absolute contrast, implied that the sensor was used to accurately reproduce scene illuminations at every pixel location. However, because most captured images are meant to be processed and displayed to human users, the main concerns are not only replicating the captured scene to the user but also maintaining an accurate perception while being aware of the complex visual operations the HVS undergoes.

In this chapter, it is suggested that, by taking advantage of behavioural characteristics of the human visual system, such as local adaptation to small changes and the eye's sharper sensitivity to relative brightness changes rather than to absolute ones, it is possible to replicate perceptual scene quality without accurately reproducing the actual captured illuminations. In fact, the Weber contrast error measure in section 4.2 is adapted to these characteristics of the human visual system, eventually leading to a new contrast error measure for quantifying fixed pattern noise non uniformity. In section 4.3 both a single pixel from a 200×100 array and the entire array are first characterised to validate previous simulation results before an investigation of the sources and forms of variation of fixed pattern noise is discussed in section 4.4. In the study of the way in which fixed pattern noise is distributed along the rows and columns, the extracted offset and gain parameters are mapped corresponding to the orientation of the pixel array. This special advantage of the mean response method in image correction is used to show row and columnar effects as well as determine the effective resolution by excluding noticeable design effects.

It is in section 4.5 that fixed pattern noise is actually corrected using the two methods introduced in section 3.5. Finally, the quantisation and rounding effects are added to the analysis to investigate the impact of the analogue to digital converter on the new contrast error measure. Because applications dictate the required image quality, the choice of the magnitude of the least significant bit or ADC step size is not fixed but leads to an interesting trade off between the usable dynamic range and the contrast accuracy or image quality - a fact further highlighted in section 4.6.

4.2 Contrast Error

Contrast error, which has been used in previous chapters as the measure to assess model accuracy, is an absolute yardstick in the determination of the estimate of the illumination of a scene. However, its calculation involves the ability to determine pixel photocurrents accurately. Since changes in the aperture during camera operation are not recorded at any point in the image and the captured images are not tagged with illumination levels and aperture settings, it is impossible to calculate the absolute brightness values within a scene. Therefore, there is a need to express the variance of pixel output as an equivalent percentage change in input.

In addition, the main market for cameras relies on sensors that capture images that are to be displayed. This suggests that the performance of the sensor should be determined by the ability to capture and reproduce a scene that is perceptually indistinguishable but not exactly the same as the actual scene. Therefore, the sensor should not be able to acquire the same details but create the same effect as those

which would be detected by the user (or human eye). However, the human visual system (HVS) is a very complex biological system with immense functionality. In order to be able to design sensors that take advantage of the HVS ways of replicating perception, it is important to understand the way the human visual system interprets and processes details within a scene and relate them accordingly.

4.2.1 HVS Adaptation

Experiments by psycho-physicists over many decades have endeavoured to characterise the behaviour of the human visual system in subjective and scientific tests to reveal the immense capability of the eye at capturing and pre-processing information. One property of interest is the eye's greater sensitivity to relative illumination changes than to absolute changes [9, 67, 82].

One such example is the fact that despite stars radiating the same amount of visible radiation, they are not seen during the day but are visible at night [3]. This is because, relative to the dark sky, the amount of radiation from the star is significant whereas relative to the sun's radiation, it becomes negligible to the eye. In other words, if a scene's illumination changed by a small percentage value, the brighter parts of the image would still appear equally and perceptually brighter than the dim parts irrespective of the actual luminance values. This property means that eyes do not have to reproduce the scene accurately but merely need to keep the relative illuminances thereby preserving the perceived effect. It will be shown in chapter 7 how this property has been utilised in colour perception and the computer graphics industry for HDR compression algorithms respectively.

This property of the HVS is very instrumental in the way fixed pattern noise is quantified since it is more important to measure image uniformity than absolute accuracy. Adaptation and sensitivity to relative changes by the human visual system means that when viewing fixed pattern noise or non uniformity, it is not as important whether a pixel accurately represents the local photo stimulus of the scene, but how, relative to other pixels, this photo stimulus is reproduced. It is the variations in the pixel responses that matter most.

4.2.2 Simulation

From the discussion above, rather than concentrating on the estimated photocurrents, the way the estimated photocurrents are spread or distributed will be a more revealing feature to our eyes. To investigate this, the same set of $1,000$ logarithmic pixels previously used in chapter 2 and simulated in CADENCE over 10 decades at 0.1 decade current intervals, was employed. The offset and gain parameters were

extracted using the offset and gain correction technique in section 2.6 and the estimated photocurrents determined. The absolute contrast values were then calculated since both the simulated and estimated currents were known.

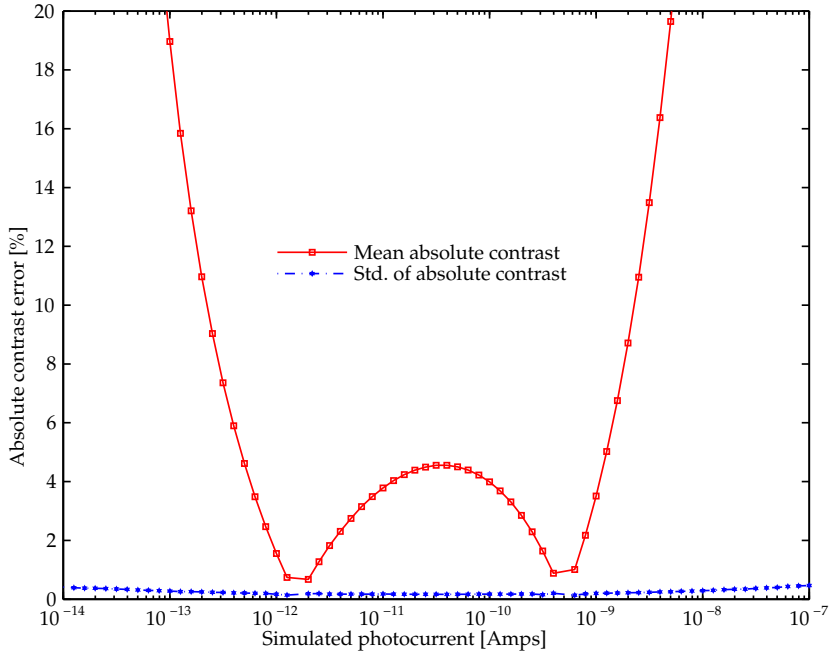


Figure 4.1: The absolute contrast values of 1,000 pixels. A comparison is made of the mean and standard deviation of these contrast values over the wide dynamic range.

Figure 4.1 shows the mean absolute contrast per photocurrent for the 1,000 simulated pixels. As previously seen, the **W**-shape is present with sharp rises in contrast error at both ends of the graph, attributed to leakage effects at the low current end and the non-logarithmic nature of the model at the high end. The figure also shows the standard deviation of the absolute contrast values that are very low at all photocurrents. It can be deduced that the absolute contrast errors occur systematically with a uniformly low spread irrespective of the photocurrent. This means despite the changing absolute errors, they all change uniformly. However, because the eye's prime stimulus is relative change, uniformly misrepresented stimuli should yield a constant output. The almost negligible variations between the estimated photocurrents will therefore be perceived as uniform photocurrents over the entire range.

4.2.3 Relative Contrast

The uniform variations of the absolute contrast values show that the model error is a systematic property along the photocurrent range inspite of the actual contrast value. The implication is that, in the capture

of a uniform gray scale image, it is irrelevant that the pixel responses are inaccurate, but if all these responses are uniformly inaccurate, then a uniform output will be perceived. A new error quantity called the relative contrast can be defined, characterised as the variation of the estimated photocurrents I_{est} from the actual photocurrent I_{act} as a ratio of the actual photocurrent,

$$\text{Relative contrast} = \frac{std(I_{est} - I_{act})}{I_{act}} = \frac{std(I_{est})}{I_{act}} = \frac{\Delta I'}{I_{act}} \tag{4.1}$$

where $\Delta I'$ is the residual noticeable error after image correction.

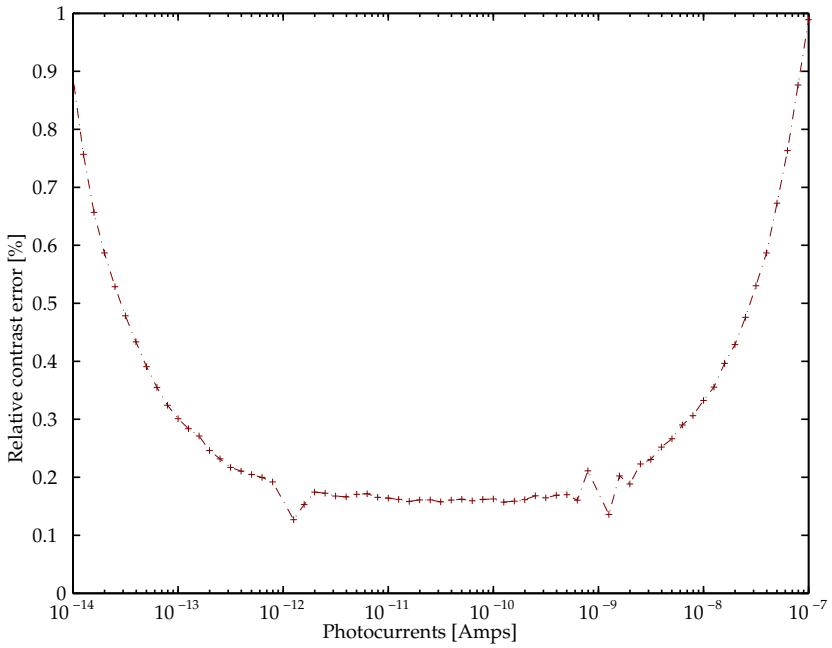


Figure 4.2: The relative contrast determined as a ratio of the standard deviation of the estimated photocurrent from the actual photocurrent to the actual photocurrent at each illumination.

Figure 4.2 shows the results of the mean relative contrast error per photocurrent calculated using equation (4.1) on the same set of simulation data used initially. The relative contrast values stay approximately constant at around 0.2% in the mid current range. With the precision of the simulation data at 0.1mV, the relative contrast values are partly limited by the simulator's inability to fully and accurately model all pixel circuit element behaviour. The 0.2% error was shown to reduce to half this value, when model data rather than simulation data was used. Experience has shown that the rest of the error and the rise at the low and high current ends is because of the greater variation of the dark leakage c_j and moderate inversion d_j parameters which, contributes to the spread of estimated photo currents to varying degrees along the current range. Despite this rise, the relative contrast values are generally low, with all

relative contrast error staying below 1% for about 7 decades of illumination.

This new contrast error target measure yields remarkable results over the entire range. However, there is a need to determine the quality of fixed pattern noise correction from the practical parameter extraction techniques described in section 3.5.1 and 3.5.2. In other words, the relative contrast errors have to be obtained without the knowledge of the photocurrents. By taking an analytical approach, the relative contrast can be determined using only pixel responses. Consider the FPN correction case where all pixel responses are being corrected to a mean response, section 3.5.1. After image correction, all pixels should ideally have the same mean response. Assuming infinite accuracy of the pixel responses, any remaining variation between the corrected pixels yc_{ij} should be attributed to model error of the two parameter model. Assume that within a set of pixels, there is an ideal corrected response yc_{mj}, from which all other pixels vary, at a photocurrent x_m. The relative contrast will then show the variations of other corrected pixel responses from this ideal mean response. Now if

$$yc_{mj} = \overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j x_m) \tag{4.2}$$

then we can represent all other corrected pixels yc_{ij}, as having variable error δx, from the ideal response where

$$yc_{ij} = \overline{a_j} + \frac{1}{2N}\sum_{j=1}^{N} b_j \ln(d_j x_i) \tag{4.3}$$

$$\text{then} \quad yc_{ij} - yc_{mj} = \overline{b_j}' \ln(x_i/x_m) \tag{4.4}$$

$$= \overline{b_j}' \ln\left[1 + \frac{\delta x}{x_m}\right] \tag{4.5}$$

But for very small values, as the ones manifest in figure 4.2,

$$1 \gg \frac{\delta x}{x_m} \therefore \qquad yc_{ij} - yc_{mj} = \overline{b_j}' \frac{\delta x}{x_m} \tag{4.6}$$

$$\frac{std(yc_{ij} - yc_{mj})}{\overline{b_j}'} = std(\frac{\delta x}{x_m}) \tag{4.7}$$

$$\tag{4.8}$$

$$\frac{std(yc_{ij})}{\overline{b_j}'} = \frac{std(\delta x)}{x_m} \tag{4.9}$$

$$\text{where} \qquad \overline{b_j}' = \frac{1}{kN}\sum_{j=1}^{N} [y_{1j} - y_{2j}] \tag{4.10}$$

and k is the logarithm of the known ratio of the calibration photocurrents. From equation (4.9), we see that the right hand side is the same as equation (4.1) and we can determine the relative contrast by dividing the standard deviation of the corrected pixels by the mean of the corresponding gain sensitivity. This method provides a photocurrent free means of determining the relative contrast and hence determines the performance from the practical image correction techniques. A similar argument can be made for the photocurrent ratio procedure described in section 3.5.2.

