Algorithms for the Enhancement of Dynamic Range and Colour Constancy of Digital Images & Video

By

Alexis Leonardo Lluis-Gómez

A Doctoral Thesis

Submitted in partial fulfilment of the requirements
for the award of

Doctor of Philosophy

Department of Computer Science
Loughborough University

September 2015

Abstract

One of the main objectives in digital imaging is to mimic the capabilities of the human eye, and perhaps, go beyond in certain aspects. However, the human visual system is so versatile, complex, and only partially understood that no up-to-date imaging technology has been able to accurately reproduce the capabilities of the it. The extraordinary capabilities of the human eye have become a crucial shortcoming in digital imaging, since digital photography, video recording, and computer vision applications have continued to demand more realistic and accurate imaging reproduction and analytic capabilities.

Over decades, researchers have tried to solve the colour constancy problem, as well as extending the dynamic range of digital imaging devices by proposing a number of algorithms and instrumentation approaches. Nevertheless, no unique solution has been identified; this is partially due to the wide range of computer vision applications that require colour constancy and high dynamic range imaging, and the complexity of the human visual system to achieve effective colour constancy and dynamic range capabilities.

The aim of the research presented in this thesis is to enhance the overall image quality within an image signal processor of digital cameras by achieving colour constancy and extending dynamic range capabilities. This is achieved by developing a set of advanced image-processing algorithms that are robust to a number of practical challenges and feasible to be implemented within an image signal processor used in consumer electronics imaging devises.

The experiments conducted in this research show that the proposed algorithms supersede state-of-the-art methods in the fields of dynamic range and colour constancy. Moreover, this unique set of image processing algorithms show that if they are used within an image signal processor, they enable digital camera devices to mimic the human visual system's dynamic range and colour constancy capabilities; the ultimate goal of any state-of-the-art technique, or commercial imaging device.

Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my supervisor Professor Eran A. Edirisinghe for his guidance, support, encouragement, and mentoring throughout all these long years of research. One simply could not wish for a better supervisor.

I would like to thank my colleagues at Apical Ltd. for their hard work and support to make these research ideas become working devices. I would specially like to thank Apical's CEO Michael Tusch and Apical's CTO Slava Chesnokov for the opportunities and support given to me at Apical Ltd. throughout these years.

I am deeply grateful towards my tennis coaches Juan M. Hernandez and Alejandro La Cruz for their support, patient, knowledge, and values in life that taught me to become an elite tennis player and being able to be awarded with sport scholarships for my academic studies. Without them, none of this could have been feasible. Also, I thank my loving, supportive, encouraging, and patient girlfriend Victoria Szychowska, whose faithful support during these years is much appreciated.

My family, and specially my parents deserve special mention for their unconditional support, hard work, guidance, encouragement, and endless love that they have given me, and continue to do so no matter how far apart we are in the world. My parents are the ones that have showed me the joy of intellectual pursuit ever since I was a child, and raised me with their caring and gentle love day and night for my success. Words cannot express how grateful I am for my parents, for their effort and hard work to provide the best possible environment to grow up and be the person that I am today.

Alexis Lluis Gómez, September 2015

Publications

Conference proceedings

Lluis-Gomez, A, S Saravi, and EA Edirisinghe. “Subjectively Optimised Multi-Exposure and Multi-Focus Image Fusion with Compensation for Camera Shake.” SPIE Photonics Europe International Society for Optics and Photonics, 2012. 84360Q-84360Q.

Lluis-Gomez, A.L and EA Edirisinghe. “Chromatic Aberration Correction in Raw Domain for Image Quality Enhancement in Image Sensor Processors.” Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on.ICCP. IEEE, 2012. 241-44.

Lluis-Gomez, AL, and Eran Edirisinghe. “A Novel Colour Management System for Image Signal Processors in Commercial Digital Cameras.” Consumer Electronics (ICCE), 2014 IEEE International Conference on.ICCE. IEEE, 2014. 41-44.

Romanenko, IV, Alexis Lluis-Gomez, and Eran Edirisinghe. “Image Matching in Bayer Raw Domain to Remove Ghosting in Multi-Exposure Image Fusion.” Consumer Electronics (ICCE), 2014 IEEE International Conference on.ICCE. IEEE, 2014. 37-38.

Journals

Lluis-Gomez, Alexis, Sara Saravi, and Eran A Edirisinghe. “Multiexposure and Multifocus Image Fusion with Multidimensional Camera Shake Compensation.” Optical Engineering 52.10 (2013): 102007-102007.

Submitted/under review: Lluis-Gomez, A.L., E. A. Edirisinghe, “Achieving colour constancy within Image Signal Processors”. IEEE Transactions on Image Processing 2015

Table of Contents

List of Figures

List of Tables

Abbreviations

3A	AE, AWB and AF
AE	Auto exposure
AF	Auto focus
AG	Analog gain
AGC	Auto gain control
API	Application programming interface
APB	Advanced peripheral bus (ARM AMBA specification)
AWB	Auto white balance correction
BP	Black point
CCM	Colour correction matrix
CFA	Colour filter array
CNR	Chroma noise reduction
CPU	Central processing unit
CWF	Cool white fluorescent
DG	Digital gain
DIS	Digital image stabilisation
DP	Defect pixel
DRC	Dynamic range compression
DSLR	Digital single-lens reflex camera
FW	Firmware
HDR	High dynamic range, synonymous with WDR
HW	Hardware
ISP	Image signal processor
LED	Light-emitting diode
LUT	Lookup table
MV	Motion vector
NR	Noise reduction
ROI	Region of interest
SoC	System on chip

SPI	Serial peripheral interface
WB	White balance
WDR	Wide dynamic range, synonymous with HDR
WP	White point
YCbCr	Method of encoding RGB colour space according to the ITU-R BT.601 or ITU-R BT.709 standards
YUV	A method of encoding RGB colour space in practice equivalent to YCbCr

Glossary

Chromatic aberration	Lateral chromatic aberration is a type of distortion in which there is a failure of a lens to focus all colours to the same convergence point. Chromatic aberration manifests itself as "fringes" of colour along boundaries that separate dark and bright parts of the image, because each colour in the optical spectrum cannot be focused at a single common point. Lateral chromatic aberration is when different wavelengths are focused at different positions in the focal plane.
Colour accuracy	The measurement of the deviation of colours captured by a camera and the expected colours using a test chart like the ColourChecker.
Colour shading	The measurement of the variation of colour throughout an image (not just the centre compared to the edges).
Colour space	A colour space is a mathematical model describing the way colours can be represented as tuples of numbers, typically as three or four values or colour components. The standard PC colour space for displays is sRGB, while webcams generally use YUV colour spaces for capture.
Defect pixel correction	A defective pixel fails to sense light correctly. Dead pixels are typically classified as minor failures and major failures based on how significant the failure is.
Demosaicing	Also known as de-mosaicing, demosaicking or debayering. This is an algorithm used to reconstruct a full colour image from the incomplete colour samples output from an image sensor overlaid with a colour filter array.
Depth of field	Depth of field is the distance between the nearest and farthest objects in a scene that appear acceptably sharp in an image.
Dynamic range	Dynamic range is the ratio between the largest and smallest possible values of a changeable quantity. For imaging it is

	measured by how many F-stops can be imaged in one scene, which can be done with a step chart like the ST-52.
Field of view	The field of view is the angular extent of the observable world that is seen at any given moment from a camera. Diagonal field of view is the angular extent from the upper left and lower right (or upper right and lower left) corners of the image.
Focus	An image, or image point or region, is in focus if light from object points is converged almost as much as possible in the image and out of focus if light is not well converged. The border between these is sometimes defined using a circle of confusion criterion.
Focus range	The focus range is the depth range that the camera can focus. For fixed focus cameras this is equivalent to the depth of field. For autofocus cameras it is the minimal distance the camera can focus on an object to the maximal distance.
Frame stitch WDR mode	ISP in WDR mode for ISP stitching of N number of exposures input.
Image aspect ratio	The image aspect ratio is the image height: width ratio using the number of pixels for height and width.
Gain	The Gain consists of the values of the Sensor Analog gain, ISP Digital gain and Sensor Digital gain.
Gamma	Gamma correction is a nonlinear operation used to code and decode luminance or tristimulus values in video or still image system.
Geometric distortion	Geometric distortion is a deviation from rectilinear projection, a projection in which straight lines in a scene remain straight in an image. The two most common forms of distortion are barrel and pincushion.
ISO	ISO standard 12232:2006. A measure of photographic sensitivity to light
Lighting temperatures	The three lighting temperatures used in these requirements are 2856 K (incandescent light, also called A), 3500 K (a common fluorescent light in the US, using the bulb Sylvania Octron 3500 K

	FO17/735/ECO, 75 CRI) and 6500 K (also called Day, using the bulb GRETAG MACBETH F20T12/65 D65).
Linear mode	ISP support for single exposure image
Lux	The lux (symbol: lx) is the SI unit of illuminance and luminous emittance, measuring luminous flux per unit area. Lux is equal to lumen per square meter.
MTF	Modulation Transfer Function is the magnitude of the Optical Transfer Function. MTF30 is the cycles/pixel where the MTF=30%, generally considered to be the lowest acceptable MTF for imaging.
Native WDR mode	ISP in WDR mode for sensor stitching of N number of exposures input.
Over sharpening	Over sharpening or under sharpening is an Imatest measurement that characterises the degree that the image is sharpened relative to a standard sharpening model. It if is negative, sharpening is applied to the original response; if it is positive, de-sharpening is applied.
Pixel aspect ratio	The ratio of the height: width of a pixel. 1:1 is a square pixel.
Relative illumination	Relative illumination, or vignetting, is a measure of the image brightness in the centre compared to the corners.
RMS edge roughness	RMS edge roughness is an Imatest measurement of how rough an edge is. An edge can be rough due to demosaicing, image scaling and spatial denoising.
Saturation	Colour saturation is determined by a combination of light intensity and how much it is distributed across the spectrum of different wavelengths. The purest colour is achieved by using just one wavelength at a high intensity, such as in a laser light. If the intensity drops, so does the saturation. An ideal ColourChecker chart has saturation of 100%.
SNR	Signal to Noise Ratio is the ratio of the signal power to the noise power corrupting the signal. The SNR can be measured spatially using a single image or temporally using more than one image.

sRGB	sRGB is a standard RGB colour space created cooperatively by HP and Microsoft in 1996 for use on monitors, printers and the Internet.
Texture acutance	Texture acutance is a measure of how well a camera can capture texture in a scene. It is defined in CPIQ Phase 3 and was designed because denoising algorithms typically don't degrade MTF but can significantly degrade texture.
White balance	White balance is a colour balance that tries to make white (and grey) objects appear white (and grey).
Zoom	Image zoom uses either optical zoom or digital zoom to decrease (narrow) the apparent angle of view of a digital photographic or video image.

Chapter 1

Introduction

One of the main objectives in digital imaging is to mimic the capabilities of the human eye, and perhaps, go beyond in certain aspects. The HVS is so versatile, complex, and only partially understood that no up-to-date imaging technology has been able to accurately reproduce the capabilities of it.

It was not until recent times that the ever-growing computational power, availability of microprocessors and imaging sensor technologies, made high quality imaging devices a reality for consumer grade electronic devices. Nonetheless, digital imaging systems are still significantly inferior, when compared with the HSV. The two most challenging features of the HVS are: its remarkable dynamic range capability to detect subtle contrast variations, and achieving colour constancy under a large variety of illumination conditions. These two factors, have become a crucial shortcoming in digital imaging since digital photography, video recording, and computer vision applications have continued to demand more realistic and accurate imaging reproduction.

The human eye is capable of adapting to lighting conditions that vary by nearly ten orders of magnitude, and within a scene, the HVS functions over a range of about five orders of magnitude [30]. In contrast, standard imaging sensors are only capable of reproducing two orders of magnitude of intensity variations [30]. To overcome the restricted dynamic range in standard digital imaging devices, researchers have proposed multi-exposure image fusion techniques that use standard dynamic range cameras to create high dynamic range (HDR) images. However, the main limitation of such current multi-exposure image fusion techniques is their inability to compensate for moving objects in a scene, camera shake, and being able to run in real time for HDR video content creation.

Colour constancy is the effect whereby the perceived or apparent colour of a surface remains constant despite changes in the spectral composition of the illumination and intensity levels [92]. The HVS is colour constant to a large extent, but this is affected by contents of the field of view, state of adaptation, intensity level of illumination and difference of the adapting chromaticity from different kinds of illumination conditions. Over decades, researchers have tried to solve the problem of colour constancy in digital imaging by proposing a number of algorithms and instrumentation approaches. Nevertheless, no unique solution has been identified; this is partially due to the wide range of computer vision applications that require colour constancy capability and the associated complexity of the HVS to achieve colour constancy [97].

Achieving colour constancy and high dynamic range imaging are not only important for digital video or still photography in consumer cameras. It is also important for computer vision related applications that require high dynamic range vision and accurate colour representation. These facts led to the motivation behind the research presented in this thesis, i.e. to attempt to address the lack of dynamic range and inaccurate colour reproduction capabilities of current digital cameras by investigating and developing advanced image-processing algorithms that are robust and feasible for practical use and implementations.

1.1 Aim and Objectives

The aim of the research presented in this thesis is to enhance the overall image quality of digital cameras by increasing dynamic range capabilities and achieving colour constancy. This is to be achieved by developing a set of advanced image-processing algorithms that are robust to a number of practical challenges and feasible to be implemented in hardware within image signal processors (ISP) used in consumer electronic imaging devises.

The specific research objectives are:

1. Carrying out a literature review to identify current image quality limitations of digital camera devices in comparison to the HVS.

2. Studying existing HDR and colour constancy approaches to identify their weaknesses, as well as analysing the challenges that the algorithms will face when implemented and used in real imaging systems.

3. Reviewing state-of-art image registration and image fusion algorithms. In particular, investigating the robustness of the image fusion algorithms in the presence of camera shake.

4. Developing an HDR solution that allows compensating for camera shake through an image registration process, and proposing a novel image fusion approach that can handle practical challenges faced by the state-of-the-art algorithms.

5. Developing a robust and low complexity multi-exposure image fusion algorithm for video HDR creation that is capable of compensating for motion due to camera shake and moving objects in the scene.

6. Investigating the challenges of colour constancy in digital imaging and reviewing the state-of-art colour constancy methods that can be used in real time applications, as well as identifying their weaknesses.

7. Developing a novel colour management system that encompasses the colour constancy problem and the reproduction of pleasing colours within an ISP pipeline.

8. Benchmarking the proposed colour constancy method to the state-of-art colour constancy methods proposed in literature, and identifying fundamental problems when trying to achieve colour constancy in digital imaging.

9. Redesigning the colour constancy method in order to solve fundamental colour constancy problems in HDR imaging with multiple illuminants or reflectances in scene.

10. Identifying the limitations of the proposed image processing algorithms and outline the future directions of research.

1.2 Contributions

This thesis consists of a number of original contributions and novel ideas mainly focused on enhancing the overall image quality of digital cameras. The key contributions, which are outlined below and presented in detail in Chapters 5 - 7, have been published as conference papers and journal manuscripts. Citations are provided next to the equations' number in order to differentiate equations from previous research, and equations that are originally derived and presented as a contribution in this thesis.

I) HDR imaging

A novel multi-exposure image fusion algorithm that has the unique capability of compensating for camera shake prior to image registration is proposed. This algorithm enables users to create HDR images with a standard dynamic range camera without the aid of stabilising devices such as tripods.

The first contribution presented in this thesis is a novel multi-exposure image fusion algorithm suitable for practical implementation in hardware for video HDR content creation (presented in Chapter 5). The proposed approach for video HDR content creation is implemented as a pre-processing technique that works in RAW Bayer domain and can be incorporated within an ISP pipeline. This novel method performs global motion estimation and correction to compensate for motion due to camera shake, followed by an image fusion process, wherein its operational logic compensates for objects moving within the captured scene. The combination of these two steps, i.e., global motion compensation in conjunction with the motion aware fusing technique, allows the removal of motion artefacts and ghosting due to camera shake and moving objects in the scene. It is shown that the proposed algorithm performs efficiently in different lighting conditions, scenes, and is robust enough to be able to perform with different imaging sensors. The proposed novel solution for extending the dynamic range of digital cameras enables digital imaging devices to create HDR content in real time.

II) Colour constancy

A novel colour management system that works within an ISP and is able to compensate for different illuminants and at the same time reproduce pleasant colours is proposed in Chapter 6. The proposed approach is designed in such a way that it is robust enough to cope with sensor-to-sensor variations, accurate enough to pass industry image quality standards, and flexible to reproduce pleasant colours in images by following Kruithof theory [95].

The proposed colour management system within an ISP pipeline encompasses three stages. First, the colour temperature of the illuminant is estimated (AWB module); secondly, a colour correction/balancing matrix is applied (CCM); and thirdly a cooling-warming effect is applied to achieve a typical camera end-user's preferred colours, and therefore, achieving a solution to the colour constancy problem in digital imaging.

In order to verify the performance of the proposed AWB module, a large database consisting of RAW Bayer images is captured using two different DSLR commercial cameras, and an image sensor commonly used in smart phones (OV8835). With this database, the proposed AWB method is benchmarked against the most known colour constancy methods in literature.

The database of RAW Bayer images is published and can be accessed at:
https://drive.google.com/folderview?id=0B2xIm02tjf9kTTFZSTB4UXB2UDg&usp=sharing

The results of the experiment show that the proposed method outperforms all other state-of-the-art colour constancy methods in terms of accuracy and robustness. That is, the proposed colour management system proves to be capable of balancing neutral and spectral colours under all types of illuminants, and is able to handle challenging scenes such as: monotonic colours, green grass colour versus CWF lighting, opposite colour objects' surface reflectance to the light colour temperature in the scene, sunset, low light conditions, absence of grey, noisy images, mixed lighting, cloudy scenes, and scenes with blue-sky images etc. It is noted that all other colour constancy methods fail under these extreme circumstances.

The last contribution presented in Chapter 7 of the thesis is a unique and sophisticated non-linear image processing method whose aim is to solve the colour constancy problem in HDR imaging with multiple illuminants or reflectances in the scene. The novelty of the proposed method is that it enables digital camera devices to increase dynamic range and colour constancy capabilities. This novel solution incorporates the proposed video HDR solution, an image segmentation process based on the image luminance levels, and a novel spatially variant white balancing correction method based on the proposed colour management system mentioned above. The results of the experiments performed demonstrate that the proposed method performs efficiently under different lighting conditions, scenes, and is robust enough to work well with different imaging sensors. Moreover, results of the experiments performed shows that the proposed method improves the subjective image quality and make high dynamic range images look more natural and well balanced in colour.

1.3 Thesis Structure

This thesis is organised in two parts. The first part includes non-contributory chapters providing basic information, framework, and fundamental background knowledge about the research problem addressed in the thesis. The second part of this thesis includes three contributory chapters, each dedicated to a particular problem in digital imaging and a novel approach to addressing the problem. A summary of the chapters of this thesis is presented below.

Chapter 1 provides an overview of the thesis, defines the research problem, states the research motivation, specifies the thesis aims and objectives, and outlines the organisation of the thesis.

The background theory and literature review are presented in **Chapter 2** and **Chapter 3**, where fundamental concepts and theories on different areas of mathematics, computing and engineering related to the Human Visual System, digital video and still photography are introduced. Moreover, these chapters provide an insight to the existing chromatic aberration correction, HDR imaging, and colour constancy solutions proposed in literature.

Chapter 4 presents a description of the framework used in the algorithms presented in contributory chapters that fit within an ISP under a common programme i.e. the improvement of image quality within an ISP pipeline that can be implemented in hardware.

Chapter 5 presents a multi-exposure image fusion algorithm suitable for practical implementation in hardware for video HDR content creation. This novel and optimised method is implemented as a pre-processing technique that works in the RAW Bayer domain and can be incorporated within an ISP. This chapter includes a literature review, experimental setup, results, analysis of experiments carried out, and explanation of the contributions made by the proposed algorithm.

Chapter 6 presents a novel colour management system within an ISP pipeline that aims to solve the colour constancy problem in digital imaging. This chapter presents a comprehensive literature review of the state-of-the-art colour constancy algorithms, a detailed explanation of the proposed colour management system, experimental setup, analysis of experiments carried out, conclusions reached, and descriptions of the fundamental colour constancy problems found in digital imaging.

Chapter 7 incorporates the research and algorithms proposed in this thesis to provide a solution to the fundamental limitations of current imaging devices in the areas of colour constancy and dynamic range. As a result, a novel spatially variant colour constancy algorithm within an ISP pipeline that solves the colour constancy problem in HDR imaging with multiple illuminants or reflectances in the scene is presented. This chapter includes a literature review, description of the problem to be solved, experimental setup, analysis of experiments carried out, and the conclusions reached.