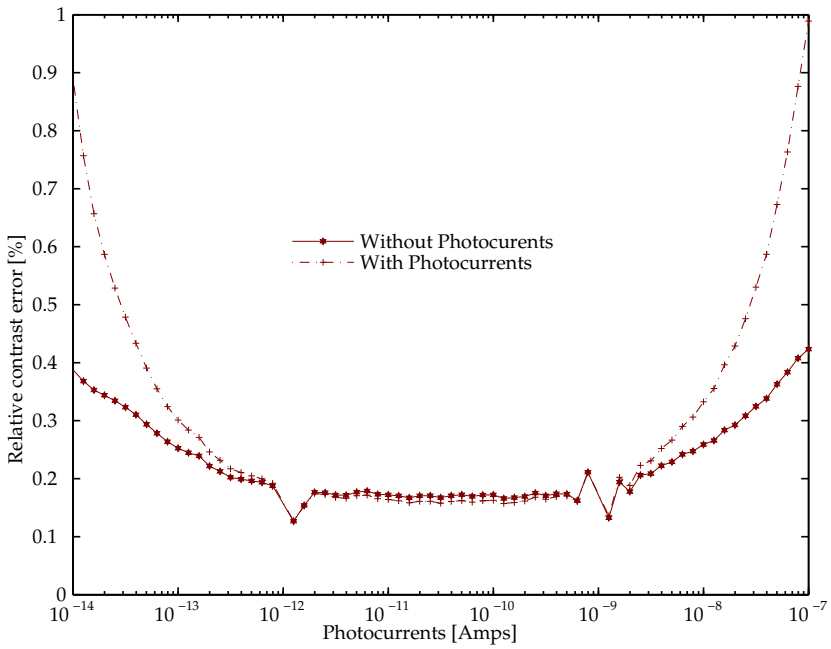


Figure 4.3: The comparison of the relative contrast obtained using the equations (4.1) and (4.9).

A comparison of the results after determining the relative contrast with the knowledge of photocurrents, equation (4.1) and without these photocurrents (mean response method), equation (4.9) is shown in figure 4.3. The two methods perform similarly for the majority of the photocurrent region. In the purely logarithmic current range, the contrast values are nearly equal except for numerical errors whereas in the low current region where dark leakage currents begin to dominate, there is increasing relative contrast error. This is because in this region the onset of the influence of leakage currents changes equation (4.4) to

$$yc_{ij} - yc_{mj} = \overline{b_j}' \ln\left(\frac{x_i + \acute{c}}{x_m + \acute{c}}\right) \tag{4.11}$$

where $\acute{c}$ is the leakage currents; thereby making the condition expressed in the inequality (4.6) no longer true. In the higher current region, the two parameter model which is the basis by which equation (4.9) is formed, is invalid. As a result, there is a growing divergence of contrast values as the currents get into

moderate inversion. However, these values are still less than the 1% contrast target. In addition, since the fixed pattern noise reduction in intended for currents below the moderate inversion region, equation (4.9) can still be used to practically quantify the success of FPN correction.

4.3 Sensor Array

So far, all the data that has been used in the fixed pattern noise analysis has been simulation data acquired from the CADENCE circuit simulator. After adapting the initially suggested absolute contrast error measure to a more meaningful relative contrast error, variations in the offset and gain components in this simulation data yielded relative contrast values that were below the 1% contrast error target over seven decades of illumination.

Nevertheless, simulations are only indicative of the FPN process but are incapable of reproducing real circuit situations since they ignore some pertinent phenomena such as temporal noise and may not fully simulate all device or circuit behaviour. In order to validate previous simulation results, the correction of fixed pattern noise has to be tested on data obtained from real pixel sensors. Particularly, real data can also confirm the amount of parameter variation especially gain mismatch that is present in practical circuits as well as the accuracy or quality of the proposed practical simple offset and gain correction. Before any of these questions is addressed, it is important to characterise the source and form of fixed pattern noise in real pixels.

The sensor of pixel size 10μm $\times 10\mu$m, is a 200 row by 100 column array of logarithmic pixels built by Choubey *et al.* [78] on a 0.35μm AMS process. With a 49% fill factor and two differential readouts, more than six decades of dynamic range were achievable during operation. Using the calibration circuit described in figure 2.1, it was possible to emulate the photocurrent flowing through the photodiode by controlling the current flowing through the pixel. By adjusting the voltage controlled current sources situated at the bottom of each column (transistor $M6$ in figures 2.1 and 4.4) and pulsing the appropriate row selects, pixel readings of the whole array were taken at 23 different voltages. In these circuits, the dark response is obtainable by merely switching off the calibration select switch in all pixels and shielding off the photodiode. Figure 4.4 also shows column amplifier circuitry and a multiplexor from which all array pixel readings were obtained before being relayed to a 2MSPS 16 bit ADC and finally imported into MATLAB for analysis. To limit the effects of temporal noise acquired along the connection between each pixel node and ADC, the pixel readings are a result of averaging 8 readings. By taking a measurement of the same pixel numerous times, the peak to peak temporal noise was found to be about 1mV. This

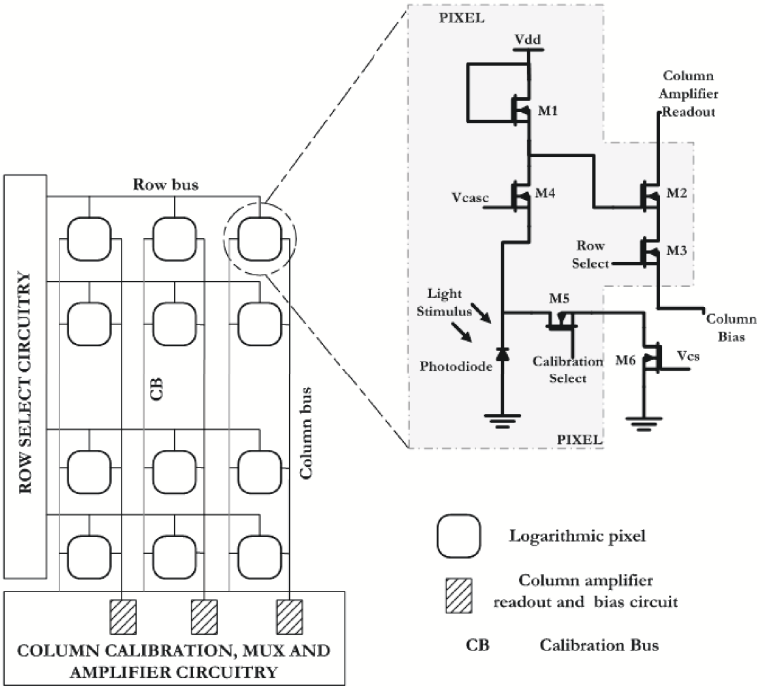


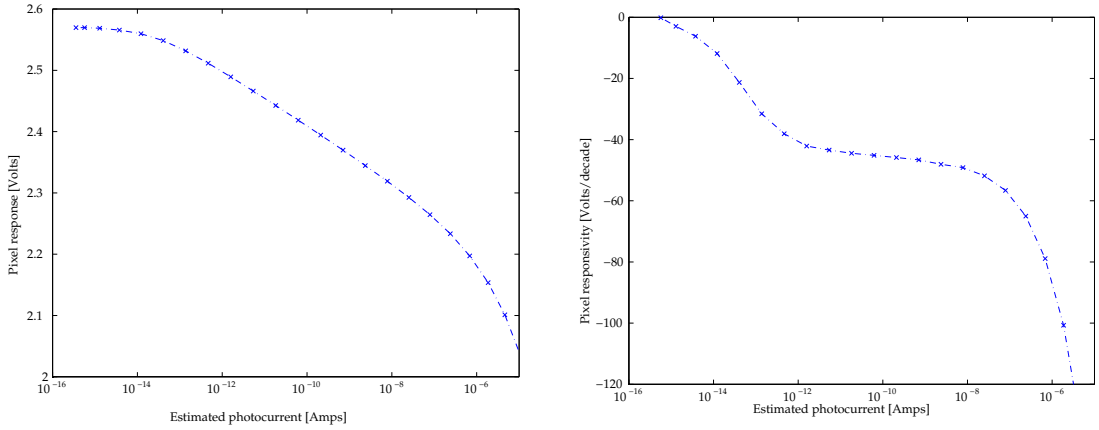
Figure 4.4: Figure showing the array structure and block configuration of the 200×100 image sensor.

means that the average temporal noise over eight readings is reduced by a factor of $\sqrt{8}$, leaving a residual 0.1768mV rms, on all pixel readings. In order to determine how wide a dynamic range was spanned, the same pixel circuit was simulated and the estimated photocurrents obtained. The pixel design, results in simulated photocurrents that extend over 10 decades from about 3×10^{-16}Amps to 1×10^{-5}Amps.

Figure 4.5(A) shows the response of a single pixel over the entire current range, while figure 4.5(B) shows its corresponding responsivity to photo stimulus. As expected, both figures are similar to the results obtained over a shorter range, figures 2.2(A) and 2.2(B), thereby confirming earlier simulation analysis. The responsivity graph also highlights the clear distinction between the different regions of transistor operation that were discussed earlier.

4.4 FPN Distribution

Another characterisation of the pixels in the sensor array above is the study of the distribution of the variations among the pixel responses as this information confirms the understanding of the nature and impact of fixed pattern noise in practical sensors. By studying variation along the rows and columns of an array, the levels of FPN can also be related to elements of the pixel architecture or design process.



(A) (B)

Figure 4.5: Figures (A) shows the response of a typical pixel from the sensor array while (B) is a real pixels' responsivity (gain versus photocurrent) over 11 decades of photo illumination.

This can help highlight the major sources of mismatch and hence alleviate them as early as at the design stage.

Figure 4.6 shows the non-uniformity of 200 pixels in a column, taken at three different photocurrents along the purely logarithmic region, 1.5×10^{-12}Amps, 6.2×10^{-11}Amps and 2.6×10^{-9}Amps. Predictably, the same pixels vary in a similar manner at the different illuminations at which they are stimulated. This is because fixed pattern noise is illumination independent in a two parameter response model thus variations exist only from pixel to pixel. Note that j is a pixel index along a dimension of an array.

$$y_{ij} = \acute{a}_j + \acute{b}_j \ln(x_i) \tag{4.12}$$

Nevertheless, the maximum pixel response non-uniformity along the column is about 50mV which is known to mainly arise from variations in the threshold voltages and subthreshold slope parameter as highlighted from the previous chapter in equations (3.5) and (3.6). This means that for a sensitivity of approximately 46mV/decade in the mid region of figure 4.5(B), this uncorrected non-uniformity leads to photocurrents being misrepresented by at least a factor of 10.

When the non-uniformity along a row of 100 pixels is investigated, at the same photocurrents as those used in figure 4.6, it is clear that there is the same illumination independence but a larger non-uniformity exists, as shown in figure 4.7. In this case, the column to column fixed pattern noise rose to about 90mV. This stems from the threshold voltage variations of the extra calibration and readout transistors that exist

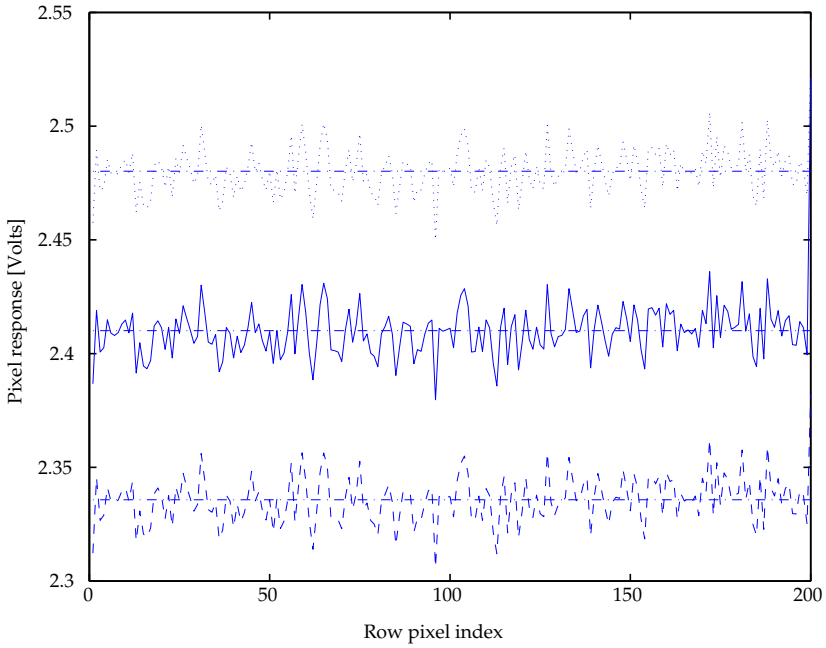


Figure 4.6: Fixed pattern noise along the column of a sensor, taken at three different photocurrents.

for each column as part of the electronic calibration and differential amplifier readout circuit respectively - figure 4.4. This mechanism only exists along each column. Again, this amount of column to column non-uniformity, if left uncorrected, would result in estimated photocurrents that would be almost 100 times brighter or darker than expected for 46mV/decade pixels. Thus, column-to-column fixed pattern noise is at least ten times the level of mismatch at the row-to-row level. This huge discrepancy is due to the fact that along a column, each pixel has a different row select transistor switch while as the row is traversed, the current source for the calibration circuit as well as the readout are different for each pixel - figure 4.4. It is these current sources that need to be carefully designed since their impact on FPN is high.

The distribution of the extracted parameters of each of the pixels across the array may also hold vital clues about the sources of the variations in the sensor. One special advantage of using the practical image correction procedure in section 3.5.1 is that the extracted a_{jref} and b_{jref} parameters could be mapped to reveal mismatch patterns. A mapping of either the offset or gain parameters is likely to show columns or rows with unusually high or low values, which might need to be excluded from the processing algorithms. The offset and gain parameters were therefore extracted from the array and mapped in figure 4.8 and 4.9. In both figures, the row and column indices are representative of their orientation while the gray colour maps[1] show the range of parameter values - these parameters being representative of the offset and gain

[1] A colour map scales all the individual values so that the highest value is displayed as white while the lowest is black.

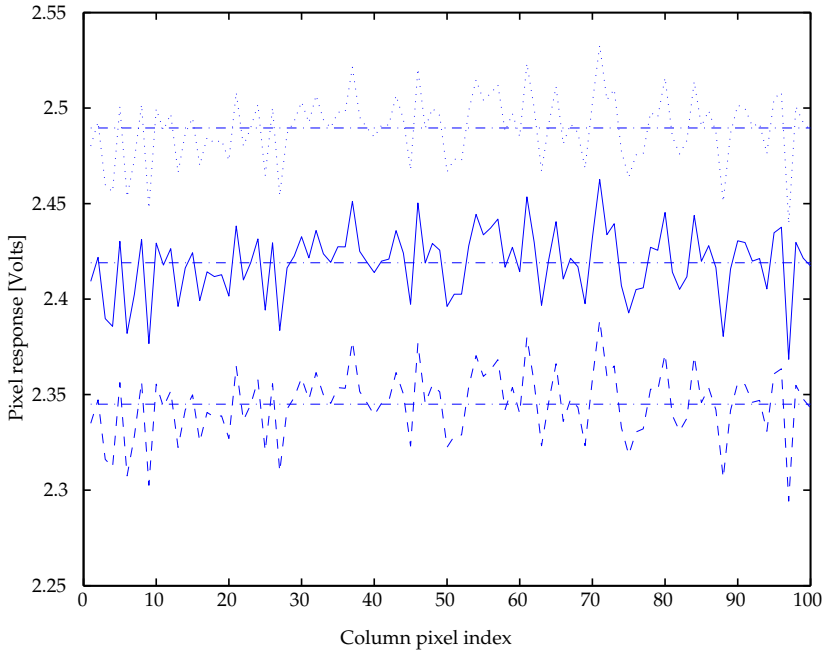


Figure 4.7: Pixel response mismatch along the row of the sensor, taken at the same three photocurrents.

mismatches.

It is clear from figure 4.8 that the bottom row of pixels have offsets that are far higher than the rest of the pixels. Although this defect does not occur in all arrays, it is usually caused by a non-responsive row select switch leading to exceptionally low or high response pixels called dead pixels. The figure also exhibits distinct vertical streaks. Although along each streak there is little variability, there is a striking non-uniformity between the stripes. This is consistent with figures 4.6 and 4.7 which have higher mismatch as a row is traversed. The reason for this pattern is the non-uniformity of threshold voltages along the transistors of the voltage controlled current (calibration) sources and column select circuitry, which are present for each column. A confirmation of this mismatch is the lower fixed pattern noise when the readings are taken at the pixel particularly at the gate of $M2$ in figure 2.1, thus excluding the column circuit.

On the other hand, there are fewer distinct features in the mapping of b_{jref} in figure 4.9, except for the same row defect at the bottom pixel row as in figure 4.8. Generally, despite the more pronounced column variations than between the row, there are less significant vertical streaks when compared to the previous case. However, a couple of pixels between the 160^{th} and 170^{th} rows do show distinctly higher gain as shown by the distinct horizontal effects. This is thought to be a random manufacturing defect as not all pixels along these rows shared this defect. Nonetheless, offset and gain variations exhibit larger

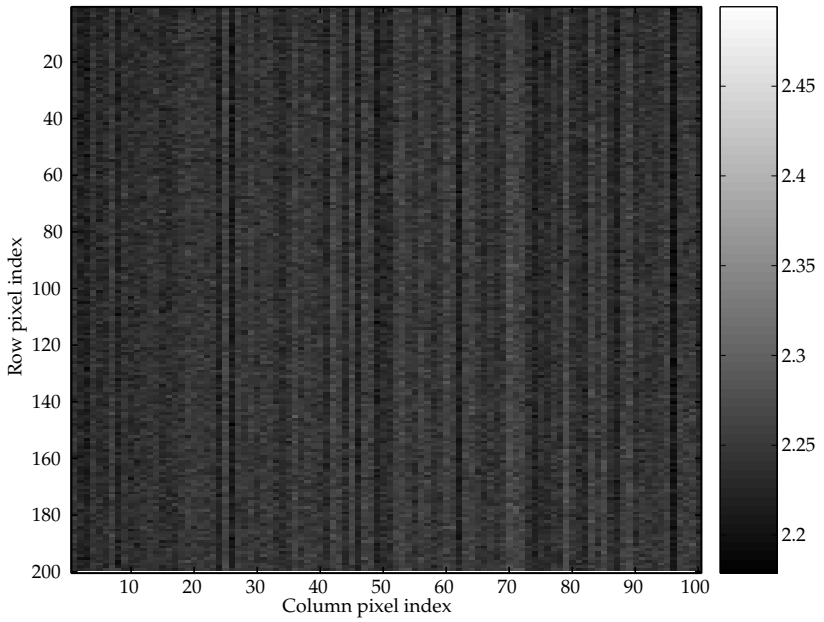


Figure 4.8: A mapping of extracted a_{jref} parameters from the array, and their calculated values shown in volts on the vertical right sided colour bar.

column to column FPN along the rows in comparison to row to row FPN.

To investigate the effect of parameter variation in a real sensor even further, the individual coefficients of the complete four parameter model were extracted for a full analysis over the wide dynamic range of photocurrents. Using the estimated photocurrents and the responses of the pixels in the array, table 4.1 was generated after the parameters had been extracted using the mean sum square non-linear optimisation technique. The statistics of the extracted parameters were compared to those in table 4.3, which are from the $1,000$ simulated pixels used in the chapter 2 and 3. In both cases, the offset and gain variations show no significant difference and in fact confirm that additive mismatch is the greatest form of fixed pattern noise [18, 83].

Parameter	a_j	b_j	c_j	d_j
Units	$Volts$	$Volts/Neper$	$--$	Amps^{-1}
Mean	2.2423×10^{0}	-4.0015×10^{-2}	8.5134×10^{-8}	2.6483×10^{6}
Std. dev.	1.640×10^{-2}	$+1.9711 \times 10^{-4}$	1.0462×10^{-8}	2.0166×10^{5}
% of Mean	$-$	0.4926%	12.2890%	7.6147%

Table 4.1: Statistics of the 200×100 imager array determined by the regressive parameter technique and the estimated photocurrents.

On the contrary, wide discrepancies occur with the variances in the c_j and d_j parameters. The difference in the variation of dark leakage currents, that are infact proportional to c_j, arise from the inability of

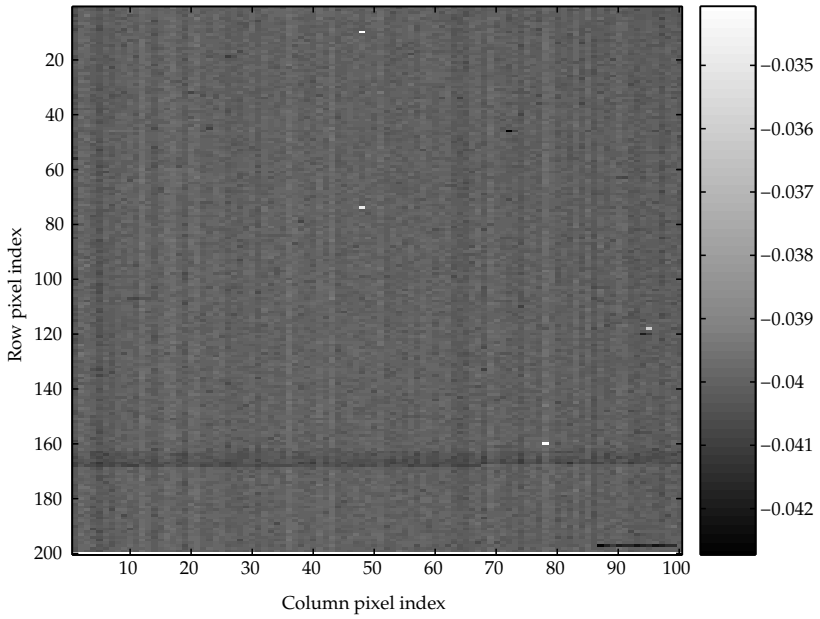


Figure 4.9: A similar mapping of extracted b_{jref} parameters along with their values, shown along the vertical right sided colour bar.

Parameter	a_j	b_j	c_j	d_j
Units	$Volts$	$Volts/Neper$	$--$	Amps^{-1}
Mean	2.2324×10^{0}	-4.0070×10^{-2}	8.1927×10^{-8}	2.5223×10^{6}
Std. dev.	1.0057×10^{-2}	$+1.4416 \times 10^{-4}$	4.7156×10^{-9}	1.0462×10^{5}
% of Mean	–	0.4178%	5.7559%	4.0272%

Table 4.2: Single column statistics from the 200×100 imager array determined by the regressive parameter technique and the estimated photocurrents.

Parameter	a_j	b_j	c_j	d_j
Units	$Volts$	$Volts/Neper$	$--$	Amps^{-1}
Mean	2.2090×10^{0}	-4.4756×10^{-2}	4.2655×10^{-8}	3.3428×10^{6}
Std. dev.	1.7634×10^{-2}	$+8.9484 \times 10^{-5}$	3.0963×10^{-10}	9.6106×10^{3}
% of Mean	–	0.2002%	0.7289%	0.2875%

Table 4.3: Comparative statistics of $1,000$ simulated pixels, determined by the regressive parameter technique and the simulated photocurrents.

the CADENCE circuit simulator to model leakage effects, hence the very small variations in the simulation. However, the same cannot be said of the moderate current parameter d_j. The striking vertical stripes in the maps of the offset and gain parameters in figures 4.8 and 4.9 suggest vast variations from column to column which are characterised by the calibration current sources. But when the four parameters are extracted for pixels in only one column, hence eliminating the effect of the calibration sources, a percentage variation of 4.0272% is still present. This variation in the d_j parameter is thought to originate from the circuit simulator since the simulator uses BSIM models for transistor theory which might not properly emulate the transition from weak to moderate inversion. Since the leakage current is the ratio of the c_j and d_j parameter (equations (3.14) and (3.15)) and the offset is affected by variations in d_j - section 3.3 - it should be expected that in the correction of FPN using the two parameter model, the results will at least have higher relative contrast values than when simulation data is used.

Overall, FPN characterisation from real sensors is in agreement with previous findings from simulation data. Offset variations show a strong overall effect on FPN, both row and column wise, although the non-uniformity of figure 4.9 and tables 4.1 - 4.2 imply that gain variations are significant enough to warrant correction in real sensors.