Chapter 8 concludes the research presented in this thesis with an insight into the future directions of research and possible enhancements to the proposed algorithms.

Chapter 2

Background Theory and Fundamentals

This chapter covers the background theory and fundamental concepts of digital camera components as an introduction to dynamic range, and colour constancy in digital imaging. These concepts are required for a better understanding of the novel methodologies proposed in the chapters that follow and the ability of digital cameras to capture light.

For clarity of presentation, this chapter is organised as follows: In section 2.1, an introduction to the HVS and digital cameras ability to capture light in order to produce images is presented, followed by section 2.2 in which colour image reproduction of the HVS and digital cameras is explained. Section 2.3 presents an introduction to digital imaging sensor characteristics, and section 2.4 provides the background theory of image signal processors. Section 2.5 introduces the concepts of image quality in digital imaging and section 2.4 summarises all the sections with the insight to make conclusions and identifying open research problems in the selected areas of research.

2.1 Human Visual System and Digital Cameras

One of the main objectives in digital imaging is to mimic the capabilities of the human eye, and possibly, go beyond in certain aspects. The single most challenging feature of optical information is the high dynamic range of intensities hitting the same light receptor from one instant to another, or hitting different light receptors in an observed scene at the same time [9]. As shown in table 2.1, this range may exceed eight orders of magnitude. This is particularly challenging when a digital camera captures scenes with a wide range of colours and intensities.

Condition	Illuminance (lx)
Clear night sky	0.001
Quarter moon	0.01
Full moon	0.1
Late twilight	1
Twilight	10
Heavy overcast	100
Overcast sky	1000
Full daylight	10000
Direct sunlight	100000

Table 2.1: Optical intensities in photometric units in real-world scenes [9]

Human eyes can see luminance levels between 10^{-3} cd/m^2 and 10^8 cd/m^2, about 220dB [11, 12]. To achieve such a wide dynamic range, the eye has three mechanisms: photoreceptor cells, cones and rods, which correspond to photodiodes with different photosensitivity; a logarithmic response allowing saturation to occur very slowly; and by the dynamic adaptation of the logarithmic curve according to the average luminance levels or ambient light [12]. In contrast, typical digital imaging devices have a dynamic range of 60-70dB with a linear response rapidly reaching saturation [10].

The human eye has a remarkable dynamic range capability to detect subtle contrast variations and to interpret scenes under a large variety of illumination conditions [8]. In contrast, a digital camera can only capture a limited dynamic range of possible variations of brightness information. As a result, when a digital camera is used to capture a scene that exceeds the maximum dynamic range of the camera, the image captured ends up being too dark in some areas or possibly saturated in others. Computer graphics and related digital imaging applications continue to demand more realistic and accurate imaging reproduction, however today's digital cameras' inability to captured high dynamic range scenes is becoming a crucial shortcoming in digital photography, video recording, and computer vision applications.

2.2 Colour Reproduction in Digital Imaging

Visible light is a form of radiant energy that travels in space and is transmitted, absorbed, and reflected by different materials. In the case of the HVS, light that reaches photoreceptors inside eyes, produce visual sensations depending on the characteristics of the light spectrum. Human eyes are sensitive to light which lies in a very small region of the electromagnetic spectrum, "the visible light", which corresponds to wavelengths of 400 to 700nm, with colours ranging from violet through red [11]. Figure 2.1 shows the visible light spectrum of human eyes (image taken from [13]).

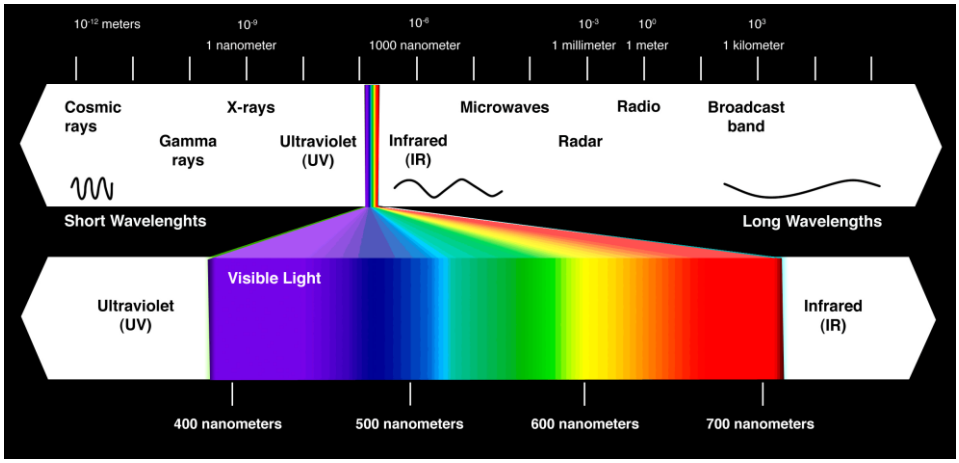


Figure 2.1: HSV Visible Light [13].

In contrast to the HVS, imaging sensors are sensitive to UV light, from 250nm, and a considerable amount of infrared light, with wavelengths up to 950nm [15]. For standard photography, capturing infrared light is not desirable because accurate colour reproduction would be impossible due to a rapid saturation in the red channel, focal point convergence, and spectrum crosstalk in the three primary colours [15]. Thus, if the aim is to mimic the human eye, sensors have to be fitted with IR-cut filters to narrow down the spectrum of light captured by the imaging sensor, and therefore match the HVS "Visible Light". Nonetheless, specialised digital cameras, mainly monochromatic, have been designed to exploit the fact of sensors being able to detect a broader spectrum than the HSV, and create camera applications for very low light, "night vision", or for medical purposes [15].

2.2.1 Colour image formation

Trichromatic theory dates back to the seventeenth and eighteenth centuries when John Dalton, Isaac Newton and Thomas Young, to name a few [18], started studying the HSV response to light, and characterised the human eye sensitivities as shown in figure 2.2(a). In HVS, the colour image formation process is through the integration of the responses of the three types of cone cells, which are responsible of detecting the spectra of the illuminant, the surface reflectance of the scene, and the sensitivity of the image acquisition system over the entire spectrum [16]. This can be denoted as:

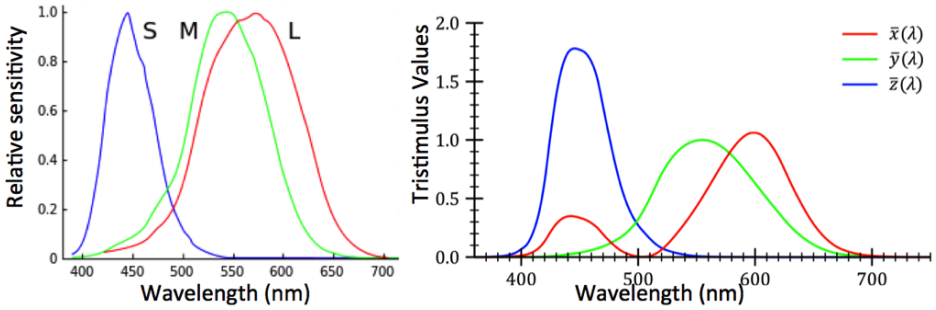
$$L = \int C(\lambda)L(\lambda)\,d\lambda, M = \int C(\lambda)M(\lambda)\,d\lambda, S = \int C(\lambda)S(\lambda)\,d\lambda \qquad (2.1)\ [125]$$

Where *l*, *m*, and *s* are the three types of cone cells in the eye, *L(λ)*, *M(λ)*, *S(λ)*, represent the relative sensitivities of the three cone cells, and *C(λ)* is the function of colour signal which is focused by the lens at the retina, and to which the light sensitive cone cells respond [17]. This same concept is used in digital images but in this case, *C(λ)* is focused into an imaging sensor with a Bayer RGGB CFA to emulate the cone cells in the human eye. The relative spectral sensitivity of an imaging sensor can then be formulated as:

$$X = \int C(\lambda)\bar{x}(\lambda)\,d\lambda, Y = \int C(\lambda)\bar{y}(\lambda)\,d\lambda, Z = \int C(\lambda)\bar{z}(\lambda)\,d\lambda \qquad (2.2)\ [125]$$

Where *X*, *Y*, *Z* are the tristimilus values defined in CIE XYZ space.

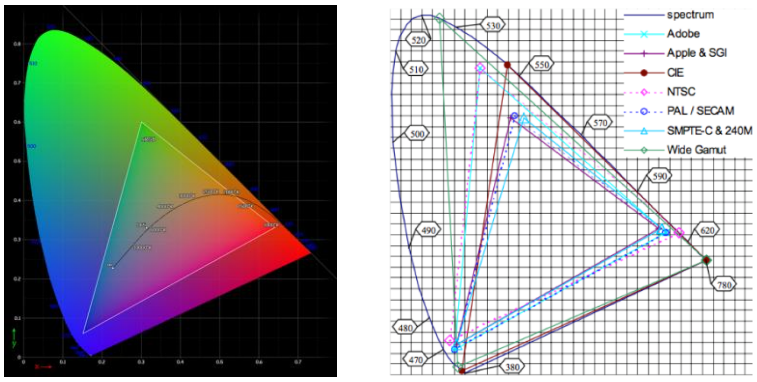
In 1931 the Commission International de I'Éclairage (CIE) proposed standards for colour representation based on the physiological perception of light, which are based on a set of colour matching functions related to red, green and blue cones in the eye as shown in figure 2.2(b) [17].



a) b)

Figure 2.2: a) The relative spectral response curves for the three types of cone cells. b) The relative spectral response curves for a typical trichromatic digital camera [17].

The colour matching functions give the amount of each primary colour necessary to generate a particular energy spectrum. Nonetheless, it is important to note that trichromatic theory has limitations and cannot accurately generate all the colours of the electromagnetic spectrum [19]. This fact is illustrated in figure 2.3(a), where the CIE 1931 chromaticity diagram is shown, along with the sRGB colour gamut space. Note that areas outside the sRGB gamut triangle cannot be accurately reproduced by the trichromatic theory. Many colour spaces have been proposed with the aim extending the accuracy of colour reproduction. Achieving all colours may not be possible without introducing more dimensions to the colour model [19]. Figure 2.3(b) shows examples of different gamut capabilities of RGB colour spaces.



a) b)

Figure 2.3: a) CIE 1931 chromaticity diagram showing the gamut of the sRGB colour space and location of the monochromatic colours. b) Examples of RGB colour spaces in the chromaticity diagram [19]

2.2.2 Colour Constancy

Colour constancy is the effect whereby the perceived or apparent colour of a surface remains constant despite changes in the spectral composition of the illumination and intensity levels [20]. The HVS is colour constant to a large extent, but this is affected by contents of the field of view, state of adaptation, intensity level of illumination and difference of the adapting chromaticity from different kinds of illumination conditions [92].

The HVS has evolved in such a manner as to maintain a dynamic white balance at all times. For example, if a room with white walls is illuminated by a tungsten light bulb, the walls may be perceived with a yellow tint but all other objects in the scene will maintain their respective colour balance, as if they were illuminated by day light. Edward H. Land demonstrated this effect by setting an experiment using “Mondrian” consisting of random sets of colour patches illuminated by different illuminants [21]. Colour constancy is assumed to be fundamental to colour vision, allowing functions ranging from the primitive foraging for ripe fruit in a tree, to the modern use of textiles, cosmetics, and packaging. In theory, colour constancy is necessary for producing a stable visual world [20].

Colour constancy in digital imaging refers to the ability of a digital camera to recognise and reproduce the colour of an object as perceived under some canonical illumination, such as white light with a flat spectrum, regardless of the illumination that is incident upon it [126]. Digital cameras have to rely on fast colour balancing algorithms, integrated into their ISP pipeline, to estimate the colour temperature of a scene illumination in terms of the camera sensor response. However, since colour constancy algorithms have to calculate the colour temperature by processing only the digital pixel values, the colour temperature calculation of the illuminant is ill posed, and even more, when dealing with the colour constancy problem in digital photography, it is necessary to take into account the colour preference of end users. Pleasing or memory colours, the colours that are in a person’s mind that are inseparable from certain common objects, may be preferred over accurate reproduction of colours [94]. Thus, the target is not to reproduce accurate representation of colours in an image, but to reproduce what the pleasing memory colours in the HSV are [94].

Achieving colour constancy in digital imaging is not only important for digital video or still photography in consumer cameras, it is also important for computer vision applications, especially if colour constant descriptors are used for object recognition. The work presented in Chapters 7 and 8 is intended so that colour constancy can be achieved in a controlled manner, and at the same time being able to produce pleasing-memory colours within an ISP in order to improve the overall image quality of a camera system.

2.3 Imaging Sensors

An imaging sensor is a solid-state device which captures light to form a digital image. Wafers of silicon are used as the base for the sensor's integrated circuit, which is built up via a process known as photolithography [3]. This process creates millions of tiny wells (pixels), which sense photons arriving at a particular location in the sensor.

Today's cameras mostly use charge coupled device (CCD) or complementary metal oxide semiconductor (CMOS) sensors. These two are the underlying technology of digital cameras. The main differences between CCD and CMOS sensors arise from their pixel readout architectures. In a CCD, charge is shifted out of the pixel array via vertical and horizontal CCDs. Then, the charge is converted into a voltage via a simple follower amplifier and serially readout before digitising the analogue signal. In a CMOS image sensor, charge voltage signals are readout one row at a time, in a similar manner to a random access memory, using row and column select circuits [1].

CCD and CMOS readout architectures have both their advantages and disadvantages. The main advantage of the CCD readout architecture is that it requires minimal pixel overhead, making it possible to design image sensors with very small pixel sizes, and since charge transfer is passive, CCDs have better temporal noise performance. Moreover, CCDs have a high output uniformity, which results in cleaner, high quality images [1]. However in CCD sensors, the charge transfer readout is serial, making the readout speed to be limited and requiring high power due to the need for high-rate, high-voltage clocks to achieve near perfect charge transfer efficiency [2]. CCD sensors are also distinct from CMOS is their use

of global shutters instead of rolling shutters. Global shutters process an entire image at once by exposing the full frame for a predetermined amount of time. This means that the entire sensor gathers an equal amount of light at once and therefore, images are free of distortion related to rapid movement or flashes of light.

In a CMOS image sensor, the readout path comprises several active devices that introduce both temporal noise and fix pattern noise; nonetheless, the random access readout of CMOS sensors provides the potential for high-speed readout and window-of-interest operations at low power consumption [1]. Unlike CCDs, CMOS sensors have circuitry at the pixel level. This extra circuitry crowds the pixels, limiting their ability to capture light and resulting in generally poorer visual image quality in comparison to CCDs [1].

Imaging sensors are colourblind, thus a mosaic of coloured filter array (CFA) has to be placed over the pixels in order to reproduce colour images. One of the most common filter arrays used in today's cameras, is the Bayer array, consisting of alternating rows of red-green and green-blue filters, known as the Bayer RGGB CFA [4]. Figure 2.4 shows an example of a sensor with a Bayer type CFA.

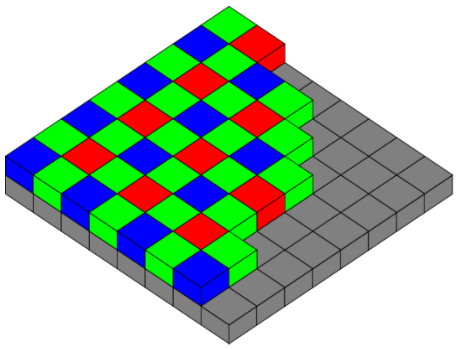


Figure 2.4: Image sensor with a Bayer type CFA [126].

The CFA, and the analogue to digital (A/D) components of a sensor, allow capturing digital colour images. However, this data still needs to be processed by an image signal processor (ISP) to interpret and enhance this data so that the final image appears identical to what a human eye can see.

2.3.1 Types of noise and dynamic range of imaging sensors

The accuracy of an imaging sensor in gathering photons and converting them into a digital signal is crucial for the quality of the recorded image. Any inaccuracies of an imaging sensor to capture light will constitute a degradation of signal, which is perceived as noise. There are several sources of noise in imaging sensors: read out, fixed pattern noise (FPN), temporal, thermal, pixel response non-uniformity, dark current, photon shot, and quantisation errors [6]. The combination of all these types of noise is considered to be the noise floor level of an imaging sensor.

Read out noise: read out is the type of noise caused by voltage fluctuations in the sensor circuitry that contributes to a deviation of the digital pixel value (DN) from the ideal DN proportional to the photon count [6].

Fixed pattern noise: The primary cause of FPN in image sensor is dark current non-uniformity [6]. If the dark current of each pixel is not uniform over the whole pixel array, the non-uniformity is seen as FPN because correlated double sampling (CDS) cannot remove this noise component.

Temporal noise: Temporal noise is a random variation in the signal generated by the sensor that fluctuates over time [6].

Thermal noise: Thermal agitation of electrons in an imaging sensor can liberate a few extra electrons [7], which cannot be distinguished from the electrons freed by photon absorption. Therefore, this effect causes a miscount of the photons count represented by the DN. Thermal electrons are freed at a relatively constant rate per unit time, and increases with exposure time [7].

Pixel response non-uniformity: this type of noise is presented as the fluctuation of the DN of output pixels at different exposure times due to the photon to output voltage conversion, i.e. nonlinear response of pixels at different exposure times.

Quantisation error: When analogue signal is converted into a digital signal, the analogue signal is rounded to a nearby integer value and this produces a rounding error of the actual signal. In practice, this is a minor contribution to the total noise of a sensor [7].

Dark Current: Dark current, also known as black level, is the current that is integrated as dark charge at a charge storage node inside a pixel. The amount of dark charge is proportional to exposure time and is also a function of temperature [6].

In order to define the dynamic range of a sensor, it is necessary to know the full well capacity, also known as the saturation charge of a sensor [6]. The full well capacity indicates the maximum amount of charge that can be accumulated on a photodiode capacitance. The dynamic range (DR) capability of a sensor is measured as $DR = 20\log(S/N)$ in decibels. Where S is the saturation point and N is the root mean square readout noise floor measured in electrons or volts [10].

Figure 2.5 shows an example of the photoconversion characteristics taken from [6], illustrating total dynamic range, signal, photon shot noise, and read noise (noise floor) as a function of incident photons. In the plot shown in Figure 2.5, it is assumed that a virtual image sensor with a pixel size of 25μm^2, a conversion gain (C.G.) = 40 $\mu V/e^-$, a full-well capacity of 20,000 electrons, a noise floor of 12 electrons, and a detector's quantum efficiency QE of 0.5. Dark current shot noise is not included in the plot. In this example the dynamic range of the sensor equals 20log(20000/12) = 64.4 dB

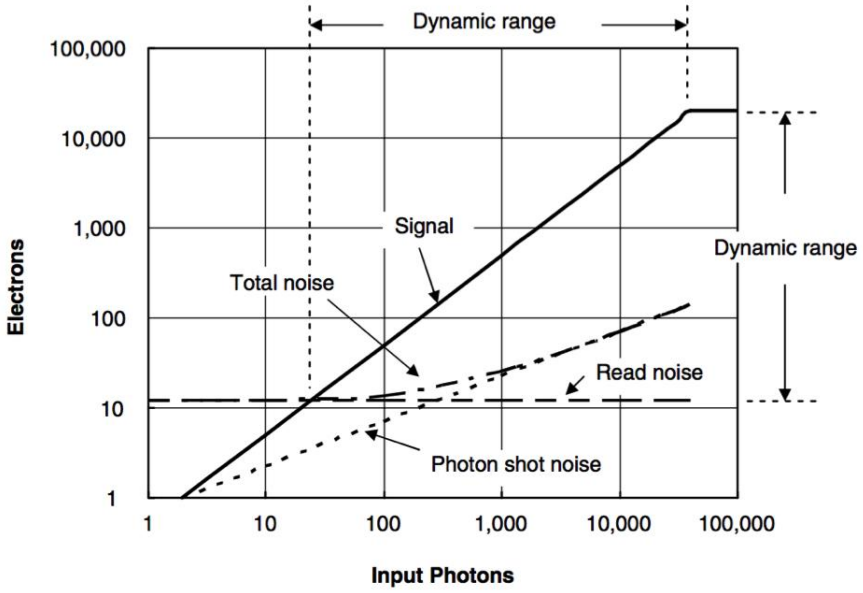


Figure 2.5: Example of photoconversion characteristics [6].

2.4 Image Signal Processors

An image signal processor is responsible for enhancing and translating the Bayer RAW data output of an imaging sensor into a final image with full colour information at each pixel. The ISP processes have to be done in real time and is often implemented as an integrated component of a system-on-chip (SoC) image processor. This allows frontend image processing to be completed without placing any processing burden on the main application processor.