4.5 FPN Correction

With the array of logarithmic pixels fully characterised, the next step is the critical task of choosing the calibration data points for this sensor. One way of doing this is to visually inspect the response curve and choose the points according to the selection criteria described in section 2.6.1. However, a more accurate method is to use information from table 4.1. For example, leakage currents I_s can be determined using $I_s = c_j/d_j$ while moderate inversion starts at $1/d_j$. It should be appreciated, however, that in real circuit situations this selection of calibration points will be preset in the electronic calibration circuit operating sequence.

In the correction of offset and gain variations from the pixel responses gathered from the sensor array, both the mean response and photocurrent ratio methods, described in sections 3.5.1 and 3.5.2, were tried on real data to test the correction of FPN using only pixel response data. Following the adaptive properties of the HVS, rather than calculate the absolute contrast, the relative contrast error determined using equation (4.9) and the mean value as a function of the photocurrent shown in figure 4.10. The results of the relative contrast values for both methods are seen to coincide over the entire dynamic range as expected. Also, sharp rises in relative contrast exist at either end of the current range. This is

attributed to the fact that the parameters c_j and d_j have significantly greater variations than simulation data (tables 4.1 and 4.3). The two parameter model does not correct these giving rise to the higher relative error.

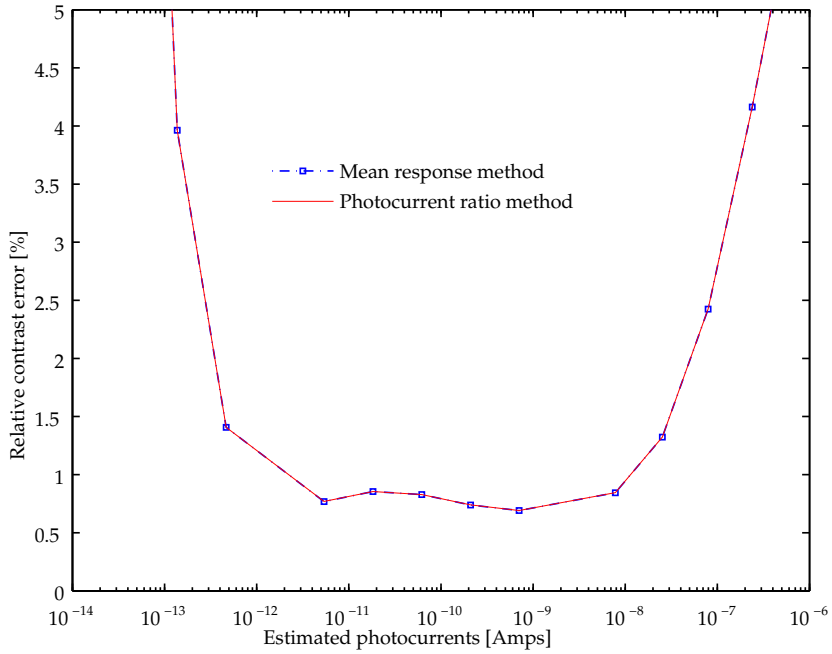


Figure 4.10: The relative contrast values from the array after fixed pattern noise correction using the two simple photocurrent independent routines.

However, one point stands out of figure 4.10 - the relative contrast values in the mid current region are higher than those achieved with simulation data, shown in figure 4.2. A possible cause is the residual temporal noise[2] that exists in measured data despite the averaging done during pixel response measurement. In the mid current region, where the pixel responsivity is about 46mV/decade and a 1% change in current corresponds to a 0.2mV voltage change, the effects of the 0.1768mV$_{rms}$ temporal noise (shown in section 4.3) should be approximately 0.88% - a value similar to that in figure 4.10.

4.6 Quantisation Effects

The correction of fixed pattern noise from a real sensor has yielded under 1% relative contrast over about 4 decades or about 5 decades for a 2% contrast accuracy. However, not all the analysis above has included the effect of analogue to digital conversion that usually occurs before the beginning of the image processing chain. The digitisation of pixel readings is a necessary step but adds quantisation or rounding

[2]Circuit simulations are incapable of emulating temporal noise.

error thereby reducing the quality of the resulting images. It is therefore important to incorporate ADC effects if a true assessment of FPN quality is to be delivered.

In order to proceed, it is necessary to know the accuracy with which the pixel readings were acquired and to what level they have to be quantised. Although the pixel responses were measured from a 16 bit 5V 2MSPS ADC [78], the accuracy of these responses is limited by the temporal noise rms value - 0.1768mV. Because ADC step sizes are usually chosen to be just higher than the noise floor, this value represents the lowest LSB value for this sensor and measuring environment. However to maintain the pursuit of a 1% contrast error target, it is desirable that quantisation is in such a way that the magnitude of the least significant bit represents a 1% change in photocurrent.

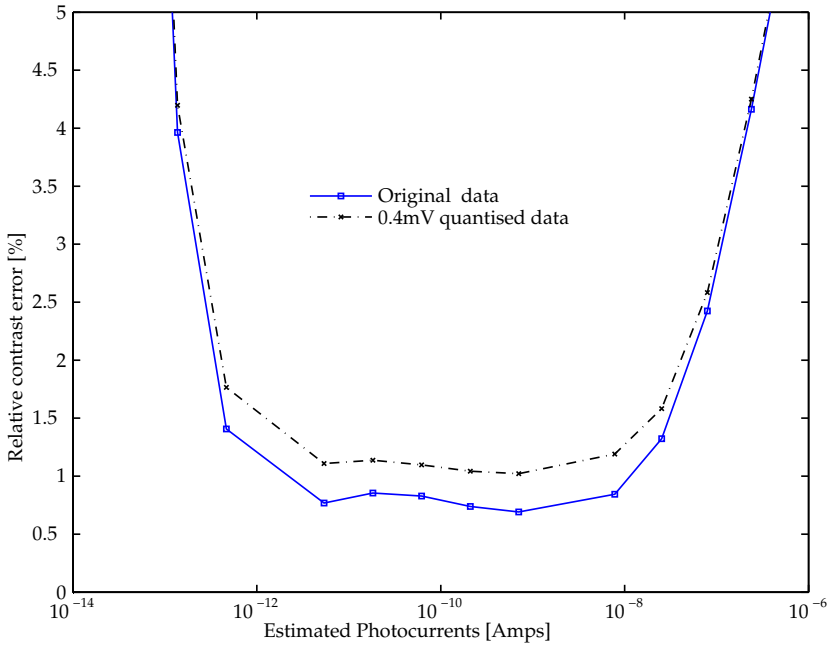


Figure 4.11: The result of ADC effects on pixel array data from the array. A comparison of relative contrast values is made against the original data.

To achieve this, consider from equation (4.12) that a 1% contrast error occurs at an illumination x_i resulting in a change in pixel response Δy, then from the pixel response y_{ij}

$$\Delta x \text{ causes } \quad y_{ij} + \Delta y = a'_j + b'_j \ln(x_i + \Delta x) \tag{4.13}$$

$$\Delta y = b'_j \ln\left(1 + \frac{\Delta x}{x_i}\right) \tag{4.14}$$

where, for a 1% contrast, $\Delta x / x_i = 0.01$. Taking an average gain responsivity of 46mV/decade, a change in pixel response of approximately 0.2mV will represent a 1% input change. Nevertheless, the maximum

error introduced by quantisation is half the voltage change represented by the least significant bit. The pixel readings should therefore be rounded to twice the 0.2mV voltage change.

Therefore, in figure 4.11 a comparison of the contrast error results is made between the pixel responses that have been quantised to 0.4mV and fixed pattern noise corrected as opposed to those that have not been quantised. The mean contrast errors per photocurrent curves follow the same general shape apart from a slight degradation in image quality in mid-current region. ADC effects increase the contrast error in this region but the difference between the quantised and original data leaves the contrast values largely unchanged outside this region. More importantly, contrast errors are generally above the 1% target and if any significant dynamic range is to be claimed, accuracy has to be forfeited.

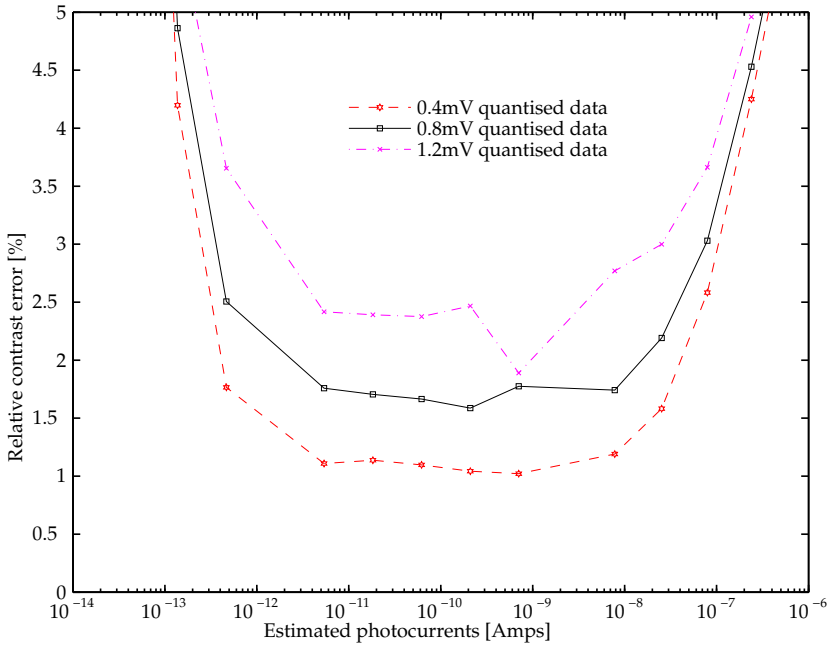


Figure 4.12: Figure showing the trade off between accuracy and dynamic range covered using different quantisation effects.

In figure 4.12, the influence of higher levels of quantisation on relative contrast accuracy is highlighted. By digitising the pixel responses from the array to 0.4mV, 0.8mV and 1.2mV to represent multiples of a 1% change in photocurrent, the effect of FPN correction was investigated. First, It is clear that as the level of quantisation increases, the accuracy of the correction at calibration suffers resulting in less smooth curves, although this is more pronounced at the higher step sizes. It is also possible to see that by quantising to an LSB of 0.4mV, a 5 decade range is attainable if a 2% accuracy is acceptable. However if this quantisation level is doubled, only 3.5 decades can be claimed for the same accuracy. Although there is a significant reduction in range for the same accuracy, this trade-off can be perceived

as an advantage and may be used to cater for the needs of specific applications that value the balance of range and accuracy, differently.

Apart from the dynamic range, the level of quantisation can also be seen to influence the number of bits with which to cover the required range appropriately. For instance, to cover a 5 decade illumination range at 2% contrast accuracy (figure 4.12) a 50mV/decade logarithmic sensor with about 100mV maximum fixed pattern noise range, would require at least a 10 bit ADC if a step size of 0.4mV is chosen. For twice the resolution, a 9 bit analogue to digital converter is required for the same accuracy and illumination range. Similarly, the ADC bit length drops with a corresponding reduction in dynamic range and accuracy.

It should also be noted that there is an approximate 0.5% difference in relative contrast between the curves in the mid current region of figure 4.12. This is because although the maximum quantisation error from quantising in 0.4mV steps, is 0.2mV, the rms quantisation error, is a factor of $\sqrt{12}$ of the step size. This is about 0.115mV in the purely logarithmic region which corresponds to approximately 0.5% change in percentage contrast.

4.7 Summary

In this chapter, the contrast error measure, previously used for benchmarking and characterising fixed pattern noise, has been changed to mimic the way the human visual system adapts to perceive non-uniformity. In particular, the eye's higher sensitivity to relative illumination changes rather than to absolute ones, was central to the adoption of a more relevant relative error measure. Using simulation data, the absolute contrast was shown to have a uniformly low standard deviation over the entire current range, which implied that contrast error was systematic. Using this relative contrast error, simulation data easily achieved the contrast target for 7 decades and was only limited by the modelling limitations of the circuit simulator. It was also proved that relative contrast error could be easily determined from practical electronic image correction techniques mentioned in section 3.5 making it a pertinent and viable error measure.

Although FPN correction on simulation data resulted in achieving the contrast target over a wide range, it still remained unproven with real pixel data. For this reason, the practical image correction techniques were tried on pixel responses from a 200×100 array. Before this was performed, the array was characterised by investigating the distribution and the sources of the various forms of non-uniformity. With typical behaviour from the pixels, column to column FPN was realised to be greater than row to

row FPN, due to the extra transistors in the column calibration circuitry. This was confirmed by maps of extracted offset and gain parameters, which also highlighted the dominance of offset mismatch as well as the significance of correcting for gain mismatch.