An ISP pipeline is often composed of advanced image processing algorithms in order to enhance the overall image quality of a system. This is required since artefacts are introduced by the camera optics, sensor spectral characteristics, and noise performance. Some of the processes used for enhancing image quality within an ISP are: linearisation, hot pixel correction, noise reduction, green equalisation, demosaicing, dynamic range compression, colour correction, vignetting correction, black level subtraction, sharpening, cross talk removal, colour space conversion, down sampling, frame stitching, digital image stabilisation, and gamma correction. Moreover, the ISP is responsible for the real time

automated controls (3A algorithms) of a digital camera. The 3A algorithms are: auto exposure, auto white balance, and continuous auto focus [5]. In addition to removing artefacts introduced to the image by the optics and imaging sensor characteristics, the ISP is also used for improving the lack of functionality that an imaging sensor can provide. That is, achieving colour constancy, panoramic imaging, lens distortion correction and extending the dynamic range.

Figure 2.6 shows an example of an ISP and its modules. The order of the ISP modules mentioned above completely depends on implantation trade-offs and flexibility of the system. Factors such as type of application, image quality, frame rate, power consumption, bandwidth, gate count, memory requirements, and time-to-market considerations usually determine the placements of each of the ISP modules used in digital cameras [6].

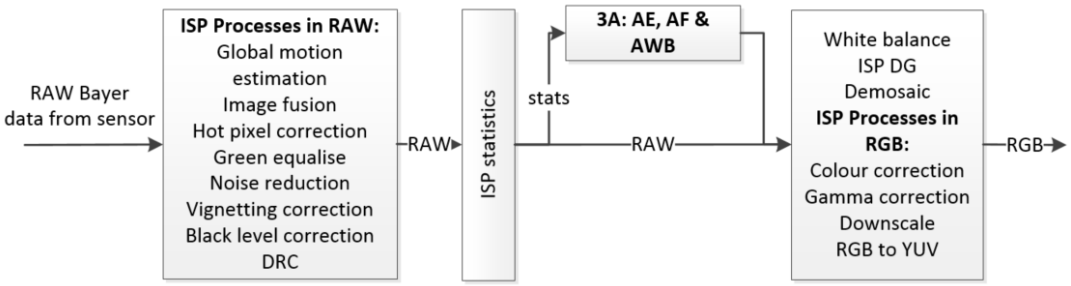


Figure 2.6: ISP block diagram example.

The image quality of digital cameras completely depends on the lens, sensor and the ISP's ability to produce high quality images. However, the ISP plays a main role in achieving the best possible image quality of a digital camera, even if the optics and imaging sensor used are not the best of their kind. The work presented in the following chapters is intended so that all of them can be directly incorporated within an ISP to improve the overall image quality of a camera system.

2.5 Image Quality Assessment

Digital images are subject to a wide variety of distortions during acquisition, processing, compression, storage, transmission and reproduction, any of which may result in degradations of the perceived image quality [22]. Previous research in image quality has shown that if the ultimate goal of digital imaging is to be viewed by humans, the only "correct" method of quantifying image quality is through a subjective evaluation [22]. However, studies have also shown that subjective image quality evaluation is usually inconvenient, time consuming, and to a great extent difficult to measure due to different societies' culture and geography of human individuals' visual preferences [22, 23]. This fact has led researchers to develop objective measurements in order to measure quantitative image quality measurements that can automatically predict, monitor, optimise, and benchmark the perceived image quality by the HVS.

There are two key aspects of image quality [24]. i) Factors affected by post-processing. Contrast, colour balance, and colour saturation. These are the factors that can be enhanced in a darkroom (digital or chemical) under the control of the photographer. For example, higher contrast images usually score higher in visual quality assessment tests, unless the image is clipped. Because of their great importance to photo finishers these factors are the focus of much image quality literature, but they are outside the scope of this research. ii) Factors intrinsic to cameras, lenses, and printers. To mention a few: sharpness, noise, dynamic range, chromatic aberrations, tonal response, exposure accuracy, lens distortions, contrast, and colour accuracy. The work presented in this thesis focus on enhancing these image-quality factors in images captured by a digital camera.

Sharpness: Sharpness is arguably one of the most important single image quality factors since it determines the amount of detail an image can convey. Sharpness can be measured as a spatial frequency response (SFR), also known as modulation transfer function (MTF). MTF is the contrast at a given spatial frequency (measured in cycles or line pairs per distance) relative to low frequencies. MTF is measured with a slanted-edge SFR chart, or using sine pattern charts that increase in frequency in a logarithmic manner [24].

Dynamic range and tonal response: Dynamic range is a strong function of pixel area, which is proportional to the number of electrons a pixel can store. Dynamic range is usually measured in decibels, and it is closely related to noise, the tonal response, and contrast. High contrast usually involves loss of dynamic range when an image is displayed. Dynamic range, tonal response, and contrast can be measured by a step chart using a transmission step wedge or by a postprocessor for step chart that uses results from up to four differently-exposed reflective step chart images [24].

Noise: noise is a random variation of image density, visible as grain in film and pixel level variations in digital images. It arises from the effects of basic physics, the photon nature of light and the thermal energy of heat inside image sensors [24]. Noise has a clear relationship to a signal, thus a signal to noise ratio (SNR or S/N) is often calculated. SNR can be defined in many ways, depending on how the signal *S* is defined [24]. Noise can be measured by using a step chart, which produces the most detailed results, or by using ColourChecker, SFR, SFRplus, and Uniformity charts [24].

Chromatic aberrations: Chromatic aberration is one of several aberrations that degrade lens performance. Chromatic aberration causes colours to focus at different distances from the image centre, and it is most visible near corners of images. Chromatic aberration can be measured by SFR, using edges at a distance from the image centre [24].

Colour accuracy: Colour accuracy is an important but ambiguous image quality factor. Many viewers prefer enhanced colour saturation; studies have shown that accurate colour reproduction is not necessarily the most pleasing [94]. Nevertheless it is important to measure a camera's colour response: its colour shifts, saturation, and the effectiveness of white balance algorithms. Colour accuracy is measured by using the widely available 24-patch GretagMacbeth ColourChecker.

Image quality assessment plays an important role in digital imaging applications. A great deal of effort has been made in recent years to develop objective image quality metrics that correlate with perceived quality measurement in the HVS [23]. Unfortunately, only limited

success has been achieved since evidence showed that human visual error sensitivities and masking effects vary in different spatial and temporal frequency and directional channels.

2.6 Conclusions

This chapter reviews some of the concepts including: digital imaging and the HVS relationship; dynamic range and colour perception; imaging sensors characteristics; image signal processors; and image quality metrics used in digital imaging. These concepts are fundamental to digital video and photography, and will be used throughout this dissertation.

Computer graphics, computer vision, and related digital imaging applications continue to demand more realistic and accurate imaging reproduction with high dynamic range scenes and colour constancy. As explained in the sections of this chapter, dynamic range and colour constancy are becoming a crucial shortcoming in digital photography, video recording, and computer vision applications. For this reason, the research carried out and presented in this dissertation aims to solve the problem of dynamic range and colour constancy in digital imaging.

Chapter 3

Literature review

The literature review presented in this chapter provides an insight into the existing chromatic aberration correction, HDR imaging, and colour constancy solutions proposed in literature. These topics cover the contributory subject areas of this thesis presented in Chapters 6, 7 and 8. Furthermore, this chapter reviews a number of fundamental concepts and theories in different areas of mathematics, computing and engineering that are used throughout the work presented in this dissertation.

For clarity of presentation, this chapter is organised as follows: Section 3.1 focuses on the review of previous work on chromatic aberrations corrections. Section 3.2 covers multi-exposure image fusion and image registration techniques, and section 3.3 presents a review of colour balancing methods found in literature that aim to solve the problem of colour constancy in digital imaging.

3.1 HDR Imaging Techniques

HDR imaging is the field of imaging that seeks to accurately capture and represent scenes with the largest possible intensity levels. Due to technological and physical limitations of current imaging sensors, the most common way to capture high dynamic range scenes is by combining multiple low dynamic range photographs through an image fusion process. The multi-exposure images are acquired with different exposure times through a process known as exposure bracketing [16].

Image fusion has been used in a wide variety of applications such as machine vision, automatic change detection, biometrics, medicine, remote sensing, etc. In general, the goal

of image fusion is to integrate multiple images so that the resulting image is clearer and more intelligible.

In 1994, Mann and Picard proposed one of the first exposure bracketing based HDR imaging techniques in digital photography [26]. This method assumes that the camera radiometric response function, *f*, of the camera has already been calibrated, in order to map intensity levels of the scene to the pixel values. Then, if z_i^p denotes the image value at pixel *p* for exposure time τi, the irradiance estimate at pixel *p* is computed as the weighted average of $f^{-1}(z_i^p)/\tau i$, *i = 1,…,T.* With this function, small weights are assigned to pixel values with very low or close to saturation values. This approach suffers from two main problems. The first is that the weights are arbitrarily determined and they are not derived from an imaging sensor noise model. The second is that the calibration of the camera response function is a problem on its own, and is prone to errors that are directly transferred to the irradiance estimator. Nonetheless, Mann and Picard's work is an important contribution, since it inspired several approaches based on exposure bracketing, whose main difference relies in the way the multi-exposure images are combined. Kay [27] and Kirk [28] present a comprehensive review and comparison of these methods.

HDR image fusion techniques can be classified in categories according to the nature of data where the image fusion process takes place. The first type of method assumes a linear camera response, and are to be used with the camera RAW Bayer data at an early stage of an ISP pipeline. [e.g., see 28, 33, and 34]. The second type of image fusion techniques, assume a non-linear camera response, gamma corrected RGB images, in which the inverse of the camera response has to be estimated. [e.g., see 26, 29, 30, and 31].

In addition to the nature of data, image fusion techniques can be classified into another three different classes: pixel level, feature level and decision level. Pixel level fusion is the combination of pixels from multiple sources into a single resolution image. Feature level fusion extracts features, e.g. edges, corners, lines, textures, etc., from different image sources and then combines them into one final image. Decision level fusion combines the results from multiple algorithms to yield a final fused rule. Methods of decision fusion include voting, statistical, and fuzzy logic based methods.

The above categorisation does not encompass all possible fusion methods, since input and output of data fusion may be different at different levels of processing. In practice, fusion procedures are often a combination of the three classes mentioned above. The advantage in pixel level fusion is that images contain original data and therefore the pixel information is preserved. On the other hand, region and decision based methods are very useful because real world objects usually consist of structures at different scales as in the human visual system [71-77].

Most image fusion techniques presented in the literature share the same shortcomings. In order to produce artefact free images, the fusion technique has to be able to compensate for motion caused by camera shake and object movements in the scene. Usually, during the fusion process, ghosting artefacts are generated due to the fractionally time difference instances of the objects' displacement within the multi-exposure images captured. Thus, the ability of registering images prior to the image fusion process is extremely important to produce high quality images.

A solution for image registration and removal of ghost artefacts was proposed by Greg Ward [31, 70]. This method is based on a median threshold bitmap (MTB), and variance estimation between the pixels of each exposure to create a segmentation mask used during the weighting average image fusion technique. This solution has demonstrated to be efficient to remove ghost artefact and recover small displacements between multi exposure images. However, this method is slow due to the iterative estimation of the weights assigned to each pixel and fails if the scene does not predominantly capture a static background.

3.2 Colour Balancing Methods and Colour Constancy

The review of colour constancy methods in this section is by no means exhaustive. This literature is intended as a summary of all major contributions on the colour constancy field. This allows the drawing of conclusions and identifying the shortcomings in achieving colour constancy in digital imaging.

Other researchers have tried to solve the problem of colour constancy in digital imaging by proposing a number of algorithmic and instrumentation approaches. Nevertheless, no unique solution has been identified; this may be due to the wide range of computer vision applications that require colour constancy and the complexity of the HSV to achieve colour constancy [97]. This led researchers in the field to identify sets of possible solutions that can be applied to particular problems in computer vision. Predominantly, previous efforts have been directed towards colour constancy approaches that can be used in real time video and still photography.

The goal of colour constancy is to achieve an illuminant invariant description of a scene taken under illumination whose spectral characteristics are unknown [97]. Colour constancy can be achieved in a two-step process. First, an estimate of the illuminant parameters is obtained; followed by the illuminant independent surface descriptor, which is parametrically calculated [35, 36, and 37]. Often illumination invariant descriptors of the scene are estimated under an illumination whose spectral characteristics are known; this is referred as the canonical illumination, which is the illuminant used by the camera to colour balance an image [38].

In Chapter 2, Equations 2.2 introduced the formula of the relative spectral sensitivity of an imaging sensor. Equation 2.2 can be rewritten in terms of a colour image with *k* number of colour channels as a product of three variables:

$$E_k(x, y, \lambda) = \int_{\omega} R\,(x, y, \lambda) L(\lambda) s_k(\lambda) d\lambda \qquad (3.11)[97]$$

Where $R(x, y, \lambda)$ is the surface reflectance, $L(\lambda)$ is the illumination property, and $S_k(\lambda)$ is the sensor spectral response, as a function of the wavelength λ, over the visible spectrum ω, and $E(x, y, \lambda)$ is the image corresponding to the k^{th} channel, $k \in \{R, G, B\}$. In the colour constancy literature, efforts are directed toward discounting the effect of illumination and obtaining a canonical colour appearance.

Hadamard, a French mathematician, stated that a problem is well posed if the following three conditions are satisfied, 1) there is a solution, 12) this solution is unique, and 3) this unique solution is stable. If any of these conditions are not satisfied, then the problem is known as ill posed [39]. Thus, achieving colour constancy is an ill posed problem since the uniqueness and the stability of the solution cannot be guaranteed due to the high correlation between the colour in the image and the colour of the illuminant. This non-independence of predictor variables directly affects the estimation of the illuminant. That is, imprecise estimation and slight variations of the predictor variables may lead to a large error in the illuminant estimation process.

The most widely accepted approach for achieving colour constancy in digital imaging is to first estimate the prevailing incident illuminant, and then apply a transformation to map the imaging device colour response to a standard colour space. The key to this process in achieving colour constancy is the illuminant estimation, since the second step is relatively straightforward. Thus, a lot of research has been done on the illuminant estimation problem.

Early solutions to the colour constancy problem focused on attempting to reduce the discrepancy between the number of known and unknowns in Equation 3.11, by adopting linear model representations of lights and surfaces [38-42]. These approaches set out the theoretical conditions under which the colour constancy problem can be solved uniquely. Unfortunately, for the case of a trichromatic device, the required conditions on lights and surfaces are not satisfied in most typical images. In general, illuminant estimation algorithms can be broadly classified as being either statistical or physics-based. Statistics-based approaches arrive at the estimated illuminant based on some sort of assumption about the distribution of colours in a scene. Physics-based approaches find the illuminant estimation by exploiting some artefacts of the underlying physics of image formation [16].

Statistical methods: within the statistical methods, simple and robust methods such as Grey-word, Max RGB, and Database Grey-world, work on the assumption that there must by a grey object in the scene in order to estimate the illuminant of the scene. These approaches are not perfect but generally give reasonable and consistent results [43]. For these reasons,

Grey-world based methods are commonly used in the literature as a benchmark when evaluating colour constancy algorithms.

There are three broad classes of colour constancy statistical methods: Gamut Mapping methods [38, 44-47], Bayesian methods [48-54], and Neural Networks [55]. The gamut mapping solution is one of the most well founded solutions to the colour constancy problem. Gamut mapping implementation is non-trivial and not always robust to real image data since these methods were developed using synthetic data. Bayesian methods, also known as probabilistic methods, share in common the fact that they attempt to capture information about the likelihood of observing a given RGB spectral response in the form of a statistical prior, and with that information chose the illuminant with the highest likelihood to colour balance an image. All Bayesian approaches operate according to the same underlying principle but they differ in terms of how they encode the prior knowledge, implementation details, and complexity. The neutral networks solution uses a large training set of images with known illuminants to learn the relationship between observed image data and scene illuminant. The shortcoming of neural networks approaches is their robustness to estimate the illuminant of the scene, and the training process where a large and general enough data set is needed to accurately predict any unseen scene.

Physics-based methods: Physics-based approaches [56-60] adopt a more general model of image formation than Equation 3.11, and estimate the illuminant of the scene by exploiting knowledge about the physical interaction between light and surfaces. Physics-based approaches are based on the observation that the sensor response from different points on a given surface will all fall on a plane in the sensor RGB space. That is, for a fixed surface $S(\lambda)$ and illuminant $E(\lambda)$ responses, defined in equation 3.11, are constrained to a two-dimensional subspace (a dichromatic plane) of the 3D RGB space [43]. This is an elegant solution to the illuminant estimation of a scene, however, these kinds of method rarely work in practice. This is because of the assumption that any scene can be reliably segmented into regions corresponding to the underlying surfaces, and determining the dichromatic planes become extremely problematic. Moreover, even if dichromatic planes can be determined, intersecting $S(\lambda)$ and $E(\lambda)$ to obtain a reliable estimate of the illuminant is often difficult and

inaccurate [43]. These difficulties have restricted the practical application of physics-based algorithms.

According to the literature review, there is not a comprehensive solution to the colour constancy problem, and methods proposed in the literature have not been able to demonstrate an understanding of the issue of preferred versus neutral colours rendering. This is understandable given the subjective nature of the problem, and thus the reason for not being able to identify a unique solution to the colour constancy problem.

The colour constancy algorithms proposed in literature have proved to perform well as a post-processing technique. However, most of these methods have never been tested as a part of an ISP that works in RAW Bayer domain. In addition, researchers have overlooked the fact that the images used in their experiments had some kind of colour correction processing before they applied their methods. Some of the processes are: static white, colour balancing (CCM), vignetting, demosaicing, gamma correction, and black level correction. This leads to a level of uncertainty, as to the ability of the proposed methods to correct colour casts on an already colour balanced image, or actually estimate and compensate colour shifts due to different colour illuminants. Furthermore, after analysing the image databases available at colorconstancy.com [105], which are the databases used by the previously proposed colour constancy approaches, none of them showed to have tested the methods to achieve colour constancy in challenging situations, also known in industry as “corner cases.” For instance, scenes with: monotonic colours, green grass colour versus CWF lighting, opposite colour objects’ surface reflectance to the light colour temperature in the scene, sunset, low light conditions, absence of grey, noisy images, mixed lighting, cloudy scenes, and scenes with a predominant colour.

3.3 Summary and Discussion

This chapter provides an insight into existing techniques for chromatic aberration correction, HDR imaging, and colour constancy, and reviews a number of fundamental concepts and theories in different areas of mathematics, computing and engineering that are used in this

thesis. Moreover, this literature review identified shortcomings and areas of improvements in chromatic aberration corrections, HDR imaging and colour constancy in digital imaging.

Chapter 4

Image Signal Processor Framework

This chapter presents a description of the framework used in the algorithms presented in later chapters that fit within an ISP under a common programme i.e. the improvement of image quality within an ISP pipeline that can be implemented in hardware.

4.1 Image quality and image signal processors

The quality of a picture captured with a digital camera is determined to a large extent by the quality of the lens and the image sensor. The progress in digital signal processing and in the understanding of the physics behind the limitations of the components of digital cameras, have allowed many artefacts introduced by the optics and imaging sensors to be corrected in the digital domain within an ISP.

There are several kinds of artefacts that degrade the quality of the image captured. These artefacts are due to imperfections added to the image caused by the lens, colour sampling, illumination, physical characteristics of image sensors, and artefacts added to the image in the digital domain, such as quantisation errors and compression artefacts. Image quality standards, either industry or academia, have not been able to classify all artefacts into certain categories, since a lot of them could overlap or be introduced by a combination of factors. For example, resolution and aliasing artefacts are related to the quality of the lens, the demosaicing process, pixel count, and sensor architecture.

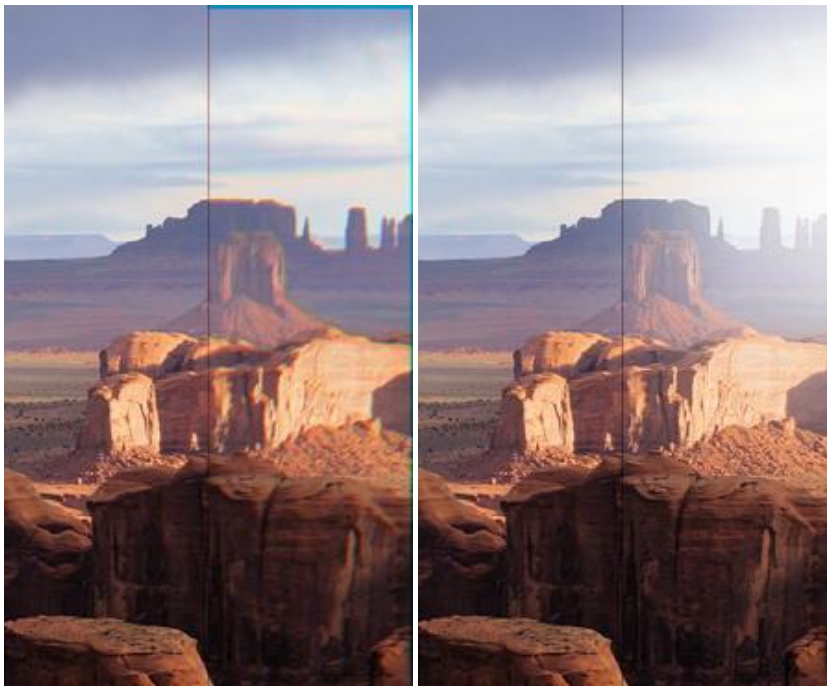
A review of image quality artefacts and how they can be corrected within an ISP is presented in the following sections.

4.1.1 Image quality artefacts introduced by the camera components

Lens artefacts: The main problem with the lenses used in consumer digital cameras is the vignetting, also known as shading or light falloff artefact. Vignetting is the effect where intensity levels drop off towards the edges of the image.

Another common lens artefact is the veiling glare, which causes loss of contrast due to light scattering within the lens elements and the inside barrel of the lens. This causes light reflections that affect the total dynamic range of cameras because intensity levels in shadow areas are increased, and sensor saturation levels are rapidly reached.

Other common artefacts usually seen on wide angle lens are chromatic aberrations and lens distortions. Chromatic aberration occurs because the index of refraction of glass varies with the wavelength of light, causing undesired colourisation of edges in the image. Lens distortions artefacts are aberrations that cause straight lines to curve near the edges of images. Figure 4.1 shows examples of image quality artefacts introduced by the optics of the camera.



Original | chromatic aberrations Original | veiling glare

Original | vignetting Original | barrel distortion

Figure 4.1: Examples of lens artefacts [62].

Colour sampling and illumination artefacts: The absorption coefficient of silicon used in imaging sensors strongly depends on the wavelength of the incident light [61]. In principle, the wavelength of the incident light characteristics can be used to separate the incoming photons into the three colour planes; red, green, and blue. However, colour separation of silicon itself is not selective enough to allow colour imaging even in low light conditions [61]. For this reason, colour separation is typically done by means of a CFA on top of every pixel of the sensor, such as the Bayer CFA. Colour images captured using a Bayer CFA rely on interpolation of data to reconstruct all colour channels. The combination of Bayer CFA light crosstalk and the demosaicing process, which interpolates all the missing pixels to reconstruct RGB images, introduce several artefacts: blurred or over sharpened images, Moiré, false colours, and zippering [63]. These kind of artefacts are shown in Figure 4.2.

The human eye interprets several combinations of red, green, and blue light as white, this effect is known as metamerism, something that does not happen in imaging sensors due to the fact that the sensor does not have equal quantum efficiencies for the various spectral bands [61]. In contrast to the human eye, image sensors perform an absolute measurement of the incoming light information [61], and thus, digital cameras have to rely on white balancing and colour correction processes to achieve colour constancy in digital imaging.

Original Zippering

a)

Original False colours

b)

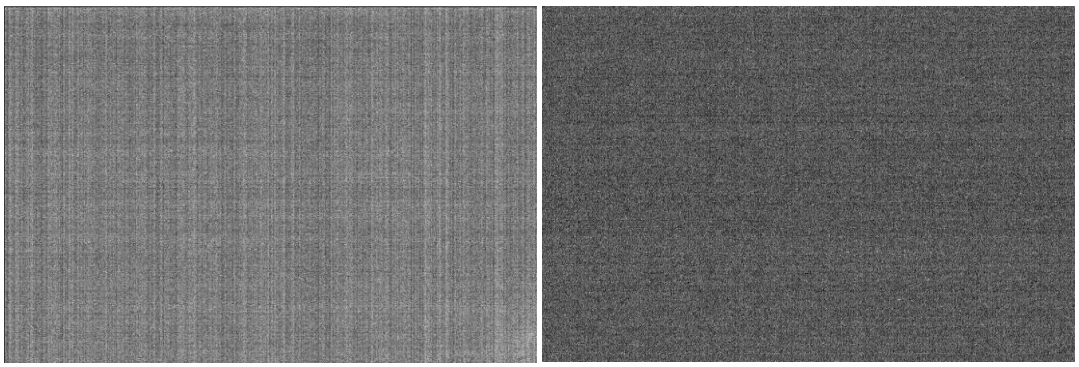
Original Moiré

c)

Figure 4.2: Examples of image artefacts caused by colour sampling [64]. In these examples the original image is mosaiced to RGGB pattern, and bilinear demosaiced to RGB.

Image sensor artefacts: Most of the sensor artefacts arise from noise sources and pixel non-uniformities reviewed in section 2.3. In still digital imaging, the temporal noise is frozen at the time of capture, and as a result, temporal noise becomes a second form of FPN. Most of the FPN in CCDs is random and it originates from dark-current non-uniformities. For CMOS

imagers the dark current is higher than for CCDs, this makes the contribution of dark current to FPN higher. Moreover, the readout architecture used in CMOS generates extra components of FPN: some of it is random, and some of it is fixed in the form of row or column-wise FPN. Figure 4.3(a) shows an example of row-column FPN. The residual noise after correction of FPN is mainly composed of photon shot and temporal noise produced by an image sensor, this is illustrated in Figure 4.3(b).



a) b)

Figure 4.3: Examples of noise produced by imaging sensors [67]. a) row-column FPN, b) residual noise after correction of row-column FPN

Other types of artefacts introduced by image sensors are the dead and hot pixels. Dead pixels are the pixels that are not sensitive to light at all, and hot pixels are the pixels that constantly report full saturation [61].

4.1.2 Artefacts correction within an ISP

The understanding of the nature and dependencies of noise artefacts and lens aberrations has helped the development of advanced image processing techniques to remove artefacts and improve the overall image quality in digital imaging.

Most of the artefacts described above can be removed or counteracted after the readout of the sensor within an ISP. Unfortunately, every time that the sensor data is manipulated, either in the analogue or digital domain, there is a possibility of decreasing the dynamic range of the system. In this case, a loss in total dynamic range has to be accepted.

Many image processing algorithms that aim to improve image quality have been proposed in the literature, however most of these algorithms work in RGB colour corrected non-linear data, and cannot benefit from the advantages of working in RAW Bayer domain. Some of the advantages of working in RAW Bayer domain are: linearity of data, noise profile characteristics, higher data precision, reduced bandwidth, reduced memory requirements, and allowing a set of algorithms to work in conjunction within a pipeline to improve the overall image quality produced by a digital camera. For this reason, the work presented in this thesis is aimed to be implemented as part of an ISP pipeline to improve two of the most challenging aspects of image quality in digital cameras: dynamic range, and colour constancy.

Ultimately it remains true: "to make a good image, one has to take a good image" [61]. But along the road from photons to digital numbers, the signal can pass through several calculation cycles to improve the quality of the end result. In order to remove the described artefacts, an ISP can adopt the following strategies.

Correction of lens artefacts: Vignetting artefacts can be removed by pre-programmed look up tables that contain the lens vignetting profile. In this profile all red, green, and blue channels are multiplied by a constant value that becomes gradually larger than 1 towards the outer edges of the image, this allows equalising the intensity levels throughout the image.

In a recent study [66], a novel image warping algorithm that works in RAW Bayer domain was proposed to solve the chromatic aberration problem. The advantage of correcting chromatic aberrations before demosaicing, is that the demosaic interpolation process can be performed in a more accurate away so that digital cameras can produce high resolution images.

Colour sampling and illumination artefacts: Camera systems that use sensors with a Bayer CFA depend on the demosaic process to reconstruct the pixels that are not sampled for each of the colour channels. Demosaic interpolation algorithms can be very sophisticated, and have adaptive interpolation schemes in which information in the vicinity of the pixel under

consideration is analysed and interpolated in a specific direction: vertical, horizontal, or diagonals [68]. For example, if the image contains objects with vertical lines, the interpolation of those specific lines would be done only along the pixels that contribute to the vertical direction of that line.

Due to the sparse sampling grids of the three colour channels and demosaic process, the resulting final image could be spatially blurred, as if it was processed with a low pass filter. To overcome the blurred appearance of the image, sharpening algorithms such as “unsharp mask” [69] can be used. The combination of advanced demosaicing and sharpening processes help to remove the colour sampling artefacts presented in Figure 4.2.

In order to cope with imperfections of a Bayer CFA, and spectral sensitivity differences between pixels, the RGB image obtained after demoisaicing needs to be corrected. These processes are known as white balancing and colour correction. In Chapters 6 and 7, a novel solution to the colour constancy problem in digital imaging is proposed to address the challenges of white balancing and colour correction in digital cameras.

Correction of image sensor artefacts: The column FPN can be measured by reading one or multiple dummy lines from the sensor. These values can be stored in digital memory in the ISP and subsequently apply the gain and offsets measured between the various column amplifiers. Since FPN is relatively constant over time, the stored digital values can be used to correct real image data [61]. Alternatively, FPN and dark current related artefacts can be corrected by subtracting a reference dark frame stored in the ISP digital memory.

Temporal and spatial noise produced by the imaging sensor can be removed within an ISP pipeline by accumulation of frames and block matching technologies that use the sensor noise profile characteristics [82].

Defect pixels can be corrected by means of look up tables. In these look up tables the XY locations of the defect pixels are stored so that the ISP interpolates a new value for each of the defect pixels by using the knowledge of the neighbouring “good” pixels within the same colour plane [65].

Dynamic range of a sensor, which is defined by the signal to noise ratio, can be improved if the sensor artefacts, noise floor, is decreased and pixel saturation level is increased. This can be achieved by the multi-exposure image fusion technology presented in Chapter 5.

4.2 Image signal processor architecture

The ISP pipeline of a digital camera is designed to exploit the processing power of dedicated hardware so that image processing algorithms can be executed to enhance the quality of image frames at very high rates per second. This results in an extremely efficient pipeline with deterministic performance which allows the increase of speed in which images can be processed, and therefore the rate at which pictures and video can be captured.

The research presented in Chapters 5, 6, and 7 covers HDR imaging and colour constancy solutions. These novel solutions share in common that all of them work in RAW Bayer domain as a part of an ISP pipeline to exploit the benefits of the RAW Bayer data, and computational power to address the artefacts mentioned in section 4.1. This strategy allows the algorithms proposed in this thesis to improve the overall image quality produced by digital cameras in real time.

In order to incorporate and demonstrate that the proposed HDR and colour constancy solutions can be implemented in dedicated hardware and used as a part of a digital camera, an ISP pipeline architecture is defined for their hardware implementation and testing. The top level architecture of the imaging system on chip (SoC) proposed is shown in Figure 4.4. In this architecture, synchronous dynamic random access memory (SDRAM) is used as an external frame buffer for the temporal noise reduction, HDR image fusion, or to stream Bayer RAW data directly to the encoder without passing through the ISP. Static random access memory (SRAM) is internally used by the ISP modules for storing pixel line memories, configuration parameters, and look up tables. The CPU in this architecture, plays an important role since it is used as the main configuration controller that drives the camera system.

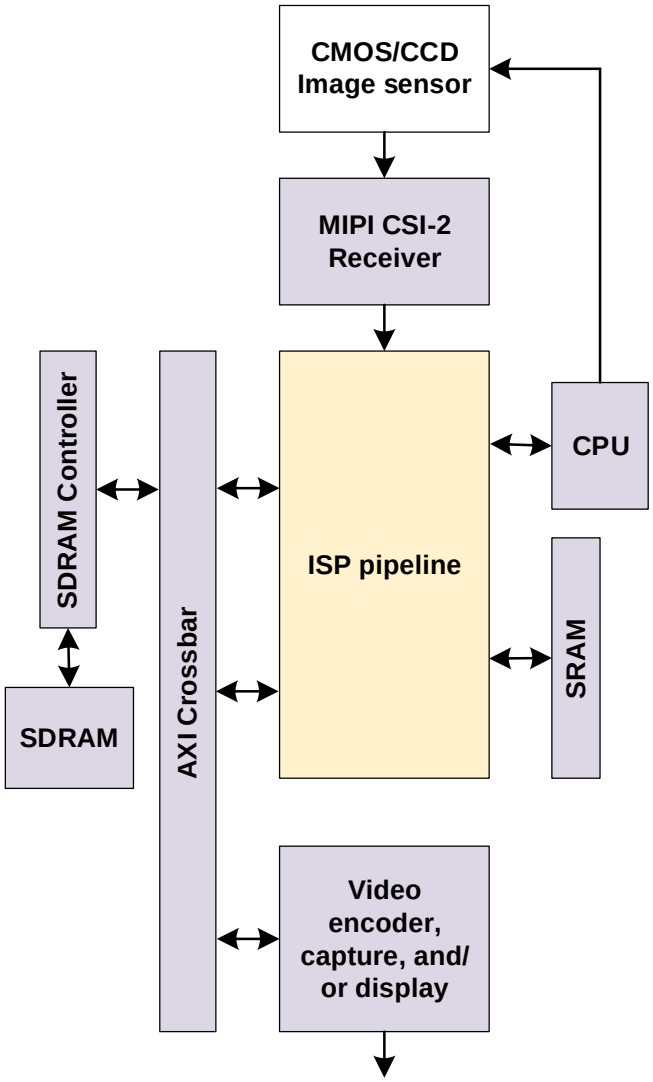


Figure 4.4: Top level architecture of an imaging SoC imaging solution

The top level ISP architecture block diagram is presented in Figure 4.5. The ISP architecture defined is an ISP solution with the minimum required modules for the testing of the algorithms developed in Chapters 5, 6, and 7. The placement of these algorithms within the ISP pipeline are highlighted in purple in the top level architecture diagram presented in Figure 4.5.

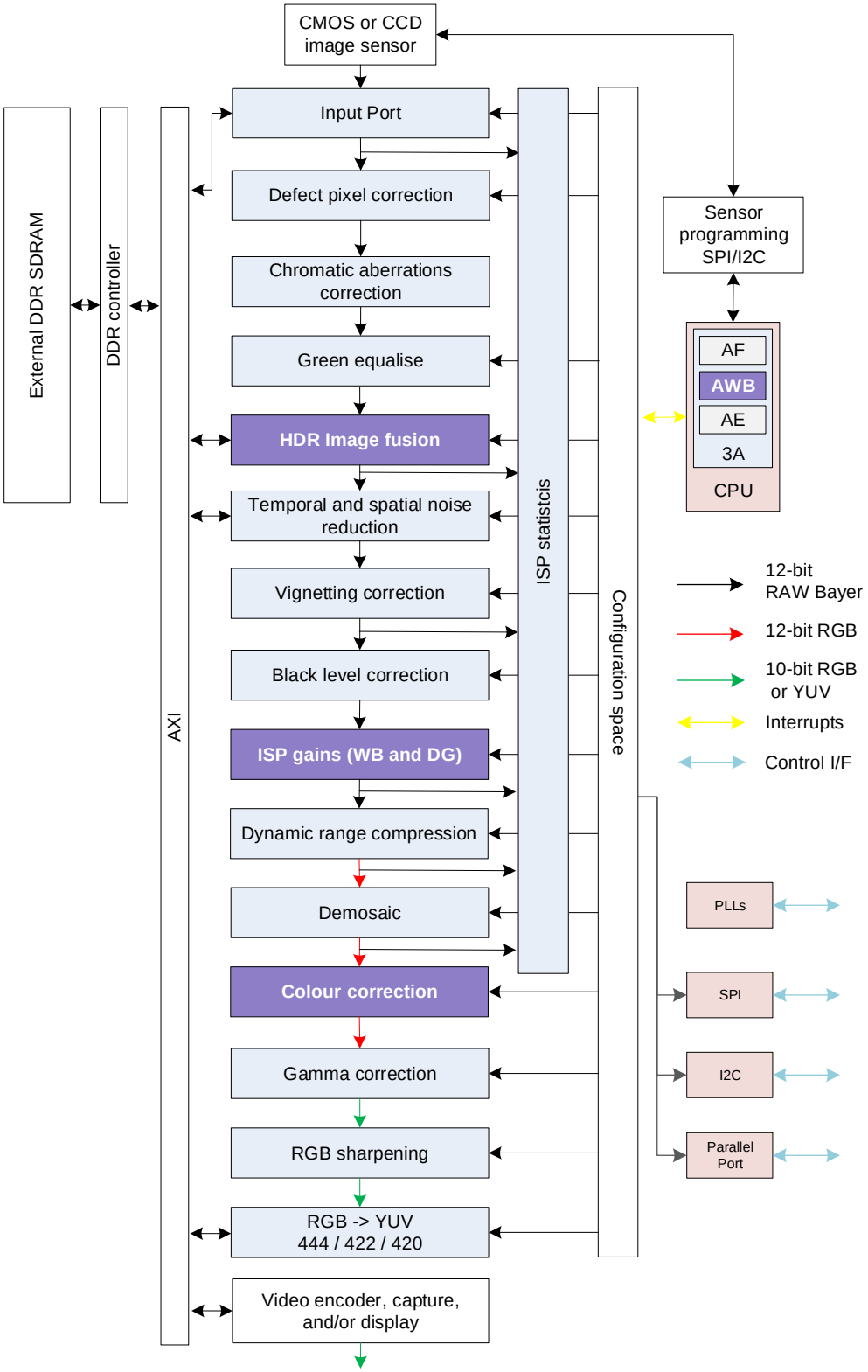


Figure 4.5: ISP top level architecture

The proposed ISP pipeline has been implemented on an Altera Stratix IV GX FPGA development kit, 530 Edition. This FPGA has a CPU NIOS II bare metal, ruining at 75MHz, and 256 kB memory. For video output, a Terasic DVI HSMC card is used, and in order to capture images, a SD capture module is used. The FPGA setup is shown in figure 4.6.

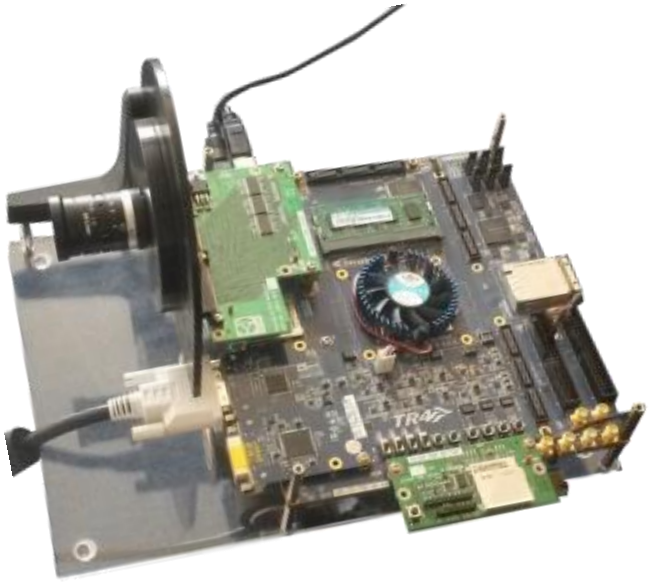


Figure 4.6: FPGA setup.

The implemented ISP runs at a processing speed of 150Mpix/sec, which is sufficient to process HD video at 60 FPS or 4K video at 15 FPS. Synthesis figures for the ISP pipeline, implemented in the Altera FPGA showed that the full ISP design has a gate count of 1122 K-gates. Out of the 1122 K-gates of the full ISP design, the proposed HDR and constancy solutions required 139 K-gates of the design. The illuminant estimation process of the colour management system proposed in Chapter 5 is run by the firmware on the NIOS II CPU. Measurements performed at 30FPS showed that the average CPU utilisation is 21%, with a peak utilisation of 69%. It requires 224 kB of memory for data and code, and the bandwidth measured is of 894 Kbytes/sec.

In Chapters 5, 6, and 7 the algorithms are benchmarked against state-of-the-art solutions. During the benchmarking process, the proposed ISP architecture is implemented in software. This ISP software simulator is used as a test bench for the evaluation of all algorithms considered. This is because not all of the algorithms in the literature can be

implemented in hardware, or when implementation details and source code is not available; some of these are only distributed in binary form. Moreover, the benchmarking tests are aimed to evaluate the performance of the algorithms by themselves and not their hardware implementation feasibility. This way, the modules to be tested can be isolated and accurately measure their performance with a common ISP pipeline framework.

4.3 Conclusions

This chapter discusses the importance of an ISP, and how the HDR and colour constancy algorithms presented in Chapters 5, 6, and 7 can be incorporated within an ISP to remove image quality artefacts introduced by the image sensor and optics of digital cameras.

In order to incorporate the HDR and colour constancy algorithms presented in the following chapters, an ISP architecture is defined and implemented in hardware in an Altera FPGA capable of processing HD video at 60 FPS or 4K video at 15 FPS. Moreover, the ISP architecture proposed in this chapter is implemented in software as an ISP simulator for the evaluation of the proposed algorithms against state-of the-art solutions on the field of HDR imaging and colour constancy.