The actual correction of fixed pattern noise was shown using the two methods; namely the mean response method and the photocurrent ratios method. A relative contrast error target of 1% was achieved over 4 decades of illumination and an extra decade at 2% contrast. This higher error was attributed to higher c_j and d_j parameter variations in the pixel sensor and to temporal noise which, despite being reduced by averaging, was still considerable.

In this chapter, the effects of the analogue to digital converter were also emulated by adding quantisation error to the pixel responses and correcting for FPN. This was to make a more accurate assessment of the image quality from a HDR imager. The choice of quantisation level was explained and applied to the real pixel responses resulting in an increase in relative contrast error and the consequential decrease of dynamic range under 1% accuracy. The level of quantisation was also shown to affect the number of bits with higher accuracy requiring a longer bit length. This was found to suit the variety of applications that have different accuracy and dynamic range needs.

Generally, it is possible to say that a high dynamic range logarithmic sensor can be easily calibrated using an electronic procedure. Wide dynamic range monochrome or gray scale images can then be corrected for offset and gain variations in a fast and accurate way; this accuracy being about as good as the human visual system can detect over 4 decades or at least 2% over 5 decade images.

Chapter 5

Colour

5.1 Introduction

The preceding chapters have been primarily concerned with achieving high quality wide dynamic range monochrome or gray scale images by correcting for fixed pattern noise in a simple and practical manner. The residual fixed pattern noise after correction remains at levels nearly undetected by the human visual system over approximately five decades of illumination. Discussions in these chapters have also assumed that logarithmic sensors capture only illumination intensity levels from an imaged scene, thereby developing an output voltage signal that has a proportional relationship with these intensities. However, this approach does not cater for the effect of the spectral properties the incident illumination. As a result, only intensity information is captured, without due regard to colour information, which is wavelength dependent.

Moreover, most natural and artificial scenes are not only perceived as varying shades of a singular colour or gray scales. With the exception of dark scenes where scotopic vision[1] is pronounced, a very large proportion of images that are captured by and for the human visual system are in photopic vision[2] - thereby yielding colour images. This is similar to the majority of scenes that commercial, industrial and scientific applications rely on for decision-making. Image sensors can no longer be viewed merely as devices that relate the captured incident illumination intensity to the signal output, but also as colour devices that gather the vast wavelength characteristics of the incident illumination. Ultimately, colour

[1] This is when, in the presence of very low light radiation, only the rods in the human visual system are used for visual sensation [6, 9].

[2] In photopic vision, the cones in the foveal region of the retina are active and excited by high illumination levels [6, 9].

image sensors must be able to output images that are sensitive to both subtle and significant colour shades such as those present in a bright red bus as compared to that of a dark red phone booth. This is especially true for some applications such as professional imaging where colour rendition is of utmost importance.

Despite the independence of the intensity and wavelength properties of visible radiation, colour image sensors still face the same problems as those from a monochrome sensor. In addition to difficulties experienced in capturing and reproducing colour, the effect suffered from the imperfections of fixed pattern noise remain significant. As before, uncorrected FPN and quantisation effects are expected to distort the image both from a contrast and colour perspective leading to misrepresented images. When considering monochrome images, the significant fact that systematic absolute errors did not affect the users' perception of uniformity was employed to reduce the effect of fixed pattern noise further. Although this was successfully illustrated when the illumination intensity and not the wavelength characteristics of the scene were considered, the same cannot be said for colour images without further investigation. Therefore, the importance of the observed levels of residual FPN need to be determined by calculating their effect on the colour of an object in a scene. This will offer and assessment of the effectiveness of the proposed fixed pattern noise correction for high dynamic range colour images; the investigation requiring a quantitative measure for colour accuracy as well as suitable colour data over a wide range of intensities.

This chapter provides the necessary background for this work, particularly exploring the Munsell colour set and some common colour error measures. To commence this investigation and fully appreciate fixed pattern noise correction in colour images, some understanding of basic colour theory is required. In section 5.2, the different means of colour capture and representation are explored after a typical colour image capture and processing chain is illustrated. Various colour models used to represent colour, are introduced in (section 5.2.2) before the MacAdam's ellipses reveal the non-linearity of human colour vision. The CIELAB colour space is then introduced in section 5.3 as a more intuitive and perceptually uniform space. It is this space that will be used to select a perceptible colour error metric for the quality of FPN correction since colour error is visually quantifiable in CIELAB, unlike in several other colour spaces. In particular the ΔE_{76} and ΔE_{00} measures are chosen for simplicity and slightly better accuracy, respectively.

For the purposes of analysis the proposed colour error metric as well as the properties of colour and its interaction with light, a set of spectral reflective colours called the Munsell colours are presented in section 5.4. The Munsell set is studied for coverage of colour gamut in a chromaticity diagram making the set representative of both natural and displayable colours. It is observed, in subsection 5.4.1 that the

sensation of colour we perceive, is independent of dynamic range and that the contrast in these scenes is largely due to variations in the external incident visible light radiation. This is followed by a visual test of the ΔE_{76} CIELAB perceptual colour error metric on each on the three primary colours, randomly chosen from the Munsell colour set.

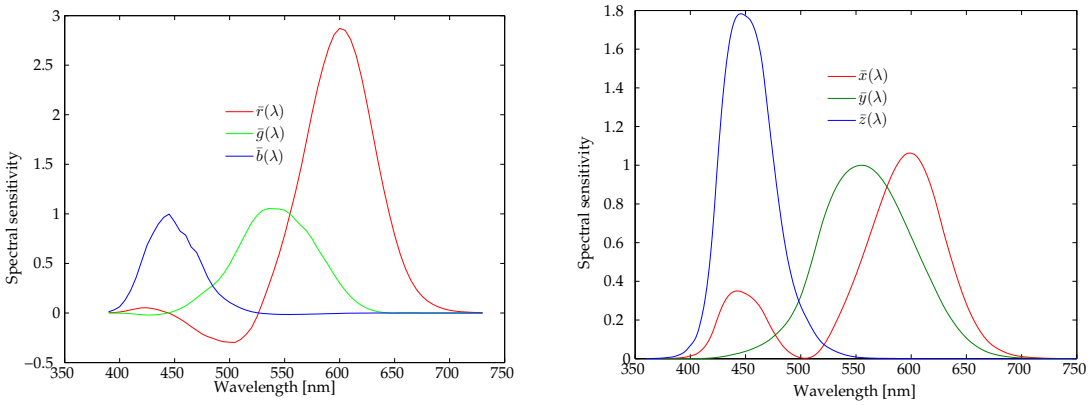
5.2 Colour Theory

The human eye, and particularly the foveal region, contains specialised receptor cells capable of obtaining information about the wavelength properties of incident light radiation at various intensities. The rods are one type of cell that specialise in acquiring information about incident intensity particularly in low illumination - scotopic vision - and are incapable of resolving wavelength dependent light properties.

Colour, as a perceptual sensation, results from visible radiation within the 350nm - 750nm wavelength range of the electromagnetic spectrum striking the retina [9,82], where three types of specialised receptor cells called cones exist. Psycho-physicists have discovered that these cells are active in high illumination - photopic vision - and are generally grouped according to their response to long, medium and short wavelengths of the visible range [6, 67]. These wavelengths generally correspond to the red, green and blue colour bands. Because a multitude of colours are perceivable with only three types of receptors, and with results from a variety of psycho-visual experiments showing that most colours can be obtained by a combination of these three colours, they are formally referred to as primary colours [84]. The phenomenon of matching a single colour using three primaries, called the trichromacy of colour, can be illustrated with results from a colour matching experiment that was adopted by the COMMISSION INTERNATIONALE DE L'ECLAIRAGE(CIE).

Figure 5.1A shows the $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ colour matching functions (CMF's) obtained when a number of standard observers[3] matched different colour samples using different weights of the three primary colours. However, not all colours could be matched using a mixture of the three primaries. Some samples required the addition of the red band to the sample to achieve a match - implied by the negative lobe of the $\bar{r}(\lambda)$ curve in the graph. Consequently, an abstract set of colour matching functions were devised by CIE, in which all the components subtended the same area under the curves, were positive and any colour could be realised by a careful summation of the primaries. The $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ colour matching functions, shown in 5.1B and derived as a linear transformation of $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$, were therefore devised and are currently used to replicate any displayable colour. Using these colour matching functions,

[3]Standard observers are individuals whose eyesight is considered average in terms of visual acuity, colour discrimination and eyesight at a particular viewing distance and angle out of a significant population.



(A) (B)

Figure 5.1: Figures (A) and (B) show the 1931 RGB and XYZ colour matching functions for a standard observer for a 2° viewing angle.

any colour under radiation a spectrum $s(\lambda)$, can be ideally specified using three tristimulus values X, Y and Z, obtained by

$$X = \int_0^\infty \bar{x}(\lambda)s(\lambda)d\lambda \tag{5.1}$$

$$Y = \int_0^\infty \bar{y}(\lambda)s(\lambda)d\lambda \tag{5.2}$$

$$Z = \int_0^\infty \bar{z}(\lambda)s(\lambda)d\lambda \tag{5.3}$$

where the Y co-ordinate was designated to carry the intrinsic luminance information contained in the colour. Ideally, the goal of all systems that need to capture and manipulate colour should therefore be to replicate the functions in figure 5.1(B) if an accurate rendition of all colours is to be achieved without any colour correction. This is assuming the entire lens, filter and other system variables are constant with respect to wavelength. However, the reason why this does not occur is because of the difficulty in accurately producing the double peaks in the red band during filter manufacture. Nevertheless, colour capture is performed using functions similar to $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ but without the negative portion or by functions that resemble linear transformations of the standard $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$. As a result, not all visible colours can be additively reproduced.

5.2.1 Colour Capture

It has been revealed from the previous section, that to capture colour information one need not collect illumination data at each wavelength but only over three broad primary colour spectra. The process

of the collection of this wavelength dependent illumination over the three bands then becomes vital to accurate colour representation and reproduction. In analogue film, three photosensitive layers or dyes enable the separate collection of the red, green and blue components of the illuminated scene. This has been replicated by the FOVEON $X3$ [85–87] sensors which employ different absorption depths of the red, green and blue energy photons to penetrate the photodiode p-n junctions at various depths. This allows the appropriate signals associated with a particular colour to be collected separately depending on absorption depths.

Likewise, in early digital still cameras, colour wheels were employed with a monochrome sensor, taking the red, green and blue images sequentially before combining them to produce a colour image [33]. Understandably, this was a slow process only viable with stationary objects and any minute motion caused significant colour fringing and aliasing. Colour splitting prisms were also employed in some cameras that had three separate sensors, one for each primary colour but they also suffered aliasing, prism effects and could not scale well with technology. Nevertheless, the majority of imagers in the market currently employ colour filters as the preferred colour capturing mechanism.

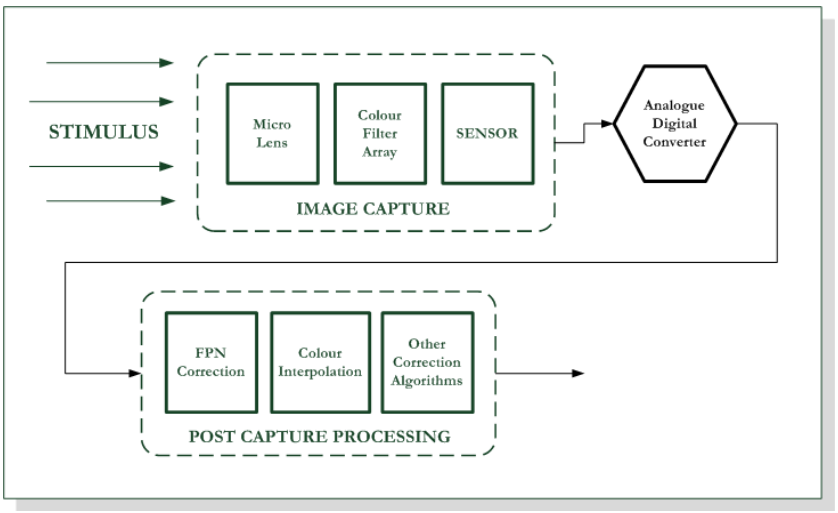
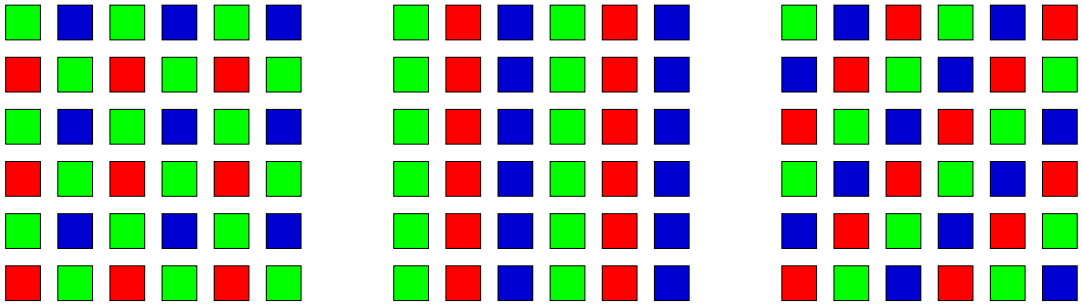


Figure 5.2: A high-level schematic of the typical colour sensor processing chain.

In figure 5.2, a schematic showing the means by which most sensors collect and process colour, highlights their position in the colour handling chain. Here, micro-lenses, placed on top of each of the filters, are used to increase colour signal strength from the filters that are positioned over the sensor. This is achieved by their convex optical design intended to avoid light falling on the photo insensitive areas of the pixel by focussing most of the incident light onto the diode junction. Although filters are meant to capture colour information most optimally with minimal attenuation of the signal, micro lenses are seen



(A) (B) (C)

Figure 5.3: Different types of colour filters arrays used in CMOS image sensors. (A) shows the Bayer pattern [5] while (B) and (C) represent the column and diagonal patterns respectively.

as complimentary to the overall design.

Generally, colour filter arrays contain different colour filters placed in a particular arrangement, on top of each pixel in the sensor, with each filter corresponding to a particular colour band. These filters, usually RGB filters - since they try to replicate the $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ matching functions - are common in digital still and video applications because they require minimal colour processing before displaying and therefore enable low data transfer rates. They are also narrow band filters that provide good detail in the subsequent images. On the other hand, CMY filters - derived from the secondary colours - are becoming more popular due to ground-breaking advances in sensor design plus improved manufacturing and understanding of different sources of noise. In addition, they enable higher SNR and better colour signal since they have broader wavelength coverage and greater light transmission. Nevertheless, their downside becomes apparent when correcting for colour since CMY filters add more noise to the process than RGB filters when applying the colour correction matrix at post processing. Nonetheless, because of the widespread use of RGB filter patterns in diverse applications, they have been the filter of choice in this work.

A variety of colour filter patterns exist, with the most common type being the BAYER pattern [5], figure 5.3(A), which consists of a checkerboard mosaic structure where twice the number of green pixels as blue or red ones are arranged in every square grid. Alleyson *et al.* [88] proved using a mathematical model that the Bayer pattern was the most optimal way of aligning three colours in a square structure. However, he showed that because spatial acuity is less critical in colour than luminance acuity, other colour processing like colour conversion and quantisation have a stronger effect on image quality than the mosaic [89]. Despite the widespread use of the Bayer filter pattern, other filter mosaics like the column filter pattern used by the FUGA15RGB [53] (figure 5.3(B)) and a diagonal stripe pattern (figure 5.3(C))

have also been successfully implemented. The choice of filter pattern usually depends on the magnitude of aliasing effects and the ease with which the magnitudes of the other two primaries can be easily and accurately estimated from the single value at a particular pixel during *demosaicing*. Generally, still applications tend to prefer different patterns from video/motion applications because in the latter, certain types of artefacts can be tolerated unlike in still photographs.

Since sensor manufacturing processes produce variations in the spectral sensitivity of the system elements, colour filters included, equations (5.1), (5.2) and (5.3) are usually modified to include the effects of wavelength dependent attenuation by the filter $f_k(\lambda)$, photodiode quantum efficiency[4] $\eta_k(\lambda)$ and lens $l_k(\lambda)$ on the captured colour where $k \in \langle R, G, B \rangle$. Thus, when a pixel array is illuminated by a spectral irradiance $s(\lambda)$, each pixel any RGB pattern will measure a red, green or blue photocurrent given by

$$I_k = \int_0^\infty \varphi_k(\lambda) s(\lambda) \tau_k(\lambda) d\lambda \tag{5.4}$$

$$\text{where} \quad \tau_k(\lambda) = f_k(\lambda) l_k(\lambda) \eta_k(\lambda) \tag{5.5}$$

φ_k represents the $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$ colour matching functions and I_k is the corresponding single colour pixel response.

5.2.2 Colour Spaces

The complicated nature of capturing and manipulating colour means that the ability to reproduce and measure colour and any variations depends on how well it can be represented or quantified in a standard, device-independent way. Colour spaces offer the means by which any colour can be described in a multi dimensional space. Most colour spaces have three degrees of freedom corresponding to either the three primary colours or an abstract but mathematically meaningful triad, leading to a gamut of colours for each colour space.

Specifying a sample colour using its RGB values - equation (5.4) - is the most common way by which colours can be standardised for use across a breadth of applications and devices. In this space, a framework based on the nature of the colour filters used at capture is defined by the amounts of red, green and blue colours that are required to reproduce any colour. Here, each saturated primary colour is chosen to be at the ends of a three dimensional unit cube. All colours that are obtained by adding any linear com-

[4]The quantum efficiency is a measure which defines how effective the sensor is at converting the incident photon energy into a voltage.

bination of these colours is contained in this cube with black being designated as the origin of the unit cube. With absolute white at the opposite end of the origin, a line drawn from black to white represents several shades of gray. This additive colour space is very relevant to display technologies where colours are projected as a sum of three primaries and as result the cameras, computer and commercial displays find it pertinent. On the contrary, printers, scanners and lithography rely on dyes, pigments and paints which have subtractive properties hence the CMY colour space is found to be more favourable [3, 90]. Theoretically, the CMY colour space is the complement of the RGB space, hence subtracting the RGB values from unity is sufficient. However, because of the overlap between the colour bands, accurate colour reproduction is better achieved using CMY filters where the non-linearities of dyes and paints are accounted for. Generally, the RGB or its complementary CMY do not have an intuitive perceptual representation of the amounts of the primary or secondary colours in a sample. This makes colour matching and reproduction difficult and non-intuitive for users.

Alternatively, a colour can be represented using its CIE XYZ tristimulus values (equations (5.1) - (5.3)) which is an industrially and mathematically abstracted colour standard. As mentioned previously, they originate from the $\overline{x}(\lambda)$, $\overline{y}(\lambda)$, $\overline{z}(\lambda)$ colour matching functions described in section 5.2, derived by setting the areas under these functions to unity and letting one of the tristimulus values proportional to the luminosity[5] function. From this representation, it is therefore possible to tell the luminance and lightness[6] of a particular colour from the Y coefficient, while the X and Z constituents contain estimates of the red and blue colours respectively. Colours in this space can also be transformed to the RGB space using a 3×3 matrix at a specified illumination. However, in the XYZ space, the corresponding amount of constituent primary colours cannot be easily discerned since they are transforms from the primary colours, which are also not inherently visually quantitative, and hence are quantitatively and perceptually non-linear. Consequently, it is also a space that does not perceptually relate with real colours but is an agreed reference of all colours as being composed of non-negative constituents.

Often, colour is also described as a combination of two separate but closely interacting components. Chroma[7], which contains all the hue or spectral properties and lightness that is dependent on luminosity of the incident light. This has been the basis for the formation of another set of colour spaces that despite being represented by a triad of parameters is primarily divided between a chroma and lightness portion. The xyY colour space, which is a derivative of the XYZ space, is a prime example. xy is the chroma

[5]The ratio of luminous flux at a specific wavelength to the radiant flux at the same wavelength.

[6]The difference between luminance and lightness, is that the former is the radiant power falling on a surface from a source weighted by a spectral sensitivity function that is characteristic of vision while lightness is the human vision's non-linear perceptual response to luminance.

[7]Chroma is a collective term that refers to the component of a colour without any luminance information.

obtained from the chromaticity equations (5.6) below and commonly represented on a chromaticity diagram while Y still contains the luminance of the sample colour. Not only does this space offer a means of comparing and profiling colour gamuts of different devices such as CRTs, other displays and printers, it also provides a more intuitive and easy way of differentiating colours that have similar chroma with different brightness's or vice versa.

$$x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z} \qquad z = \frac{Z}{X+Y+Z} \tag{5.6}$$

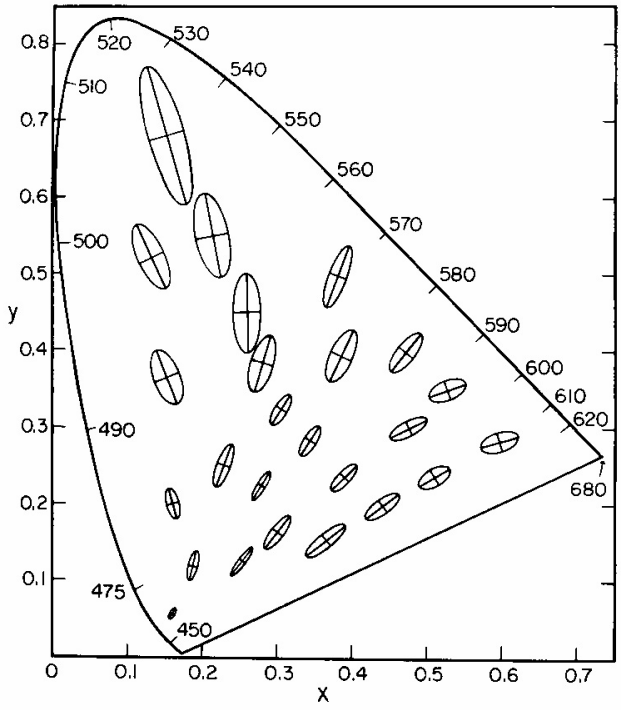


Figure 5.4: 3rd Step MacAdam's ellipses on a 1931 chromaticity diagram showing how different colours have varying perceptual tolerances to chroma shifts [6].

Nevertheless, fully characterising colours in a perceptual and intuitive manner remains a challenge. In an experiment to determine the HVS tolerance to chroma shifts, a sample colour was viewed while keeping the brightness constant over a 2° angle. This colour sample was then progressively changed, triggering a stimulus when the just noticeable threshold was passed. Figure 5.4, shows ellipses of varying sizes at different regions of the chromaticity diagram that represent the colour difference threshold from the original colour (at the center of the ellipse), as seen by the standard observer. It is clear that perception is non-linear in the xyY and its parent XYZ spaces since different tolerances are seen in the greener regions than for example the redder and even the more blue areas. This is because the human visual

system responds to stimuli differently in different parts of the spectrum thus accommodating more shades in the green portion of the visible spectrum than in the red and blue parts respectively.

Analogous to the non-uniform xyY colour space, the HSV (hue, saturation and value) and HSL (hue, saturation and lightness) colour models also separate the colour chroma and lightness into its three dimensional components. In these models, hue represents the dominant wavelength present in the colour while saturation is the purity or amount of white light added to the hue, both collectively known as chroma and usually shown as a planar colour disc. Along the circumference of this disc, are the pure, saturated colours varying along a 360° angular locus while the radius represents the range of saturation. Perpendicular to this colour plane, is the lightness L or value V component which is void of any colour and varies from 0 (denoting black) to 1 (representing white) through different normalised shades of gray. However, just like in the xyY model, the colour gamut created by this non-linear representation is still perceptually non-uniform, shows the difficulty in quantifying colours and is non-intuitive since one cannot tell how different any two samples are.

5.3 CIELAB

The perceptual non-linearities or inadequacies of the colour models described in section 5.2.2, leaves no clear standard colour space that bears a perceptual relationship to the way the human visual system views and interacts with colour. Therefore, when comparing colour differences in the most commonly used colour models (the RGB and CMY colour spaces) a non-linear perceptual relationship exists, mainly attributed to the fact that human perception of luminance is highly adaptive and approximately logarithmic [3, 6, 9]. The advantage of having a colour space with uniform perceptibility is that distances of a sample colour from a desired colour in that model can be viewed as proportional to how much visual difference is associated with the colours. The CIELAB uniform colour space was therefore developed in response to this need. In CIELAB, the lightness component L, is separated from the circular chroma plane that contains two coefficients a and b which represent the degree of redness-greenness and blueness-yellowness respectively - a clear similarity with the human visual system's opponent colour mechanism [3]. By approximating human vision, a perceptually linear space becomes a device independent model where better colour balance corrections are possible by modifying individual components in the respective colours.