Chapter 5

Multi-Exposure Image Fusion for Real Time HDR Content Creation

This chapter presents a novel, robust and low complexity multi-exposure imaging algorithm for HDR content creating in real time. A special emphasis of the proposed approach is its unique ability withstand motion within the set of multi-exposure frames due to camera shake or moving objects in the scene, a fundamental drawback of the state-of-the-art algorithms.

5.1 Introduction

Multi-exposure image fusion is a well-known approach adopted to create high dynamic range images and emulate the human visual system using standard dynamic range cameras. The main limitation of current multi-exposure image fusion techniques is their inability to compensate for moving objects in a scene, camera shake, and being able to run in real time for HDR video content creation. Previous attempts to solve camera shake have been able to accurately align the multi-exposure images that have static backgrounds prior to their fusion. Nonetheless, image alignment cannot solve the issue of ghosting artefacts due to moving objects.

HDR photography or HDR video can be achieved by capturing images at different exposures and then fusing them to produce HDR images. However, in order to produce artefact free images, the fusion technique has to be able to compensate for motion caused by camera shake and object movements in the scene. During the fusion process, ghosting artefacts are generated due to the fractionally time different instances of the objects' displacement, within the multi-exposure images captured. Previous attempts have being made by other

researchers to resolve ghosting artefacts as in [78], and Chapter 5, where the HDR content creation with camera shake was solved. However, its complexity and memory requirements would not allow its implementation for real time HDR processing. Moreover, the approach presented in Chapter 5 was able to improve ghosting artefacts but it was not able to completely remove them in fused areas. That is, objects appeared in short and long exposures with slightly different positions. Thus, an ideal practical solution to the problem of multi-exposure image fusion that addresses the problem of HDR creation in real time has yet to be found.

In previous research [82, 83] a practical method to perform spatio temporal noise reduction in RAW images within an ISP was proposed. The spatio temporal method relied on the idea of matching, blending, and recursive accumulation of image data into a frame buffer to improve signal to noise ratio. Errors due to motion were handled by the noise reduction engine. The spatio temporal noise reduction method [82], was adopted and extended so that it can be utilised for the purpose of multi-exposure image fusion in real time with local motion compensation [83]. It is noted that this HDR solution [83], is a practical application in which accumulation of data is carried out by the imaging sensor, and not by a frame buffer. Thus, the problem of accumulating images to improve signal to noise ratio, is transformed into the problem of matching images taken at different exposures in order to produce artefact free HDR images.

The spatio temporal HDR solution [83], performs motion estimation and compensation in two stages. First, a robust optical flow in a Laplacian Pyramid style method is used for coarse motion estimation by adopting a robust optical flow method [83]; followed by a block-matching process, which allow the matching of pixels to be at a sub-pixel level. Once the multi-exposure images were motion compensated, a temporal filter [81] is used to perform the image blending to produce a wide dynamic range image. Overall, this method produced impressive results and can run in real time [83], however, its complexity and memory requirements make this approach very costly in terms of hardware implementation.

In this chapter, in order to overcome the shortcomings presented from the spatio temporal HDR solution [83], a robust and low complexity multi-exposure imaging method for video

HDR creation is proposed. In this method proposed, motion within the set of multi-exposure frames due to camera shake or moving objects in the scene is compensated by a global motion estimation process, followed by a fusing process that can compensate for local motion of objects in the scene. The proposed method exploits the benefits of the Laplacian Pyramid for motion estimation processes, and the linearity of Bayer RAW domain data for better noise predictability, accurate decisions when fusing, and decreased memory requirements.

For clarity of presentation, Chapter 5 is organised as follows. Section 5.2 presents the functional details of the proposed system. Section 5.3 provides experimental results and a detailed evaluation of the system's performance. Concluded the proposed research, section 5.4 gives an insight to further work.

5.2 Proposed Image Fusion Method

In the proposed HDR method, all processes are performed in the Bayer RAW domain as it allows more accurate calculations when fusing the images due to the linear nature of the data, and predictability of the noise characteristics.

The first step in the proposed approach is to perform global motion estimation and correction to compensate for motion due to camera shake. The second step of the proposed method is to blend the set of multi-exposure images, wherein its logic, it compensates for objects moving within the scene. The combination of these two steps, global motion compensation in conjunction with the motion aware fusing technique, allows removing motion artefacts and ghosting due to camera shake and moving objects in the scene.

Once the multi-exposure images are motion corrected and fused, a dynamic range compression (DRC) method is used to create the resulting HDR image. Figure 5.1 shows the block diagram of the proposed multi-exposure image fusion method.

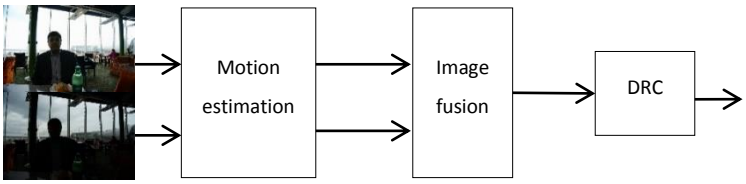


Figure 5.1: Block diagram of the proposed multi-exposure image fusion block algorithm.

5.2.1 Global motion estimation

Global motion estimation is performed to remove possible artefacts due camera shake. This is achieved by adopting the idea of the multi-scale motion estimation technique proposed in [81 and 82], which is based on linear matching of images in a transform space. Global motion estimation is performed by first constructing a Laplacian Pyramid of two levels to reduce the number of computations during the motion estimation calculation, followed by an intensity matching process where I_s (short exposure image) and I_l (long exposure image) are equalised in the intensity domain.

The intensity matching process is performed so that short and long exposures have the same intensity levels to allow the motion estimation algorithm to be able to compare the images in a more accurate way. The intensity matching process is performed as follows (Equations 5.1 - 5.9):

$$p_{Is(i)} = p(I_s = i) = \frac{n}{n_i}, 0 \leq i < L \tag{5.1}$$

$$p_{I_l(i)} = p(I_l = i) = \frac{n}{n_i}, 0 \leq i < L \tag{5.2}$$

$$I_{cfdIs}(i) = \sum_{j=0}^{i} p_{Is}(j) \tag{5.3}$$

$$I_{cfdIl}(i) = \sum_{j=0}^{i} p_{\hat{I}_l}(j) \tag{5.4}$$

Where *p* is the image the probability of an occurrence of a pixel of level i in the image, P_{ls} and P_{ll} are the histograms of I_s and I_l, and I_{cdfls}, I_{cdfll}, are the cumulative histograms of I_s and I_l. *L* is the number of intensity levels in the image, in this case 0 to 255 possible values, *n* the total number of pixels in the image and n_i the occurrences of grey level *i.* The cumulative histograms are used for determining thresholds for creating the equalised intensity masks, $Imask_{dk}$ and $Imask_{br}$, as follows:

$$u = \sum_{j=0}^{i} I_{cfd_{\hat{l}l}}(i) < L * 0.95 \tag{5.5}$$

$$Icut_{dk} = \arg\min_{index}\left(\left|I_{cfd_{Is}} - u\right|\right) \tag{5.6}$$

$$Imask_{dk}(x,y) = \begin{cases} 1 \; if \; I_s(x,y) > Icut_{dk} \\ 0 \; otherwise \end{cases} \tag{5.7}$$

$$Icut_{br} = \arg\min_{index}\left(\left|I_{cfd_{Il}} - (1 - \langle Imask_{dk}\rangle) * n\right|\right) \tag{5.8}$$

$$Imask_{br}(x,y) = \begin{cases} 1 \; if \; I_l(x,y) > Icut_{br} \\ 0 \; otherwise \end{cases} \tag{5.9}$$

Where *u* is the number of unclipped pixels in I_l, and $Icut_{dk}$ and $Icut_{br}$ are the thresholds calculated for the intensity masks.

Once the intensity matching is performed, I_s and I_l are represented in a transform space of vectors. $\hat{V}h_s$, $\hat{V}v_s$, $\hat{V}h_l$, $\hat{V}v_l$ are composed of the mean of each row and each column of $Imask_{dk}$ and $Imask_{br}$ as in equations 5.10 and 5.11.

$$\hat{V}h_s(x) = \sum_{i=0}^{x} Imask_{dk}(x,i), \qquad \hat{V}v_s(y) = \sum_{i=0}^{y} Imask_{dk}(i,y) \tag{5.10}$$

$$\hat{V}h_l(x) = \sum_{i=0}^{x} Imask_{br}(x,i), \qquad \hat{V}v_l(y) = \sum_{i=0}^{y} Imask_{br}(i,y) \tag{5.11}$$

Where $x = \{0, 1, \dots, n_{col}\}$ and $y = \{0, 1, \dots, n_{rows}\}$.

In order to estimate the delta offsets, Δ_h and Δ_v, due to possible motion from the multi-exposure images, the vectors, $\hat{V}h_s$, $\hat{V}v_s$, $\hat{V}h_l$, $\hat{V}v_l$, are analysed as follows:

$$\widehat{me}_h(i) = \sum_{s=-d}^{d} \langle \left| \hat{V}h_s - (e^{(-(d,\dots,d)+s)^2} \otimes \hat{V}h_l) \right| \rangle \times e^{\left(-\frac{\left(-\left(-\frac{nh}{2}+.5,\dots,\frac{nh}{2}+.5 \right) \right)^2}{(nh/4)^2} \right)} \quad (5.12)$$

$$\Delta_h = \frac{\arg\min_{index}(\widehat{me}_h)}{d} \times 2 \quad (5.13)$$

$$\widehat{me}_v(i) = \sum_{s=-d}^{d} \langle \left| \hat{V}v_s - (e^{(-(d,\dots,d)+s)^2} \otimes \hat{V}v_l) \right| \rangle \times e^{\left(-\frac{\left(-\left(-\frac{nv}{2}+.5,\dots,\frac{nv}{2}+.5 \right) \right)^2}{(nv/4)^2} \right)} \quad (5.14)$$

$$\Delta_v = \frac{\arg\min_{index}(\widehat{me}_v)}{d} \times 2 \quad (5.15)$$

$\widehat{me}_v$ and $\widehat{me}_h$ are the motion estimation vectors for the vertical and horizontal directions, *d = delta,* the amount of displacement to be measured in pixels, *i = s +delta, nh* is the number of columns in the image, and *nv* is the number of rows in the image. It is noted that in Equations 5.13 and 5.15 the calculated delta offsets are multiplied by 2 in order to compensate for the downscaled images during the construction of the two level Laplacian pyramid.

An example of the proposed motion estimation process is presented in Figure 5.2. Figures 5.2(a) and 5.2(b) illustrate the two Bayer RAW multi-exposure images prior to motion estimation. Figures 5.2(c) and 5.2(d) show the *Imasks* calculated using equations 5.9 and 5.11. Figures 5.2(e) and 5.2(f) depict the ability of $\widehat{me}_v$ and $\widehat{me}_h$to accurately estimate the image's vertical and horizontal displacement vectors.

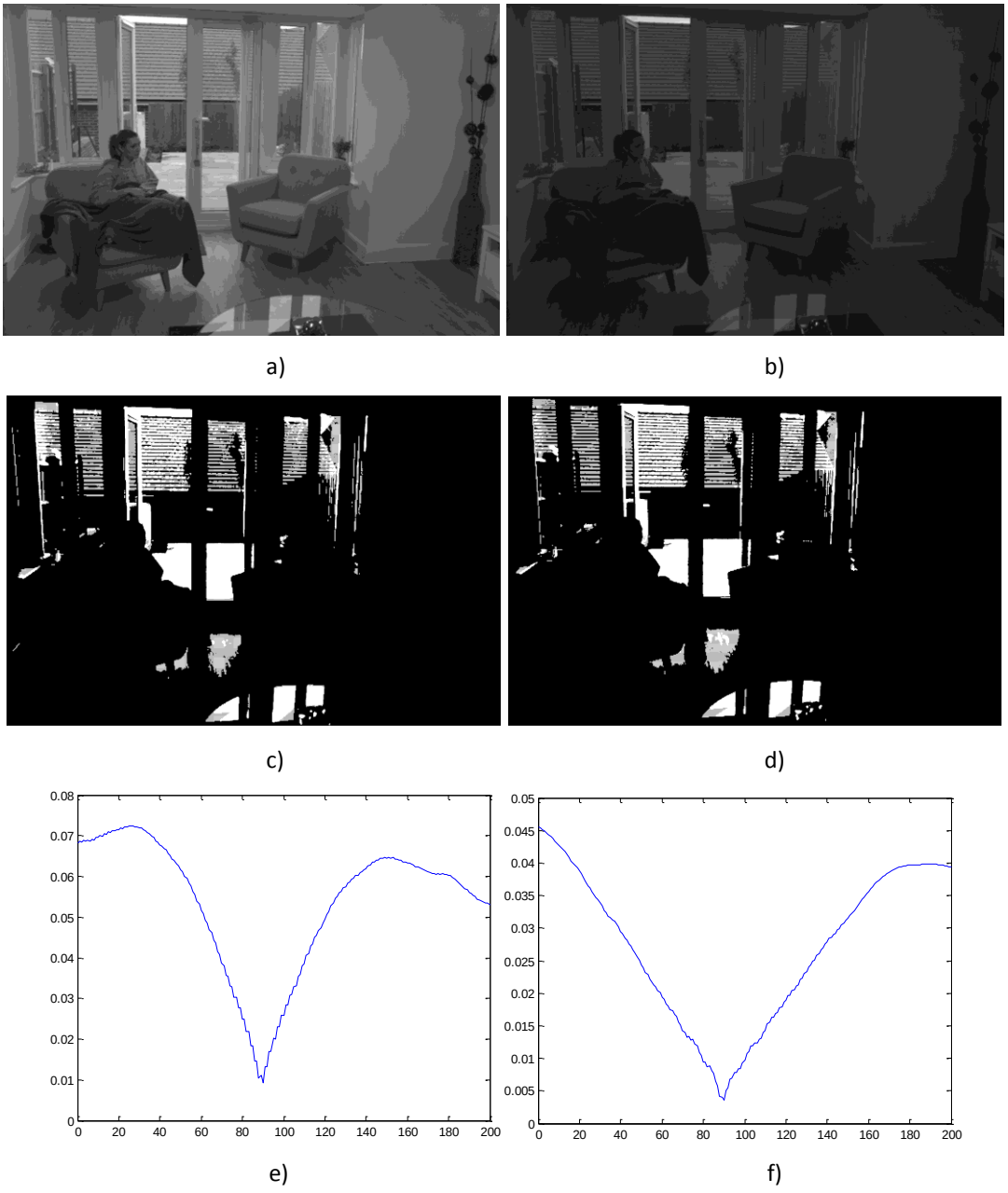


Figure 5.2: Motion estimation process. (a) and (b) over and under exposed images to be registered. (c) and (d) Illustration of intensity equalisation for the motion estimation process. (e) and (f) plots of horizontal and vertical motion vectors, with delta parameter =100.

The resulting delta offsets for the horizontal and vertical directions that are calculated during the motion estimation stage are propagated and used by the image fusion method to produce the HDR image.

5.2.2 Image Fusion

In this research, a motion aware image fusion method is proposed in order to produce high dynamic range images. The fusion process uses the global motion compensation delta offsets previously calculated in section 5.2.1, and incorporates a motion aware logic for removing motion artefacts and ghosting due camera shake and moving objects in the scene.

In order to blend the multi-exposure images, it is necessary to calculate the exposure ratio between I_s and I_l prior the fusing process; knowing the exposure ratio allows constructing fusing rules to decide which pixels with certain level of intensity will be chosen from I_s or I_l to create the final HDR image, I_{hdr}.

The exposure ratio (*ER*) set from short to long images can be calculated as follows (Equation 5.16).

$$ER = \frac{ETshort \times AGshort \times DGshort \times Aperture}{ETlong \times AGlong \times DGlong \times Aperture} \tag{5.16}$$

Where *ET* is exposure time, *AG* is analogue gain, and *DG* is the digital gain set by the sensor.

After the calculation of exposure ratio is done, luminance of I_s and I_l, represented by $Lmax_s(x,y)$ and $Lmax_l(x,y)$, are calculated using equations, 5.17 and 5.18 so that the fusing rules, $\alpha(x,y)$, can be applied on I_s and I_l to generate I_{hdr}.

$$Lmax_s(x,y) = argmax\left(\sum_{i=0}^{2}\sum_{j=0}^{2} I_s(x-1+i, y-1+j) \times ER\right) \tag{5.17}$$

$$Lmax_l(x,y) = argmax\left(\sum_{i=0}^{2}\sum_{j=0}^{2} I_l(x-1+i, y-1+j)\right) \tag{5.18}$$

The fusion rules in this approach consist of intensity thresholds, th_s and th_l, that determine the contribution of I_s and I_l to the final image as shown in Equation 5.19.

$$\alpha(x,y) = \begin{cases} 1 \; if \; argmax(Imax_s(x,y), Imax_l(x+\Delta_h, y+\Delta_v)) > th_s \\ 0 \; if \; argmax(Imax_s(x,y), Imax_l(x+\Delta_h, y+\Delta_v)) < th_l \\ \frac{th_l - th_s}{-1} \times \left(1 - argmax(Imax_s(x,y), Imax_l(x+\Delta_h, y+\Delta_v)\right) + th_s, Otherwise \end{cases} \quad (5.19)$$

In order to compensate for local motion or non-matching areas from the set of multi-exposure images, the absolute difference, *ΔI,* is calculated using Equations 5.20, 5.21 and 5.22 so that non-matching areas in the image can be identified and therefore compensated.

$$Iave_l(x,y) = \frac{1}{9} \times \left(\sum_{i=0}^{2} \sum_{j=0}^{2} Imax_l(x-1+i, y-1+j) \right) \quad (5.20)$$

$$Iave_s(x,y) = \frac{1}{9} \times \left(\sum_{i=0}^{2} \sum_{j=0}^{2} Imax_s(x-1+i, y-1+j) \right) \quad (5.21)$$

$$\Delta_I(x,y) = \begin{cases} \frac{|Iave_l(x+\Delta_h, y+\Delta_v) - Iave_s(x,y)| \times \delta}{ER} \\ 0 \; if \; |Iave_l(x+\Delta_h, y+\Delta_v) - Iave_s(x,y)| \leq I_{th} \end{cases} \quad (5.22)$$

Where I_{th} is an intensity threshold that can be used for masking very dark areas of the image. I_{th} is usually used for pixels below the black level of the sensor that may contain high level of noise, and δ is a signal amplifier that can be used for fine-tuning the detection magnitude response.

If high levels of noise are present in I_s when calculating *ΔI*, false motion will be detected due to the noise introduced by the imaging sensor. To address this shortcoming, a modelling of the sensor noise profile is used as a thresholding method to determine if the absolute difference of I_s and I_l is due to motion in the scene, or noise in I_s.

Sensor noise modelling has been investigated, [86, 87, 88, and 89], where the noise profile of an image *I,* sampled by the imaging sensor's A/D converter, is assumed to be composed of different sources of noise.

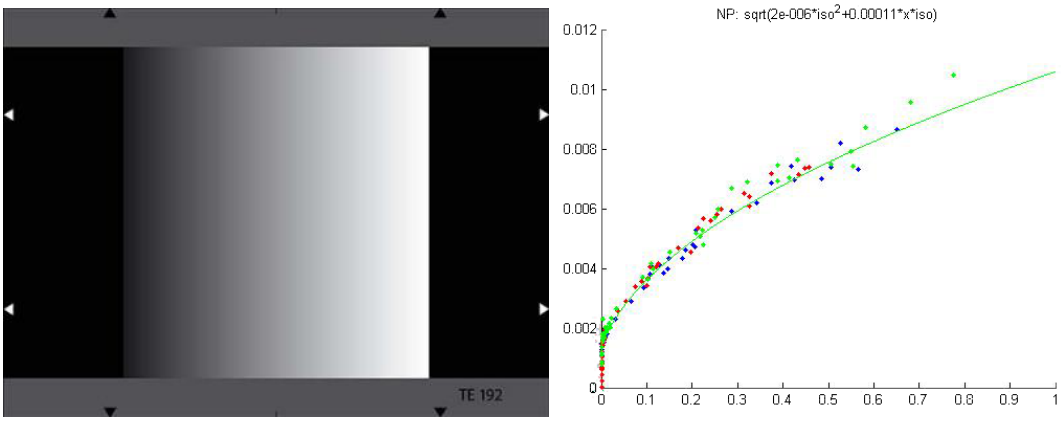
Assume that image data representing the actual scene without noise added to the image is defined as $I_{p,}$ the analogue gain is defined as $n_{ag}(q)$, which is characterised by a Gaussian distribution with standard deviation σ_a, fixed pattern noise is $n_{fpn}(x,y)$ [5], and photon shot noise $n_q(I)$ *is* described as a random process with a Poissonian distribution with a standard deviation of σ_q.