Just as in the HSV and HSL, the L component represents greyscale luminance that varies from black to white usually denoted by a $0 - 100$ range and dictating the inherent contrast in a particular colour only

this time it is characterised according to human vision. The a and b components form the chroma on a perpendicular circular plane where the greater their values, the more red and blue there is in a colour sample while the less the values, the more green and yellow are present respectively. All colours are therefore in this circular plane unlike in the linear models, where the colours were embodied in a unit cube. The significant difference between CIELAB and the HSV and HSL models is that in defining the components, human methods of perception are employed unlike the latter. This representation makes understanding the colour concept more intuitive, since colour differences or similarities are viewed as easily changing variables in a three dimensional colour sphere.

In this conceptual colour model, the way our eyes respond to different spectral stimuli presents the main challenge to the definition and evaluation of colour differences between two samples. Because of the varying levels of human visual system's sensitivity, we tend to perceive errors in some hue samples better than changes at other wavelengths as shown in figure 5.4. The differently sized ellipses show the lack of perceptual uniformity among the colours which the CIELAB space seeks to standardize.

Developed in 1976, the CIELAB space model is the most commonly used perceptual metric and, since RGB and CMY are modelled on the output physical devices, transformations from these colour spaces were first defined under a specified illuminant or lighting condition. After linearly transforming a colour from RGB (since most filters produce colours in this space) to its corresponding XYZ values, the same colour can then be defined in CIELAB space as;

$$L = 116\left[f\left(\frac{Y_k}{Y_w}\right)\right] - 16 \tag{5.7}$$

$$a = 500\left[f\left(\frac{X_k}{X_w}\right) - f\left(\frac{Y_k}{Y_w}\right)\right] \tag{5.8}$$

$$b = 200\left[f\left(\frac{Y_k}{Y_w}\right) - f\left(\frac{Z_k}{Z_w}\right)\right] \tag{5.9}$$

$$\text{where} \quad f(t) = \begin{cases} \sqrt[3]{t} & \text{if } t > 0.8856\% \\ 7.787t + 16/116 & \text{if } t \leq 0.8856\% \end{cases}$$

with $X_kY_kZ_k$ representing the XYZ components of the colour while $X_wY_wZ_w$ are those of the corresponding reference white illuminant under which it was imaged.

5.3.1 ΔE Error Measures

Colour error formulae have been developed based on numerous experimental data, yielding error metrics known as ΔE error that remove subjectivity during comparison of colour samples. These error metrics

are based on the differences between the L, a and b components of CIELAB and have huge importance in determining how far off both imaging and print devices have drifted from intended colour rendition profiles. However, due to the varying quality of the different experimental data sets and the subjective focus of varying applications, several colour-difference formulae have come to existence and there has not been one standard for CIELAB colour error measurement.

The ΔE_{76} error was developed together with CIELAB space model and has become the most widely used colour error metric. For any two colours $L_1a_1b_1$ and $L_2a_2b_2$ under a single illumination, the perceptual difference between them, defined by the Euclidean distance in the colour sphere, is called ΔE_{76} and can be easily quantified using;

$$\Delta E_{76} = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2} \tag{5.10}$$

where ΔL, Δa and Δb represent the difference between the lightness, redness-greenness and blueness-yellowness components of the two colours. In the definition of CIELAB, an error of 2 ΔE_{76} units is expected to be twice as pronounced visually, as that from a single unit error for the average observer. Alternatively, only the change in chroma between two samples may be evaluated in special scenarios or applications where brightness is not as important, or less revealing of image features, as a measure of colour difference. In these cases, a ΔE^*_{76} is obtained using the equation (5.11) below, thereby reducing the variables from the colour sphere to a two dimensional non-perceptually uniform colour plane.

$$\Delta E^*_{76} = \sqrt{\Delta a^2 + \Delta b^2} \tag{5.11}$$

Nevertheless, the fact that the experimental data used for CIELAB was gathered under a set of viewing conditions which were low dynamic range over large uniform patches of material samples [28, 91] and limited to the data acquisition capabilities a few decades ago [92], several improvements have been sought.

The CMC (Colour Measurement Committee of the Society of Dyes and Colourists of Great Britain) designed and employed a ΔE_{cmc} error equation constituting hue, saturation and lightness components based on L, a and b. Intended for the textiles industry, it allowed for the user-setting of l and c factors to voluntarily indicate the eye's varying sensitivity to lightness and chroma respectively. The resulting

equation is defined by

$$\Delta E_{cmc} = \sqrt{\left(\frac{\Delta \acute{L}}{l S_L}\right)^2 + \left(\frac{\Delta \acute{C}}{c S_C}\right)^2 + \left(\frac{\Delta \acute{H}}{S_H}\right)^2} \tag{5.12}$$

where $(l:c)$ is defaulted to $(2:1)$, $\Delta \acute{L}$, $\Delta \acute{C}$, $\Delta \acute{H}$, S_L, S_C and S_H are weighting functions for the differences in the lightness, chroma and hue components of the two colours to be compared. Specifically, the chroma determined the Euclidean distance between the *a* and *b* components while the hue was denoted by the angle subtended between them. Later, an equation similar to CMC but based on a different set of data, incorporated viewing condition parametric factors K_L, K_C and K_H to specify the ΔE_{94} error equation (5.13), with S_L, S_C and S_H as additional (but differently defined from ΔE_{cmc}) weighting functions. However, because validity of this error measure was under very particular conditions for example $D65/2$ lighting, the need for colour samples to have direct edge contact while subtending a 4° to the assessor and even only when ΔE_{76} was already under 5 units, it is less commonly used but remains a reference metric for some studies.

$$\Delta E_{94} = \sqrt{\left(\frac{\Delta \acute{L}}{K_L S_L}\right)^2 + \left(\frac{\Delta \acute{C}}{K_C S_C}\right)^2 + \left(\frac{\Delta \acute{H}}{K_H S_H}\right)^2} \tag{5.13}$$

In 2000, assumptions made in earlier colour error metrics were corrected; for example instead of assuming L accurately represented perceived differences in lightness over the entire range, ΔE_{00} varied this component with a dependency on where, within the colour range the sample fits. This resulted in a slightly more accurate error measure which was also closer to perceptual uniformity for samples compared under the same illuminant [92]. Equation (5.14) below shows ΔE_{00} where S_H, S_L and S_C are still the weighting functions for the differences in the lightness, chroma and hue components of the two colours to be compared while K_H, K_L, K_C are parametric factors set to unity. This time, an R_T term is added to compensate for the interactions between the chroma and hue differences.

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta \acute{L}}{K_L S_L}\right)^2 + \left(\frac{\Delta \acute{C}}{K_C S_C}\right)^2 + \left(\frac{\Delta \acute{H}}{K_H S_H}\right)^2 + R_T \left(\frac{\Delta \acute{C}}{K_C S_C}\right)\left(\frac{\Delta \acute{H}}{K_H S_H}\right)} \tag{5.14}$$

Overall, because the ΔE_{cmc} and ΔE_{94} error measures are particularly suited to the textile and graphic arts industries respectively, the ΔE_{76} and ΔE_{00} are still the most widely accepted units of colour error. Although variations of the CIELAB colour space, such as S-CIELAB [91], have been developed to improve uniform perceptibility of ΔE_{76}, it has remained the dominant colour space for comparing uni-

formity of colours across most industries and applications. Consequently, it has been adopted in this work as the unit of preference for quantifying colour error in logarithmic image sensors although in a few cases, the most recent ΔE_{00} is quoted for comparison; hence showing that the conclusions are largely independent of the error measure.

Just as no clear standard exists as a colour error formula, there is no indication of what amount of colour error is acceptable for excellent or even good image quality across the different applications. At this point, the relationship between just noticeable difference (JND) and ΔE_{76} error comes to mind as a fair visual reference. Although researchers such as Kang [93] have concluded that the JND is a single unit of ΔE_{76}, Mahy *et al.* [94] suggested that JND was equivalent to 2.3 units of error, whereas others have proposed ranges of acceptable quality as more representative. For example, Abrardo *et al.* [95] proposed explicit values in their work on the VASARI (Visual Arts System for Archiving and Retrieval of Images) scanner that imaged artwork for archival purposes. Hardeberg *et al.* [96], on the other hand, suggested broader less subjective ranges while performing colorimetric characterisation of printers and scanners. Table 5.1, shows these authors' views, resulting from experimental findings. While it is important to note that any error target is subjective and should take into consideration the intended application, it is reasonable to say that a ΔE_{76} of less than 3, will sufficiently retain colour quality for most applications.

Authors	CIELAB ΔE_{76} units and measure of quality				
Abrardo *et al.* [95]	$0-1$	$1-3$	$3-6$	$6-10$	>10
	Limit of perception	Very Good	Acceptable quality	Sufficient quality	Insufficient quality
Hardeberg *et al.* [96]	<3		$3-6$	>6	
	Hardly Perceptible		Perceptible But acceptable	Not Acceptable	

Table 5.1: A table showing some image quality interpretations based on the ΔE_{76} error measure.

5.4 Munsell Colours

In response to the need for coherent nomenclature for every perceptible and distinguishable colour, a set of colours known as the Munsell colours were introduced in the first decade of the 20^{th} century. The standardisation of the Munsell colour notation for colour communication followed the adoption of a colour system that created a relationship between perceptible colours using the three attributes - the hue, chroma and value. With the notation similar to that used in the HSV colour space, Munsell colours first specify the hue defined as angular variables around a 360° locus, then a value attribute for lightness varying from $0-10$ and lastly, a chroma component represented by the radius of the hue circle - Appendix A is

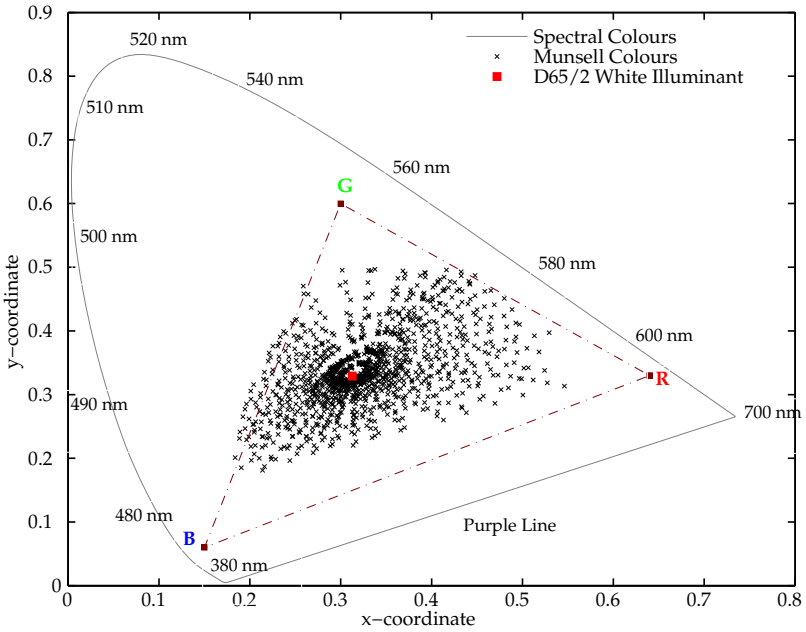


Figure 5.5: The distribution of the Munsell colours on a 1931, 2° chromaticity chart, with a corresponding $D65/2$ white illuminant.

a detailed listing of the $1,269$ Munsell colours illustrating their notation and the relationships between the colours. Palettes (chips) painted in colours corresponding to the Munsell notation were illuminated with $D65/2$ lighting[8] before spectral readings reflected from them onto a UV/VIS/NIR spectrophotometer [97] were measured. The XYZ values of the Munsell colours were then evaluated for each chip using equations (5.1), (5.2) and (5.3).

If any investigation using the Munsell colours is to yield meaningful results, it is vital to establish the range of hues spanned by these colours as well as the gamut they occupy. The calculated XYZ values of the Munsell colours can be used to determine this gamut by obtaining chromaticity co-ordinates - equations (5.6) - and plotting a chromaticity diagram [3]. Figure 5.5, shows a chromaticity diagram with all the Munsell colours relative to the $D65/2$ white reference and the location of pure spectral colours[9]. Note that these colours are derived from monochromatic light while the purple line is an abstract combination of reds and blues for completeness. The figure also shows the standard display gamut [98] represented by the dashed triangle[10]. It is clear that although the colour gamut occupied by the Munsell colours does not completely cover the typical display gamut, it envelopes approximately

[8] $D65/2$ is a term used to refer to lighting conditions that simulate daylight when viewed at a 2° foveal angle by a standard observer.

[9] All colours in the chromaticity diagram are a result of mixing of other colours unlike the pure spectral or monochromatic colours from the electromagnetic spectrum that appear along the edges of the diagram.

[10] The edges of the triangle are the red$(0.64, 0.33)$, green$(0.30, 0.60)$ and blue$(0.15, 0.06)$ co-ordinate points for a CRT/HDTV display at $D65/2$ illumination.

80% of the gamut with colours fairly evenly distributed among the several hues. Hence using this colour set, credible conclusions can be drawn from ensuing analysis.

Overall, the Munsell data colour set has been successfully used in wider research even further validating their use as an authentic tool for analysis with researchers such as Romney who mathematically justified the entire Munsell colour set as adequate for precise characterisation in three dimensional colour space [99]. Similarly Kimball and Roy [100] showed a very close correlation of the Munsell colour coordinates with results from physical reflectance data, modelled with opponent colour theory. These works, among others [101], make the Munsell colours a satisfactory tool for the analytical work in this thesis.

5.4.1 Colour & Dynamic Range

The first investigation with the Munsell colour set is the link between colour and dynamic range. It is interesting to note that as the luminance intensity (and not wavelength), that is reflected off a colour patch varies, it is perceived as the same hue by the viewer. For example, the colour of an object stays the same when viewed in a well-lit room as when the light level is dimmed. The dynamic range of the surrounding light changes but the proportion in each colour appears constant. This seems to imply that the dynamic range and colour properties are independent of each other and that dynamic range is dominated by illumination variations.

The theoretical explanation of the independence of the dynamic range and colour perception can be sought from the incidence-reflectance model of image composition [3] - equation (5.15). Since the images we perceive are dependent on both the intensity (ι) and wavelength (λ) properties of the incident light on the subject as well as the subjects' spectral reflective and absorbent properties, it is possible to say that the latter component $\Re(\iota,\lambda)$ is mostly constant and largely defines the ratios of reflected and absorbed light and hence colour. This is because the light reflectivity or absorption of an object largely doesn't change over time. On the other hand, the incident $\Im(\iota,\lambda)$ light varies widely both spectrally and in intensity hence dictating the perception of dynamic range from the viewed object.

$$\Psi(\iota,\lambda) = \Re(\iota,\lambda) \cdot \Im(\iota,\lambda) \tag{5.15}$$

Assuming a sample scene, the reflective component $\Re(\iota,\lambda)$ of the image $\Psi(\iota,\lambda)$, will contain the colour properties since objects are perceived the way they are depending on the wavelengths they reflect and absorb. Similarly, the incidence component $\Im(\iota,\lambda)$, is associated with the light falling on the patch

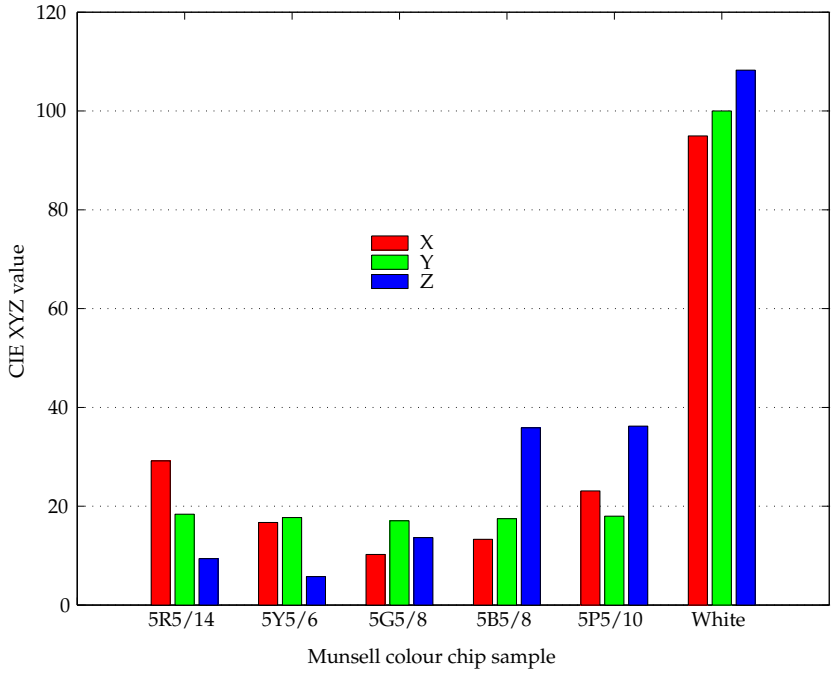


Figure 5.6: Figure showing the XYZ values of some Munsell saturated colours - Red, Yellow, Green, Blue and Purple with white for comparison. It illustrates the low dynamic range exhibited by typical colours.

before it is reflected. Whether more/less light falls on this patch is irrelevant to the ratios of light that are reflected hence the same colour sensation is maintained. Of course, the human visual system performs a host of other functions such as colour constancy to balance for varying luminance wavelengths, making the overall changes appear minimal.

This can be investigated by studying the separate components of Munsell colour samples. Figure 5.6 shows the XYZ coordinates of five Munsell colours including an added white colour under $D65/2$ lighting for comparison. The colours, named and identified by their Munsell colour notation and index (out of the 1269 colours) in table 5.2, represent some of the more saturated hues of the red, yellow, green, blue and purple Munsell colours. The figure reveals that the ratios between the XYZ components of the colours are no more than a factor of five. In addition, the luminance (represented by the Y component) in the white colours is at most 2 decades from the luminance of any of the other sample Munsell colours.

The two decade range is far from the immense dynamic ranges exhibited by natural and commercial scenes (sometimes over 5 decades) some colour images could have - suggesting that colour is probably an inherently low dynamic range concept. In fact, the defining difference between the colours is the ratio between the components themselves that sets them from each other. For example the ratios of the yellow

Colour Name	Munsell Notation	Munsell Index
Red	$5R5/14$	70
Yellow	$5Y5/6$	321
Green	$5G5/8$	587
Blue	$5B5/8$	804
Purple	$5P5/10$	1064

Table 5.2: CIE colour samples and their respective Munsell indices.

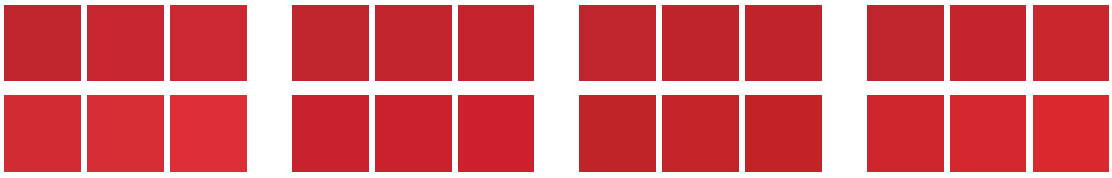
sample have the X and Y nearly the same, while the blue sample has higher Z ratios with respect to the other blue components than the purple. Nevertheless, it is possible to conclude that the dynamic range in the colours of the scenes we perceive, can generally be viewed as ratios in reference to a widely varying luminance.

5.4.2 Colour Chips

It has been noted that if the image quality along the image processing chain of a typical sensor shown in figure 5.2 has a tolerance of least 3 ΔE_{76} units by the end of that chain, then it is generally acceptable for average applications - see table 5.1. Although a particular tolerated error value depends on the sensor application and design, it is not possible to say how perceptually acceptable one level of colour accuracy is from another, without a visual comparison. In other words, a 4 ΔE_{76} error is twice as perceptually erroneous as 2 ΔE_{76} units, but whether this is noticeable is unknown. Using some of the Munsell colours, acceptable levels of colour quality can be both quantitatively and perceptually set for an image sensor.

By simulating the changes that cause this colour difference from an original colour and making a perceptual comparison, we can relate a particular error metric to a visual result. In figure 5.7(A), a random red Munsell colour chip has only its CIELAB L component changed to create a colour error equivalent to 1, 2, 3, 4 and 5 ΔE_{76} units while the a and b components are changed in (B) and (C) respectively. All the CIELAB components are then collectively changed in (D) to simulate the same levels of colour error. In all these cases, the original colour is shown at the top left, with the error increasing row wise. The same simulation is also performed on randomly chosen green and blue colours from the Munsell set to test for any visual variations within the separately altered primary colours.

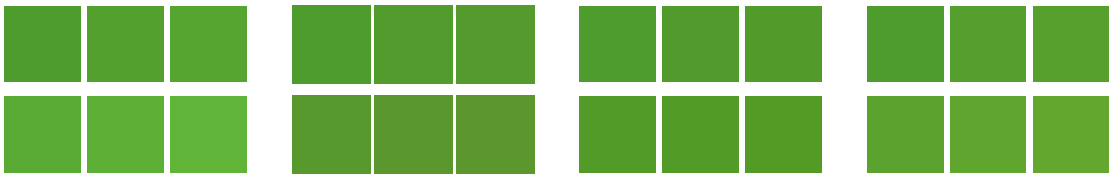
It is apparent that at levels of less than three ΔE_{76}, perceptible colour differences are very hard to conclusively distinguish. This is true for all colours despite whichever component of CIELAB the colour error originates. However, as this error begins to increase from 3 units, visible but still minimal colour differences begin to show in all the different scenarios and for all the colours. This is still consistent with the colour quality ratings in table 5.1. In addition, some scenarios show less visible changes for the same



(A) (B) (C) (D)

Figure 5.7: Effects of added errors on the perception of a randomly chosen red colour of Munsell notation 2.5R 6/12. The original colour is shown as the top left of each of the four images and errors that amount to 1,2,3,4 and 5 CIELAB ΔE_{76} respectively shown row-wise. In (A), only the lightness L component has error, while in (B) and (C) the a and b components are affected. Finally, all the components in (D) have error incorporated.

colour and level of ΔE_{76} error value - evidenced by greater visibility in changes related to lightness and less to do with the hues themselves. This is because, although the ΔE_{76} is an independent unit of error that represents visual changes irrespective of the cause of the colour change, some slight imperfections still exist in creating full uniformity [92]. Nonetheless, experience has shown that the difference between the ΔE_{76} and other CIELAB error measures is small compared to the uncertainty in interpretation of the results in terms of subjective assessments of acceptability. Overall, the subjective quality is still considerably high such that if a logarithmic sensor can produce colour images with less than 3 ΔE_{76} units of colour distortion at each pixel location, these errors will be mostly indistinguishable.



(A) (B) (C) (D)

Figure 5.8: Effects of added errors on the perception of a randomly chosen green colour of Munsell notation 7.5GY 8/8 . The original colour is shown as the top left of each of the four images and errors that amount to 1,2,3,4 and 5 CIELAB ΔE_{76} respectively shown row-wise. In (A), only the lightness L component has error, while in (B) and (C) the a and b components are affected. Lastly, all the components in (D) are error incorporated.

5.5 Summary

This chapter has laid the foundations for the assessment of high quality fixed pattern noise corrected colour images from a logarithmic sensor. Beginning with basic colour theory, the means of capturing or sensing colour information using colour filter arrays were detailed as part of the sensor colour processing chain. Along with the numerous ways in which standard colours are represented, the equations that

(A) (B) (C) (D)

Figure 5.9: Effects of added errors on the perception of a randomly chosen blue colour of Munsell notation 10B 8/4. The original colour is shown as the top left of each of the four images and errors that amount to 1,2,3,4 and 5 CIELAB ΔE_{76} respectively shown row-wise. In (A), only the lightness L component has error, while in (B) and (C) the a and b components are affected. In (D), all the components have error incorporated.

defined the most commonly used RGB & XYZ colour models were also highlighted. The merits and drawbacks of these colour spaces were then discussed particularly, their perceptual non-linearity and lack of intuitiveness in colour representation using Macadam's ellipses on the chromaticity diagram.

The CIELAB colour space was discussed as a perceptually uniform, device independent space from which many colour difference formulae are based and on which most of the colour work in this thesis was focused. Despite the suggested alternatives to the ΔE_{76} error metric for various industries and applications, it has remained the most widely used and quoted metric in the colour imaging field. Consequently, a measure of subjective visual quality was deduced, nominating less than 3 ΔE_{76} units as acceptable for logarithmic sensors.

The Munsell colours were then introduced as a valid and authentic tool set, from which logical deductions could be drawn after analytical simulations had been performed. This is because they sufficiently covered the colour gamuts of typical displays and had been used in various research studies. The set was also used to infer insights about dynamic range of colour, the ratios of the colour components and how perceptually incorrect certain values of be ΔE_{76} error were from others. Overall, this chapter has helped define a pertinent colour error metric and a data set, fit for analysing the quality of FPN correction in high dynamic range colour images.