Given the above, the noise of an image *I,* can be represented as in Equation 5.23. Thus the standard deviation for a sensor noise can be defined as in Equation 5.24, where I_{max} is the maximum digital value output of the imaging sensor

$$I(x,y) = I_p(x,y) + n_{ag}(q) + n_{fpt}(x,y) + n_q\left(I_p(x,y)\right) \qquad (5.23)[82]$$

$$\sigma^2 = \sigma_a^2 + \sigma_q^2 \times \frac{I_s(x,y,q)}{I_{max}} \qquad (5.24)[82]$$

In order to calculate σ^2 of a sensor, a series of images thought all exposures times of the image sensor is captured so that the total dynamic range of the sensor is covered. Then, these images are analysed by plotting the standard deviation versus intensity level of *I,* and a best-fit curve is used to determine the noise profile (*np*) of a sensor. Figure 5.3(a) shows the image used for calculating the noise profile, and Figure 5.3(b) illustrates how Equation 5.24, the noise profile, can be determined experimentally. The red, green and blue dots represent noise variances for the corresponding RGGB pixels, measured throughout the whole range of intensities of a characterised imaging sensor [82].



a) b)

Figure 5.3: a) Image captured for estimating the noise profile at ISO100. b) Plot of standard deviation vs. intensity calculated from the image captured.

Once the noise characteristics of the imaging sensor are analysed, the imaging sensor's noise profile is calculated using Equation (5.25) to determine if Δ_I magnitude is due to noise or motion.

$$\Delta_{Inp}(x,y) = \begin{cases} 0 \; if \; \Delta_I(x,y) < \dfrac{1}{np\left(\Delta_I(x,y)\right)} \\ \Delta_I(x,y) \; otherwise \end{cases} \tag{5.25}$$

Once Δ_{Inp} is calculated, $\alpha(x,y)$ can be rewritten as in Equation (5.26) so that the non-matching areas from the multi-exposure images are taken into account during the fusing process as shown in Equation (5.27). It is noted that with this logic, motion detection will be handled by using short exposure data only, and thus, motion artefacts can be removed. The compromise of this method is that areas detected as motion will be noisier in I_{wdr} but no ghosting artefacts or discontinuities between the multi-exposure images will be visible.

$$\propto (x,y) = \propto (x,y) + \Delta_{Inp}(x,y) \tag{5.26}$$

$$I_{wdr}(x,y) = I_s(x,y) * \propto (x,y) + \left(1 - \propto (x,y)\right) * I_l(x + \Delta_h, y + \Delta_v) \tag{5.27}$$

The motion aware image fusion process proposed in this chapter is illustrated in Figure 5.4. Figures 5.4(a) and 5.4(b) show I_l and I_s respectively, where local motion of objects is created

by a shaking hand in the scene. Figure 5.4(c) shows α calculated using Equation (5.19); in this figure, a colour bar at the right of the image is provided to indicate the blending weights between I_s and I_l. 1 means, a pixel is purely based on I_s, 0 means the pixel is purely based on I_l and all others values are a blend between I_s and I_l. At this point the motion detection logic is disabled. Figure 5.4(d) illustrates the absolute difference ΔI, which is used for detecting motion within the multi-exposure images. It is noted that ΔI clearly illustrates how due to noise in I_s, blending regions produce false positive detections, when trying to identify motion. Thus, it is required to use noise characteristic knowledge when trying to detect motion. Figure 5.4(e) depicts ΔI thresholded by the noise profile as in Equation 5.25. In Figure 5.4(e), the colour bar is also provided to illustrate the noise regions where values of 1 indicates noise in the image, and 0 a clear signal.

Figure 5.4(f) shows the alpha blending mask and absolute difference after false motion detections due to noise are removed as in Equation 5.26. Figure 5.4(g) illustrates a crop of an image after the whole fusing process is performed with motion detection is disabled; red arrows have been drawn in (g) to illustrate the ghosting artefacts. Figure 5.4(h) illustrates the resulting image crop after the whole fusing process is performed with motion detection enabled.

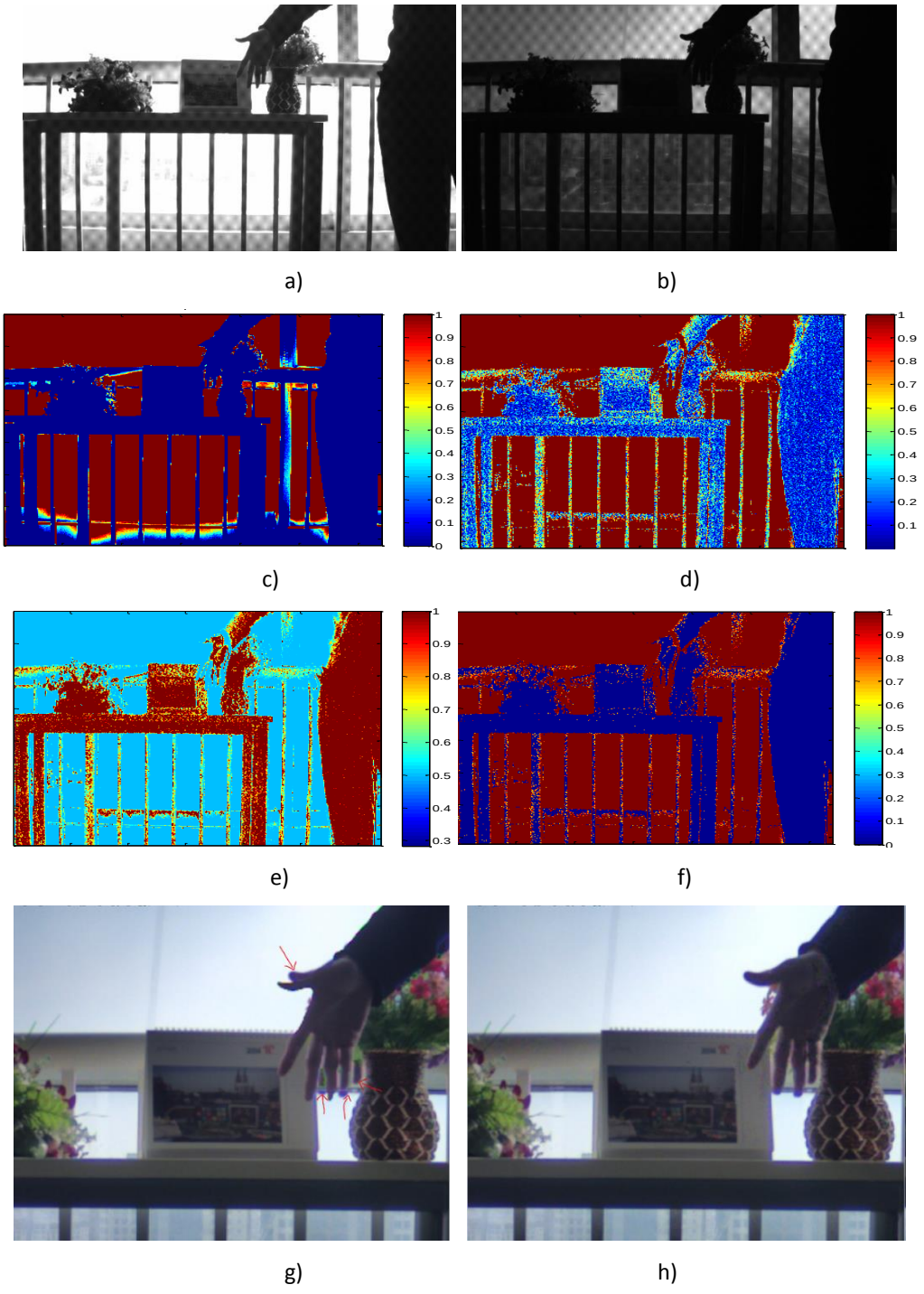


Figure 5.4: Image fusion process. (a) and (b) Original set of multi-exposure images. (c) Result of the fusion rules alpha mask without motion detection logic enabled (d) Absolute difference prior the application of the noise profile. (e) Noise profile calculated on the absolute difference. (f) Result of the fusion rules with motion detection logic enabled. (g) Output image crop of the image fusion method with motion detection disabled. (h) Output image crop of the image fusion method with motion detection enabled.

5.2.3 Dynamic range compression

In order to be able to visualise the contents of the wide dynamic range image produced during the image fusion stage, a local histogram equalisation or local tone mapping technique is applied to I_{hdr}.

The dynamic range compression algorithm used in this research is an algorithm from a commercial product called iridix©. Iridix core performs dynamic range compression or tone mapping of an input image based on a model of the human visual system. It is used in a camera pipeline to produce most natural images under a wide range of capture conditions, typically by revealing shadow detail which would otherwise be under-exposed in high contrast situations. The purpose of the iridix transform is to map the image content from an input source such that it remains fully visible on an output display, without loss of content. Iridix is specifically designed to preserve colour, sharpness and boost the contrast of the source image, and it is based on a space-variant or pixel-by-pixel processing type algorithm. The algorithms inside iridix are adaptive, meaning that a single set of parameters can be used to process any source image or any video sequence under different lighting conditions.

The dynamic range compression algorithm is an important part of the evaluation of the system, though not the subject for the proposed research. The objectives for the inclusion of the dynamic range compression algorithm in this research are: the ability to compress the dynamic range of I_{hdr} so that shadow areas in I_{hdr} are boosted and therefore matched to the look of the corresponding areas in I_l; the ability to represent highlighted areas of I_{hdr} matching the look of corresponding areas in I_s; and the minimisation of any low spatial frequency artefacts.

It is noted that in this research, the objectives of assessing wide dynamic range compression techniques are quite difficult to formalise and describe by using quantitative metrics. Thus such issues are not investigated assuming that the dynamic range compression algorithm serves its purpose. Hence, this work concentrates on the noise measures of the resulting image I_{hdr}, and the ability to compensate for global and local motion within the set of multi-

exposure images in comparison to the work presented by Romanenko [83]. The results of the experiments are presented in the following section.

5.3 Experiments, Results & Analysis

In this section, experimental results are presented in order to demonstrate the performance of the propose method for creating HDR content and prove its ability to remove ghosting artefacts that are due to moving objects in the scene and camera shake.

The proposed method was tested on an extended set of Bayer RAW images captured with cameras used for different applications such as: video conferencing, CCTV surveillance, digital photography (DSLR) and in mobile phones. Each of the videos captured consisted of multi-exposure frames obtained by using different exposure ratios. This allowed generating short and long exposures fames so that they could be processed by the proposed method to generate the HDR image outputs.

In order to verify the performance of the proposed approach, the peak signal to noise ratio (PSNR) values were measured as a reference to evaluate the quality of the fusion process. Moreover, each of the test sets was visually examined for spatial artefacts such as ghosting and fusion discontinuities.

In addition to the PSNR values measured, the proposed approach's performance was also compared against the HDR method proposed in [83]. This comparison consisted of PSNR values measured in the outputs of the HDR methods, and subjective analysis for identifying ghosting artefacts and fusing discontinuities.

5.3.1 Experiment setup

In order to evaluate the performance of this research, the proposed method was placed inside an ISP simulator so that the results could be visualised in the gamma corrected RGB domain. The ISP modules used in conjunction with the proposed method included black level

correction, grey world white balance, dynamic range compression (iridix), bilinear demosaic, colour correction, and standard sRGB gamma correction. The simplified ISP used in the experiments is shown in Figure 5.5

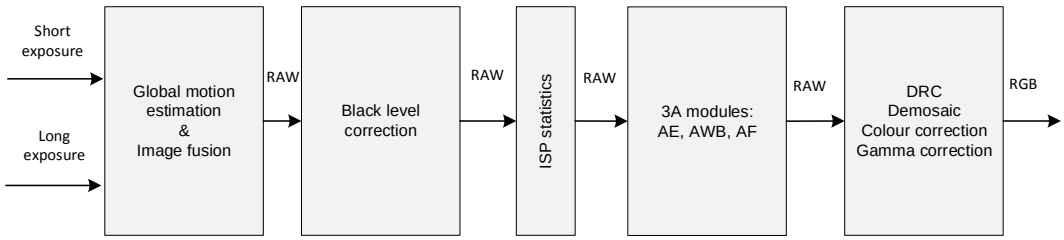


Figure 5.5: ISP used for the experiments.

Table 5.1 presents the characteristics of each of the imaging sensors used, and Table 5.2 presents the parameter set in the proposed approach during the experiments conducted. The noise profile of each sensor was calculated as described in Section 5.2.3, and all other parameters of the proposed approach were experimentally found. It is noted that once these parameters were tuned for each of the imaging sensors, none of the parameters were modified, when testing in different scenes and lighting conditions.

	Sensor model		
	OV8835 – used in mobile applications	**DSLR Sony NEX-5**	**AS3372- used in video conferencing and surveillance**
Resolution	8MP	13MP	2MP
FPS	7.5FPS	7.5FPS	60FPS
Data bits	10	12	12
Pixel size	2.7μm	3.5μm	1.4μm
Max signal to noise ratio	36.6dB	60+dB	59dB

Table 5.1: Imaging sensor characteristics used in the experiments.

	Sensor model		
	OV8835	**DSLR Sony NEX-5**	**AS3372**
Motion estimation: Delta	100 pixels	100 pixels	100 pixels
NP curve fitting formula a, b, and c coefficients: $\boldsymbol{np(x) = \sqrt{b \times ISO^2 + c \times x \times ISO}}$	a = 0.34 b = 6.8× (10^-5) c = 1.2× (10^-2)	a = 0.29 b = 1.15× (10^-5) c = 1.00× (10^-2)	a = 1 b = 1.01× (10^-8) c = 2.18 × (10^-5)
Stitching points:	th_l = 0.85 th_s= 0.93	th_l = 0.85 th_s= 0.93	th_l = 0.85 th_s= 0.93
Absolute difference (Δ_i) intensity threshold:	I_{th} = 0.050	I_{th} = 0.050	I_{th} = 0.080
Signal amplifier:	δ = 64	δ = 64	δ = 64
Exposure ratio	8 and 16	8 and 16	8 and 16

Table 5.2: Parameters of the proposed approach used in experiments.

5.3.2 Results and analysis

Experiments were carried out to evaluate the ability of the proposed method to produce artefact free HDR images. Figure 5.6 shows an example of one of the sets of multi-exposure images used in the experiments captured with the OV8835 [90] imaging sensor. In Figure 5.6, all processes performed by the proposed method in order to generate HDR images are shown. The over and under exposed RAW Bayer images obtained with an exposure ratio of 16 are shown in Figure 5.6(a) and (b). The intensity equalisation mask used during the global motion estimation process used for the over and under exposure images are shown in Figures 5.6 (c) and (d). Plots of the global motion estimation performed are shown in Figures 5.6 (e) and (f). Figure 5.6 (g) shows the fusing mask with motion aware logic enabled, and Figure 5.6 (h) shows the final HDR image.

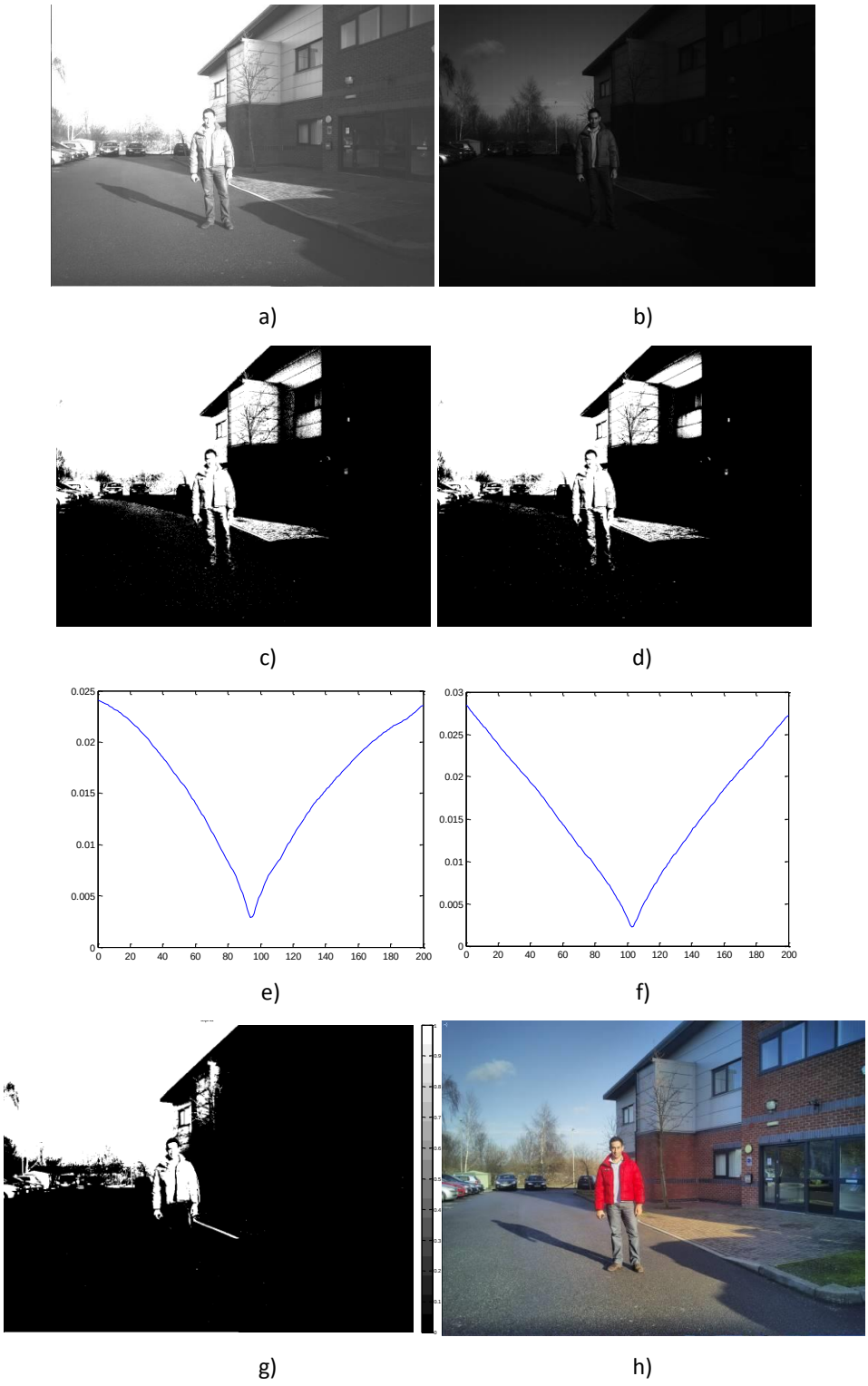


Figure 5.6: Step by step illustration of the HDR proposed method. a) and b) the over and under exposure RAW Bayer images. c) and d) The intensity equalisation mask. e) and f) the plots of the global motion estimation in the horizontal and vertical directions. g) Fusing mask with motion aware logic enabled. h) The final HDR image.

In the set of multi-exposure images shown in Figure 5.6, PSNR values were measured and it was possible to see the ability of the proposed method to extend the dynamic range of the output image from ≈35dB to ≈46dB, without showing visible artefacts due to camera shake or moving objects in the scene. This is better illustrated in Figure 5.7, where the ability of the proposed method to compensate for global and local motion estimation is shown.

Figure 5.7 (a) shows the crop of the output image with global motion estimation and motion aware logic disabled, and Figure 5.7 (b) shows the output image with global motion estimation and motion aware logic enabled. In this comparison, it is possible to see the importance of having motion estimation compensation in HDR imaging. Discontinuities and artefacts are observed around the car and face areas in Figure 5.7 (a), whereas Figure 5.7 (b) does not show any of these artefacts.

a)

b)

Figure 5.7: a) Crop of the output image with global motion estimation and motion aware logic disabled. b) Crop of the output image with global motion estimation and motion aware logic enabled.

In order to show the importance of HDR imaging, Figure 5.8 shows two examples of the scenario where dynamic range compression (DRC) is applied to the short exposure images. Note that if the noise floor level is not considered, the short exposure image with DRC applied represents the same dynamic range as the HDR image produced by the proposed method.

As can be observed in Figure 5.8, if DRC is applied to the short exposure image, large amounts of noise in the shadow areas are revealed. Figures 5.8 (b) and (g) show the under exposed images, Figure 5.8 (c) and (h) are the short exposure images with DRC applied, and Figure 5.8 (a) and (f) shows the HDR image produced by the proposed method. Zoomed in crops of the short exposure images with DRC applied (Figure 5.8 (e), (j)) and HDR images (Figure 5.8 (d), (i)) are provided for better illustration of the ability of the proposed method to improve signal to noise ratio.

Figure 5.8: PSNR values calculated from short exposure and fused images with DRC applied.

Another example of the proposed method is presented in Figure 5.9. In this case the images were captured with a DSLR camera, Sony NEX-5 [91], with the exposure ratio set to 8. It is observed how fast moving objects are best represented partially in a short exposure image, and partially in the long exposure image. This scenario is a common situation where fusion techniques fail to deal with object displacements, and therefore ghosting artefacts are produced. Figures 5.9 (a) and (b) show the over and under exposed images processed by the ISP for illustration purposes, Figure 5.9 (c) shows the fused image without motion compensation, Figure 5.9 (d) shows the fused image with motion compensation enabled, and Figure 5.9 (e) shows the zoomed in crops of (c) and (d).

a) b) c) d)

e)

Figure 5.9: Ability of the proposed method to compensate for motion. a) and b) shows the over and under exposed images. c) The fused image without motion compensation. d) The fused image with motion compensation enabled. e) Zoomed in crops of (c) and (d).

The proposed method was also tested in comparison to Romanenko's HDR method [83], since it was proved that its fusing of multi-exposure images was very accurate, even though, memory and logic required were very high. Figure 5.10 – 5.13 show images produced by Romanenko's HDR method and images produced by the proposed approach. This comparison is important since one of the aims of this research is to be able to produce as good results as Romanenko's HDR method, but with lower logic and memory requirements so that it can be implemented in hardware for real time HDR processing. In Figure 5.10 – 5.13 (a) the short and long exposure images are shown as a reference to the processed HDR images. Figure 5.10 – 5.13 (b) present the HDR image outputs from Romanenko's HDR method, and Figure 5.10 – 5.13 (c) the HDR image outputs from the proposed approach.

As it can be observed in Figure 5.10 – 5.13, the proposed method can produce similar results to Romanenko's HDR method. This is confirmed by the PSNR values measured in all test sets processed with the two methods. PSNR values were measured in the short exposure image with DRC applied, long exposure image, HDR image produced by Romanenko's HDR method and the proposed approach. The average of PSNR values measured over the whole set of multi-exposure images for each sensor are shown in table 5.3.

	Sensor model		
	OV8835	**Sony NEX-5**	**AS3372**
Long exposure	38dB	46dB	40dB
Short exposure + DRC	23dB	33dB	28dB
HDR output from [83]	36dB	45dB	41dB
Propose approach HDR output	35dB	42dB	40dB

Table 5.3: Average PSNR values measured in all tests sets.

a) b) c)

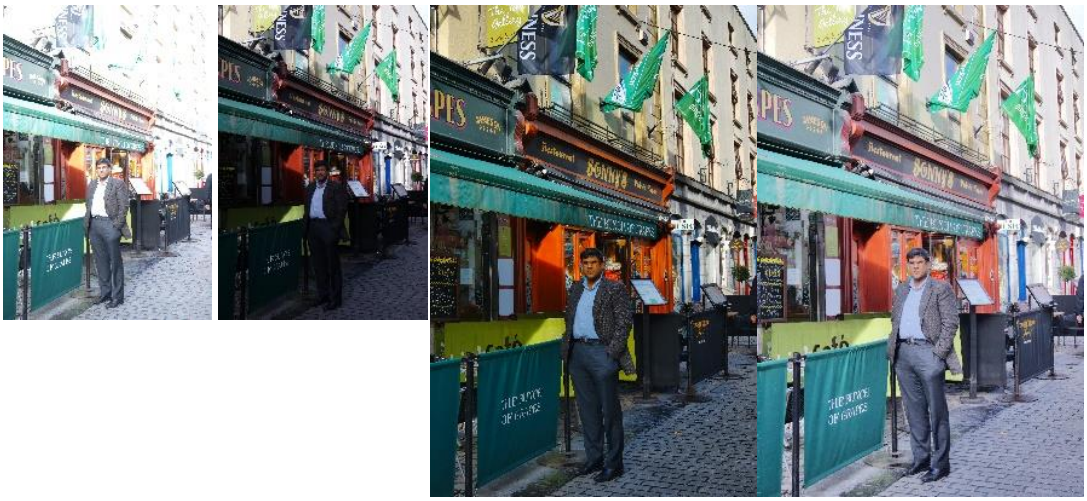
Figure 5.10 Performance comparison of the proposed approach versus the HDR image method proposed in [83]. a) Multi-exposure images. b) HDR image output from [83]. c) HDR image output from the proposed approach.

a) b) c)

Figure 5.11 Performance comparison of the proposed approach versus the HDR image method proposed in [83]. a) Multi-exposure images. b) HDR image output from [83]. c) HDR image output from the proposed approach.

a) b) c)

Figure 5.12: Performance comparison of the proposed approach versus the HDR image method proposed in [136]. a) Multi-exposure images. b) HDR image output from [136]. c) HDR image output from the proposed approach.



a) b) c)

Figure 5.13: Performance comparison of the proposed approach versus the HDR image method proposed in [83]. a) Multi-exposure images. b) HDR image output from [83]. c) HDR image output from the proposed approach.

According to the PSNR values measured, the proposed approach was able to match the performance of Romanenko's HDR method. This is very valuable since the proposed approach does not require large amounts of memory storage, bandwidth, and logic for its implementation in hardware.

The proposed HDR solution for video was implemented in hardware and incorporated to an ISP of a digital camera as described in Chapter 4. Figures 5.14 - 5.18 present frames of videos

captured with the proposed HDR method in an imaging lab at different lighting levels, and in real case scenarios in day light. In these figures, it is possible to observed that the proposed HDR algorithm can cope with moving objects within the scene and camera shake.

The setup of the imaging lab is presented in Figure 5.14. The lab scene contains a saw tooth signal test chart [129] used for measuring dynamic range, a moving train, and objects within a box at 2 lx. While capturing the videos, the camera was handheld and the train was set to a constant speed.

Figure 5.14:Lab setup for HDR video experiments

Figure 5.15: HDR scene in at 800 lx.

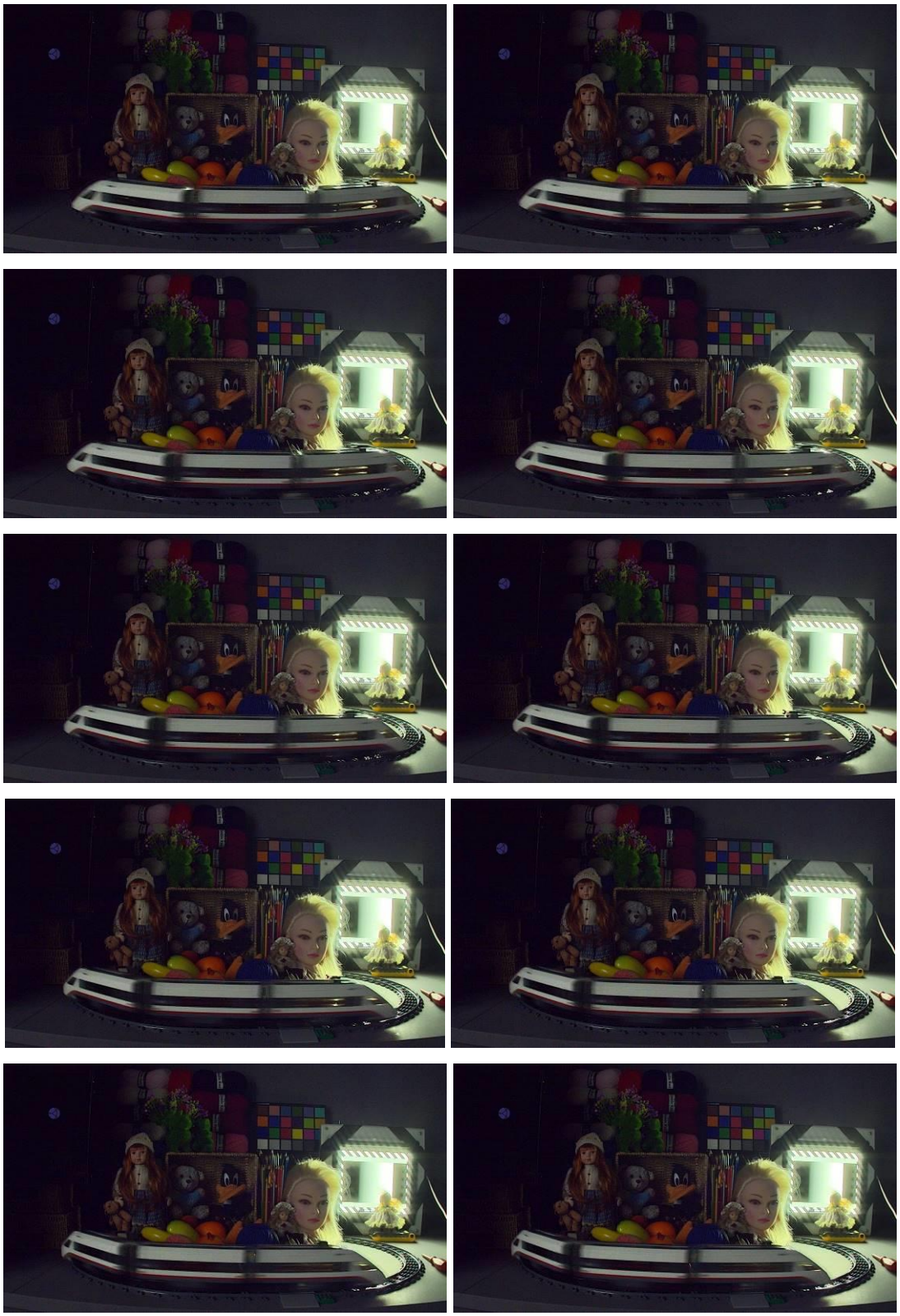


Figure 5.16: HDR scene at 5lx.

Figure 5.17: : HDR scene day light

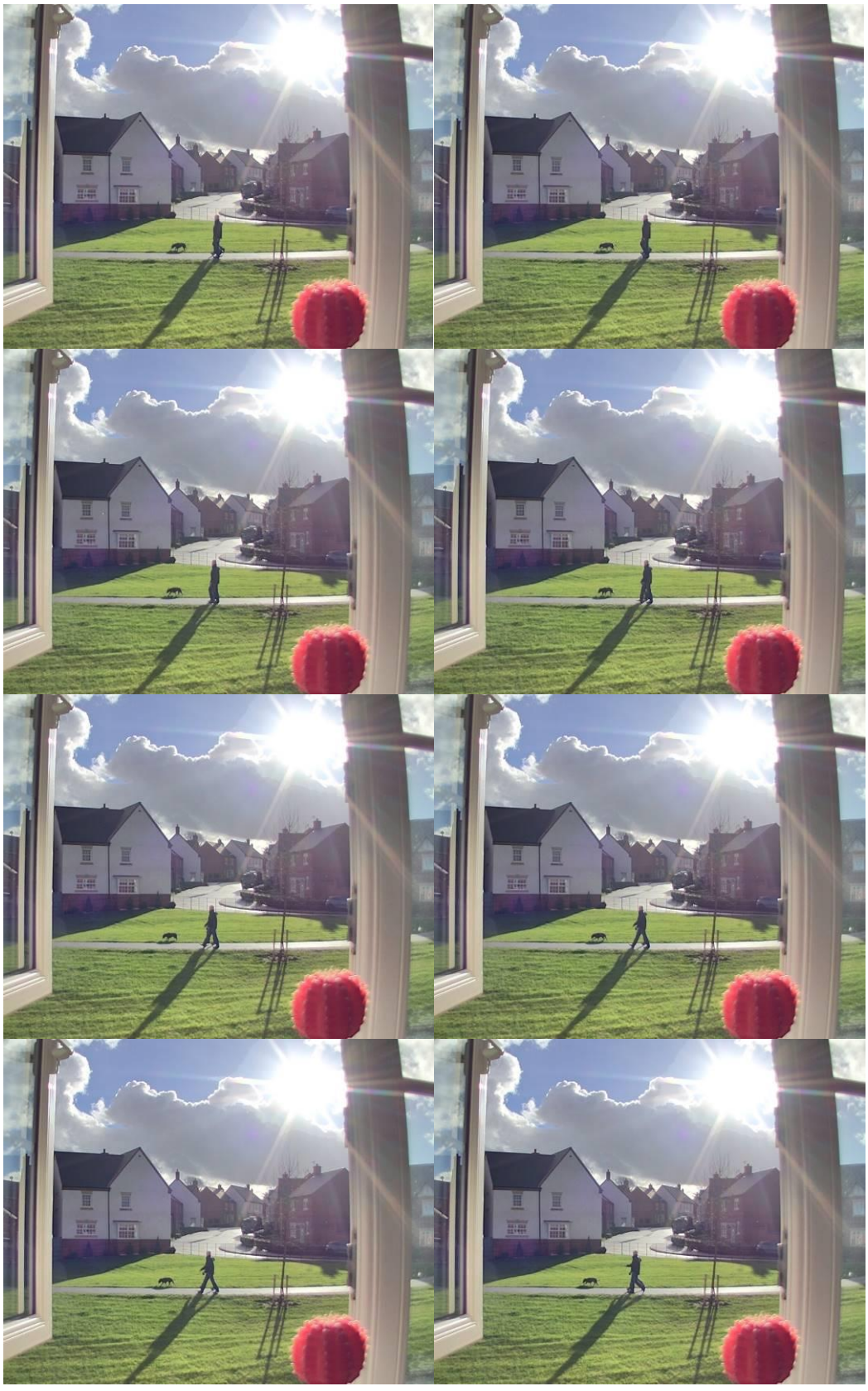


Figure 5.18: HDR scene day light

The limitations of the proposed algorithm are similar to the proposed spatio-temporal algorithm [83]. That is, in order to detect motion of objects in a pair of images, the objects should be present in both scenes and should be captured in both images. This limitation sets a limit on the range of optimal exposure ratios usable when capturing multi-exposure images. In the experiments performed, an exposure ratio of 1:16 was the maximum optimum value before the proposed method failed to accurately compensate for motion. This is because if images are taken at very different exposure levels, dark objects may not have details captured in the short exposure image, and thus making the motion estimation process impossible. Another limitation of the proposed method was its inability to be able to compensate for large, hundreds of pixels, rotational displacements due to camera shake.

5.4 Conclusions

A multi-exposure image fusion algorithm suitable for practical implementation in hardware was proposed in this chapter. The proposed algorithm was able to perform fusion of images taken at different exposures so that the total dynamic range of an image captured with a limited dynamic range imaging sensor is increased.

The proposed approach performs global motion estimation and correction to compensate for motion due to camera shake, followed by the fusing of the set of multi-exposure images, where in its logic, it compensates for objects moving within the scene. The combination of these two steps, global motion compensation in conjunction with the motion aware fusing technique, allows removing motion artefacts and ghosting due to camera shake and moving objects in the scene.

The proposed algorithm proved to be efficient in different lighting conditions and scenes and was proven to work well with different sensors.

After further analysing the results from the experiments performed, it was shown that the limitation of the proposed approach is its inability to be able to compensate for motion at exposure ratios bigger than 1:16. This problem can be solved if the fusion system was

extended to three-exposure image fusion instead of a two exposure image fusion exercise. In three-exposure image fusion system, exposure ratio between short to medium and medium to long exposures could be set to 1:16 and 1:16, and therefore extend the total exposure ratio to 1:256 to be able to capture 120dB of dynamic range. Unfortunately, imaging sensors capable of capturing RAW Bayer data at high frames per second are not widely available and thus, it was not possible to test this option. Another limitation found in this research, was the ability to compensate for large rotation displacements within the set of multi-exposure images. Thus future research should focus on proposing new HDR methods that include efficient motion estimation algorithms capable of compensating for rotational displacements in order to create HDR content in real time.

Chapter 6

A Novel Colour Management System for Image Signal Processors

This chapter presents a novel colour management system within an ISP pipeline that encompasses the colour constancy problem and the reproduction of pleasing colours to enable digital camera devices to mimic the HVS's colour constancy capabilities.

6.1 Introduction

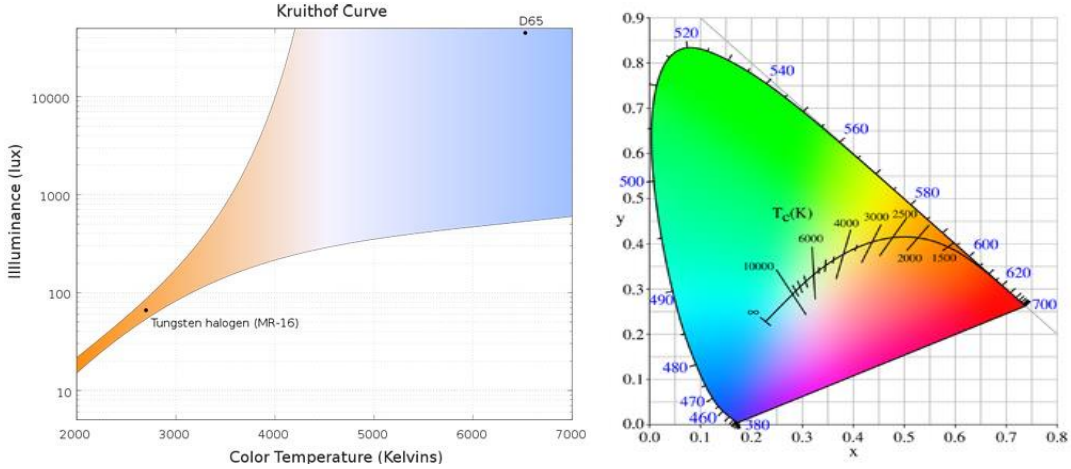
Colour constancy can be defined as "the constancy of the perceived colours of surfaces under changes in the intensity and spectral composition of the illumination" [92]. The human visual system is colour constant to a large extent, but this is affected by the contents of the field of view, state of adaptation, intensity level of illumination and difference of the adapting chromaticity from different kinds of illumination conditions [92]. Colour constancy in digital photography, also known as automatic white balancing (AWB), refers to the ability of a digital camera to recognise and reproduce the colour of an object as perceived under some canonical illumination [99], such as white light with a flat spectrum, regardless of the illumination that is incident upon them.

Digital cameras have to rely on fast colour balancing algorithms integrated into their ISP pipeline, to estimate the colour temperature of a scene illumination in terms of the camera sensor response. The main purpose of an AWB algorithm is to estimate the colour temperature of a scene to compensate for undesired colour casts, and allow other ISP's pipeline algorithms to perform more accurately. However, since colour constancy algorithms have to calculate the colour temperature by processing only the digital pixel values, the calculation of temperature is ill posed. An AWB algorithm needs to accurately estimate

colour temperature given many possible illuminants and surface reflectances. Furthermore, an AWB module needs to be robust, run in real time with limited computational resources, and has to have a full understanding of the camera system. That is, the imaging sensor, optics, and modules of the ISP that runs before AWB have to be perfectly tuned so that colour statistics are as accurate as possible when estimating colour temperature of a scene [93].

When dealing with the colour constancy problem in digital photography it is necessary to take into account the colour preference of end users. Pleasing or memory colours [94] may be preferred over the accurate reproduction of colours. Thus, the target is not to reproduce accurate representation of colours in an image, but to reproduce what the pleasing memory colours of the majority of the consumers are [94]. That is. colours in a processed image should be reproduced closer to human memory colours, and not as colours reproduced in a light box of a photo lab. Since preferred colour reproduction is required instead of an accurate representation of them, the problem of colour constancy can be limited to finding the chromaticity of illumination. Despite the aim of preferred colour reproduction, it is still important to be able to estimate the chromaticity of illumination accurately, i.e. neutral colours in the image should be 100% neutral grey.

Determining preferred colours is a subjective task and it may be impossible to determine a method that will satisfy all viewers. For this reason, Arie Andries Kruithof proposed what is popularly named as, the Kruithof curve [95, 96]. In the Kruithof curve, the illuminance and colour temperature have a relationship to a preferred colour temperature given a scene with certain illuminance as shown in Figure 6.1 (a). Similar transition of colour temperature from orange/red to blue can be seen in the Planckian Locus chromaticity space as shown in Figure 6.1 (b). Kruithof’s curve relationship holds because it is expected to have cool colour temperatures in very bright conditions, and warm colour temperatures in low light conditions. However, recent research papers have shown that the Kruithof's curve is not always useful. It seems to be due to Kruithof’s omission to consider the influence of activities to the preference of lighting condition [96]. However, this is a starting point for determining and comprising a margin of preferred colours when trying to achieve colour constancy and preferred colours within a single system.



a) b)

Figure 6.1: a) Kruithof curve [128]. b) The CIE 1931 x, y chromaticity space [127]. It is also shown the chromaticities of the black body locus (Planckian Locus) at various light temperatures, and lines of constants correlate to the colour temperature.

The goals of this research are: 1) to evaluate and benchmark the ability of current state of the art colour constancy methods to estimate the colour temperature when they are used as an AWB module of an ISP pipeline, and 2) to propose a new ISP colour management system where colour constancy and pleasing colours can be achieved.

Colour constancy algorithms have been proposed and proved to perform well as a post-processing technique. However, from the literature reviewed, it seems like these methods have never been tested as a part of an ISP that works in the RAW Bayer domain, which has the colour channels R, Gr, Gb, and B. In addition, researchers seem to have overlooked the fact that the images used in their experiments had some kind of colour correction processing before they applied their methods. That is, static white, colour balancing, gamma correction, green equalise, black levels, which leads to a level of uncertainty of the proposed methods being able to correct colour casts on a balanced image, or actually estimate/compensate colour shifts due to different colour illuminants. Moreover, after analysing the image databases available in [105], which are the databases used by the previously proposed colour constancy approaches, none of them showed to have tested the methods to achieve colour constancy in challenging situations, also known in industry as “corner cases.” For instance scenes with: monotonic colours, green grass colour versus CWF lighting, opposite

colour objects' surface reflectance to the light colour temperature in the scene, sunset, low light conditions, absence of grey, noisy images, mixed lighting, cloudy scenes, and scenes with a predominant colour in the scene.

Another fact that researchers have overlooked when testing colour constancy methods is the dependency of other modules of the ISP's pipeline that affects accurate colour reproduction. Colour constancy will be affected by sensor-to-sensor variation, and by the ISP modules such as: black level correction, green equalisation, vignetting correction (which is a major problem in mobile phone cameras) and chroma noise. Hence, depending where the AWB module sits in an ISP pipeline, other ISP modules could have an impact on the final reproduction of colours in the image.

Most of the existing colour constancy methods used in digital cameras work in a 2-step operation. The first step is to estimate the illuminant, colour temperature, of the scene which is the focus of much research; the second step is to correct the image by either using the fixed diagonal transformation model based on von Krie's coefficient rule [108] shown in Equation 6.1, or by a linear non-restrictive transformation that can deal with non-orthogonal camera sensitivity function such as colour filter arrays in the sensor [93] shown in Equation 6.2.

$$[R' \quad G' \quad B']^T = \begin{bmatrix} \Gamma_{RR} & 0 & 0 \\ 0 & \Gamma_{GG} & 0 \\ 0 & 0 & \Gamma_{BB} \end{bmatrix} . [R \quad G \quad B]^T \qquad (6.1)\ [99]$$

$$[R' \quad G' \quad B']^T = \begin{bmatrix} \Gamma_{RR} & \Gamma_{RG} & \Gamma_{RB} \\ \Gamma_{GR} & \Gamma_{GG} & \Gamma_{GB} \\ \Gamma_{BR} & \Gamma_{BG} & \Gamma_{BB} \end{bmatrix} . [R \quad G \quad B]^T \qquad (6.2)\ [99]$$

So far, most of the colour constancy algorithms and AWB methods proposed in the literature have not been able to demonstrate an understanding of the issue of preferred versus neutral colours rendering. This is understandable given the subjective nature of the problem, and thus the reason for not being able to identify a unique solution to the colour constancy problem.

In this research a novel colour management system is proposed within an ISP pipeline that encompasses the colour constancy problem and the reproduction of pleasing colours in a final image processed by the ISP. The proposed ISP based colour management system has three stages: 1) the AWB module, where colour temperature is estimated and corrected, 2) colour balancing for achieving colour constancy in non-orthogonal camera colour functions, and 3) a cooling-warming approach to achieve pleasing colours in the final image processed by the ISP.

For clarity of presentation, this chapter is organised as follows: Section 6.2 provides a discussion of on the six colour constancy algorithms evaluated. Section 6.3 presents the proposed colour management system within an ISP pipeline. In Section 6.4, presents the experimental setup, and the experimental results obtained after testing the benchmark colour constancy methods described in Section 6.2 and the proposed novel colour constancy method presented in Section 6.3. Section 6.4 presents the ability of the proposed system to produce pleasant reproduction of colours based on Kruithof curve theory. To conclude this chapter, Section 6.5 presents the conclusion and future work.

6.2 Review of Colour Constancy Approaches

Within the context of research presented in this chapter only six of the fastest and most efficient colour constancy algorithms will be presented: Grey World (GW) [35], Probabilistic (Bayesian approach) [104], Max-RGB [98], Grey Edge 1st and 2nd order [101], Weighted Grey Edge [102], and Shades of Gray [100]. These algorithms estimate the R, G, and B components of the illumination vector *e* projected on the sensitivity of the camera's sensor function $\hat{I} = [\hat{I}r, \hat{I}g, \hat{I}b]^t$, and calculating and normalising the gain factor Γ for each colour channel as follows:

$$\hat{\mathrm{I}} = \frac{\mathrm{e}}{||\mathrm{e}||} \qquad (6.3)[99]$$

$$\Gamma_i = \frac{\max(\hat{\mathrm{I}})}{\hat{\mathrm{I}}_i}, i \in \{R, G, B\} \qquad (6.4)[99]$$

The Grey-world is regarded as a simple algorithm on the basis of its implementation. It is commonly used as a benchmark algorithm for colour constancy, and is based on the assumption that the average reflectance of a scene as captured by the camera is grey. The grey world algorithm estimates the illumination vector *e* by calculating the mean of the R, G, and B channels (Equation 6.5), and normalising the vector *e* to obtain the R, G, and B components of $\hat{I}$ using Equation 6.3. When the illumination is uniform over the field of view, the grey world algorithm is unbiased [99]. The illumination vector *e* is defined as:

$$e = \left[\frac{\sum R}{nm}, \frac{\sum G}{nm}, \frac{\sum B}{nm}\right]^T \quad (6.5)[99]$$

Where R, G, and B represent each of the colour channels of the image, and nm is the height and width of the image.

The Max-RGB approach to colour constancy is based on Land's explanation [98] as a mechanism to achieve colour constancy in digital photography similar to the way that the human visual system does. Land proposed that the HVS achieves colour constancy by detecting the area of highest reflectance in the field of view separately for long (red), medium (green) or short (blue) wavelengths, which correspond to the three types of cones in human eyes. The area of highest reflectance is then normalised by the response of each cone by the highest value [98]. The estimation of the illuminant is similar to the Grey-world algorithm in Equation 6.5, except for the fact that it estimates the maximum instead of the mean to get maximum sensor responses for each channel (Equation 6.6), and normalising the vector *e* to obtain the R, G, and B components of $\hat{I}$ using Equation 6.3.

$$e = [\max(R), \max(G), \max(B)]^T \quad (6.6)[99]$$

The Max-RGB algorithm produces accurate results when the scene contains a white patch, which reflects the entire spectrum of light evenly, or when the maximal object reflectance is the same for the R, G, and B channels. In practice, one should take care and make sure that the chosen maximal values accurately represent reflectance information from the scene. That is, clipped or saturated pixels due to the camera's limited dynamic range should

not be taken into consideration. It is usually a good practice to ignore pixels above a certain threshold (e.g. 95% of the camera's dynamic range) [99].

Finlayson and Trezzi [153] showed that Grey-world and max-RGB are two different instantiations of a more general colour constancy algorithm based on the Minkowski norm. Their method is labelled as Shades of Gray and it is computed by calculating and normalising the Minkowski norm for each colour channel to form the estimated illumination vector. In their setting, the grey world algorithm is obtained by setting $p = 1$, while max-RGB is the result of $p = \infty$. Shades of gray, is estimated as shown in Equation 6.7, followed by normalising the vector e to obtain the R, G, and B components of $\hat{I}$ using Equation 6.3. Finlayson and Trezzi concluded that using the Minkowski norm with $p = 6$ gave the best estimation results on their data set given [99].

$$e = \left[\left(\frac{\sum R^p}{nm} \right)^{\frac{1}{p}}, \left(\frac{\sum G^p}{nm} \right)^{\frac{1}{p}}, \left(\frac{\sum B^p}{nm} \right)^{\frac{1}{p}} \right]^T \quad (6.7)[99]$$

R, G, and *B* represent each of the colour channels of the image, and nm is the height and width of the image.

Another approach to colour constancy is the Grey Edge method [102]. It considers the 1st and 2nd order image derivatives, instead of using the image itself. The assumption in this method is based on the observation that the distribution of the derivatives of images forms an ellipsoid in the RGB space, of which the long axis coincides with the illumination vector. The Grey Edge method assumes that the average of the reflectance differences in a scene is achromatic. This method estimates the illumination vector e by first applying to the image a Gaussian filter with standard deviation σ to produce a $I\sigma$ scaled image (Equation 6.8), and then calculate Minkowski norm of each colour channel of the first order derivative of $I\sigma$, as shown in Equation 6.9, and normalising the vector e to obtain the R, G, and B components of $\hat{I}$ using Equation 6.3.

$$I\sigma(x,y) = e^{\left(-\left(\frac{(x-x_0)^2}{2\sigma_x^2}+\frac{(y-y_0)^2}{2\sigma_y^2}\right)\right)}, \quad I\sigma \in \{R,G,B\} \tag{6.8}[99]$$

$$e_i = \left(\sum_x \sum_y \left(\sqrt{\left(\frac{\partial}{\partial_x} I\sigma_i\right)^2 + \left(\frac{\partial}{\partial_y} I\sigma_i\right)^2}\right)^p\right)^{\frac{1}{p}} \quad p = \begin{Bmatrix} 1, grey\ edge \\ \infty, \max edge \end{Bmatrix} \tag{6.9}[99]$$

$$e_i \epsilon [R,G,B]$$

Another approach based on the use of distinct edge types is the Weighted Grey Edge [102]; this method improves the performance of edge-based colour constancy by computing a weighted average of the edges. Since such methods often assume that the scene is illuminated by a white light source, the automatic detection of such edges can become erroneous when the colour of the light source is different to white. The Weighted Grey Edge method is computed using an iterative photometric weighting scheme that sequentially estimates the colour of the light source and updates the computed edge weights. The foundation behind this approach is to fully exploit the information that is enclosed in the image, and simultaneously to increase the accuracy of the illuminant estimation and (specular) edge detection [102]. The Weighted Grey Edge method is computed using Equation 6.10, 6.11, and 6.3 to normalise *e*. *w(f)* is the weighting function and *k* is the value used to control the weighting.

$$\left(\int \left|w(f)^k f_{rgb,x}(x)\right|^p dx\right)^{\frac{1}{p}} = fe_{rgb} \tag{6.10}[99]$$

$$e = [fe_r, fe_g, fe_b]^{\mathrm{T}} \tag{6.11}[99]$$

The Bayesian approach based framework provides a description on how to use all of the information about the illuminant constrained in the sensor response. Three probability distributions play key roles. These are the prior, the posterior, and the likelihood. The prior probability describes what is known about the parameters before observing the illuminant data, while the posterior probability describes what is known after observing the illumination data. The Bayesian approach estimates parameters described by the vector *x*,

then the prior information is the probability density *p(x).* The likelihood *p(y|x)* expresses the relation between the data *y* and the parameters *x*. The likelihood may be thought of as the rendering Equation expressed as a probability distribution. Given the prior *p(x)* and the likelihood *p(y|x)*, the posterior rule, probability *p(x|y)* is computed by using Bayes' theorem shown in Equation 6.12.

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \qquad (6.12)[113]$$

Bayesian Colour Constancy Revisited [104] is one of the newest probabilistic approaches to colour constancy. This method is an improved version of the Bayesian approach from [103]. The Bayesian Colour Constancy Revisited approach models the observed image pixels with a probabilistic generative model, and decomposes them as the product of unknown surface reflectances with an unknown illuminant. Then, by using Bayes' rule, it obtains a posterior for the illuminant, and from this, an estimate with the minimum risk. That is, the minimum expected chromaticity error. In order to estimate the illuminant, it calculates the likelihood of the observed *Y* image data for an illuminant *L* as follows: Let *Y* be an image pixel with three colour channels. The pixel value is assumed to the reflection of a single light source of a Lambertian surface with power *Lrgb*, and the proportion of light reflected *Xrgb*. Thus the model for a given pixel is *Yrgb = lrgbXrgb*, and, since it is assumed that the illumination reflectances are independent, then *p(X,l) = p(X)p(L)*. The reflectances *p(x)* are modelled by assuming exchangeability of reflectances. With this assumption, it is possible to define a set of probabilities over the reflectances histogram *(n1,…,nk)* in *Y*, where *nk* are the number of reflectances in the k^{th} bin of the histogram. The prior illumination distribution *p(L)*, is given by using the reflectance distribution of the Grey-world algorithm. In order to estimate the posterior, Equation 16 is used where η is an exponent that helps with cross validation. In [104] the value of $\eta = 0.001$. The posterior probability *p(Y|L)* is derived as follows:

$$p(X) = \propto f(n_1, \dots, n_k) \qquad (6.13)[104]$$

$$f(n_1, \dots, n_k) = \prod_k m_k^{vk} \qquad (6.14)[104]$$

$$vk = n\frac{\tanh(\lambda n_k)}{\sum_s \tanh(\lambda n_s)} \quad (6.15)[104]$$

$$p(Y|L) = \int_x \left(\prod_i p(y(rgb)|L, x(rgb)) \right) p(X)dX \quad (6.16)[104]$$

$$= |L^{-1}|^n p(X = L^{-1}Y) \quad (6.17)[104]$$

$$p(L|Y) \propto |L^{-1}|^n p(X = L^{-1}Y)p(Y) \quad (6.18)[104]$$

$$p(L|Y) \propto p(Y|L)^\eta p(L) \quad (6.19)[104]$$

6.3 Proposed ISP Colour Management System

The proposed colour management system within an ISP pipeline encompasses three stages. First, the colour temperature of the illuminant is estimated (AWB module); then, a colour correction/balancing matrix is applied (CCM); followed by a cooling-warming process to achieve camera end-users preferred colours depending on the colour temperature of the scene. The diagram in Figure 6.2 shows how the proposed system fits into the pipeline of an ISP. The proposed approach is designed in such a way that it is robust enough to cope with sensor-to-sensor variation, on either mobile phones or industrial cameras, and at the same time flexible enough to be able to produce pleasant colours in images by following Kruithof curve theory, and also to exceed industry image quality standards.

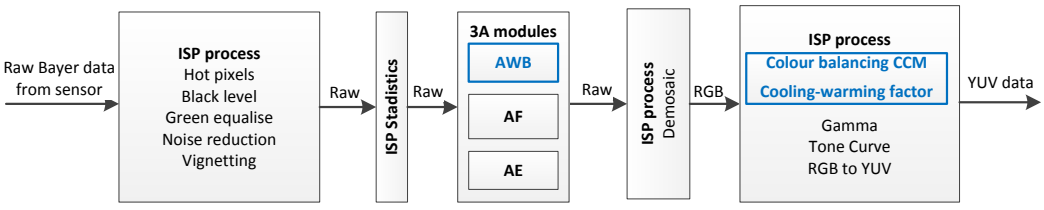


Figure 6.2: The integration of the proposed colour constancy method inside an ISP pipeline.

The system works by estimating the colour temperature of the illuminant and compensating accordingly so that the image is 100% neutral under any type of illuminant. This is followed

by the calculation of the colour temperature in the AWB module; the image's colours are balanced by the colour correction matrix (CCM) depending on the colour temperature estimated in the AWB module. After these two steps, colours in the image should be subjectively and objectively constant under any lighting condition. The last step of the system is to make the image more pleasant to camera users by adding a cooling-warming effect to the image depending on the colour temperature estimated in the AWB module.

In order to estimate colour temperature, a new Bayesian method adapted from a machine learning technique to estimate the illuminant of a scene is proposed. As mentioned in section 6.1, the colour constancy problem is ill posed if the only information used to estimate colour temperature is the pixel intensity values of an image. For this reason, this system uses the luminance dimension as an extra piece of information when estimating colour temperature. The luminance dimension is characterised by the relationship of the camera exposure value (EV) to brightness (lux) of a scene. This is because low colour temperatures are more likely to be present in low lightning conditions, and high colour temperatures are more likely to be present in very bright lighting conditions.

In the following sections, a full description of the proposed approach is presented.

6.3.1 Sensor characterisation

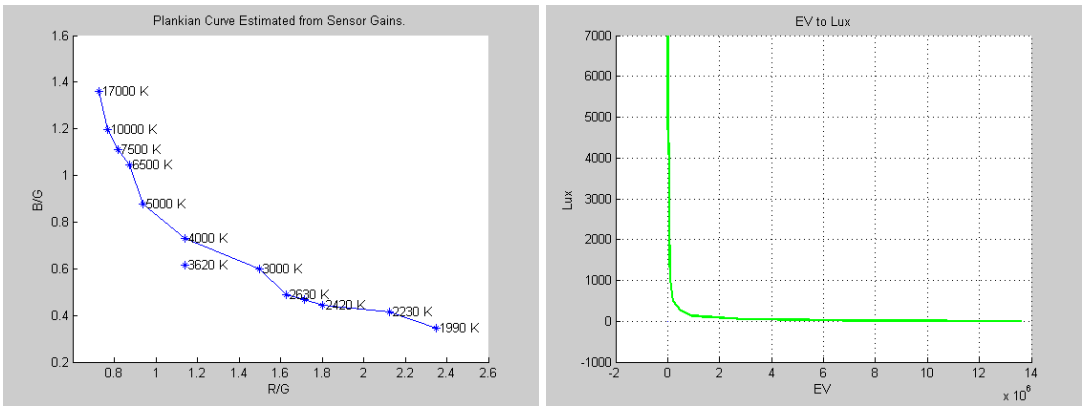
In this stage, the behaviour of a camera system under different colour temperature lighting is characterised. A large set of illuminants ranging from 1800 kelvin (k) to 17000K, including cool white fluorescent (CWF) lighting are considered. The images used for the characterisation are taken under a control environment using professional equipment: photo lab light box, CIE standards colour temperature light bulbs, and a 24-ColourChecker Macbeth chart for colour measurements.

The behaviour of a camera sensor is analysed by plotting the ratio of the colour channels red to green and blue to green from all illuminants captured. This allows knowing the Planckian

curve or response to colour temperatures of the sensor. This analysis is shown in Figure 6.3(a).

In order to estimate the colour matrices for the colour balancing stage, three lighting conditions are considered: incandescent (A), fluorescent (TL84), and daylight (D65).

The final measurement taken for characterising the sensor is the relationship of EV to lux levels in the scene. The EV value is calculated as the product of exposure time, digital and analogue gains, and aperture. The lux levels are recorded by using a lux meter in the scene. The lux levels measured range is from 10 lux to 25000 lux, which is done by increasing the light intensity in the photo lab. Figure 6.3 (b) shows the EV to lux relationship.



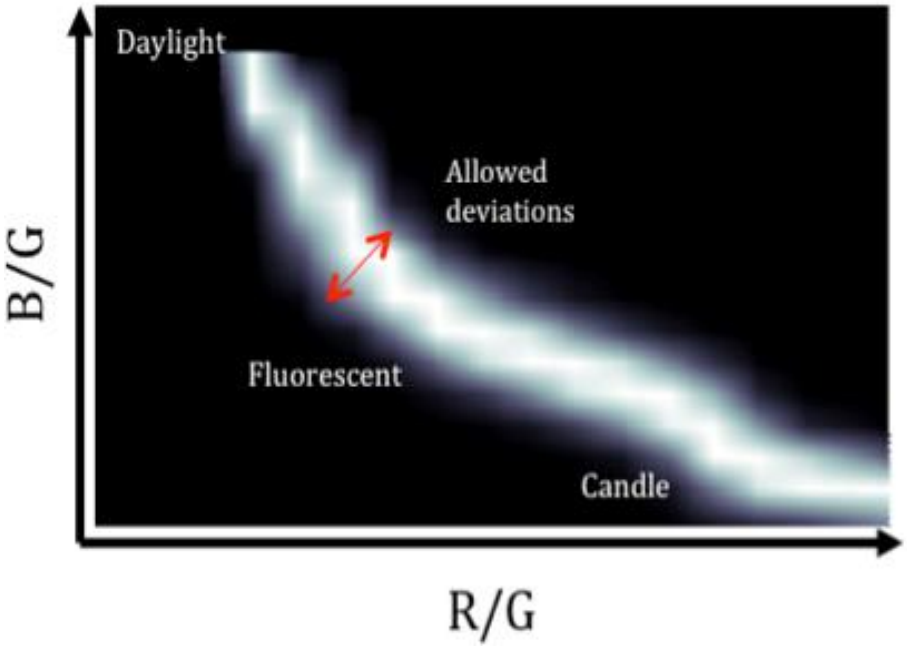
a) b)

Figure 6.3: SONY NEX-5 colour response. a) Characterisation of the Planckian locus curve. b) Relationship of EV values to Lux levels

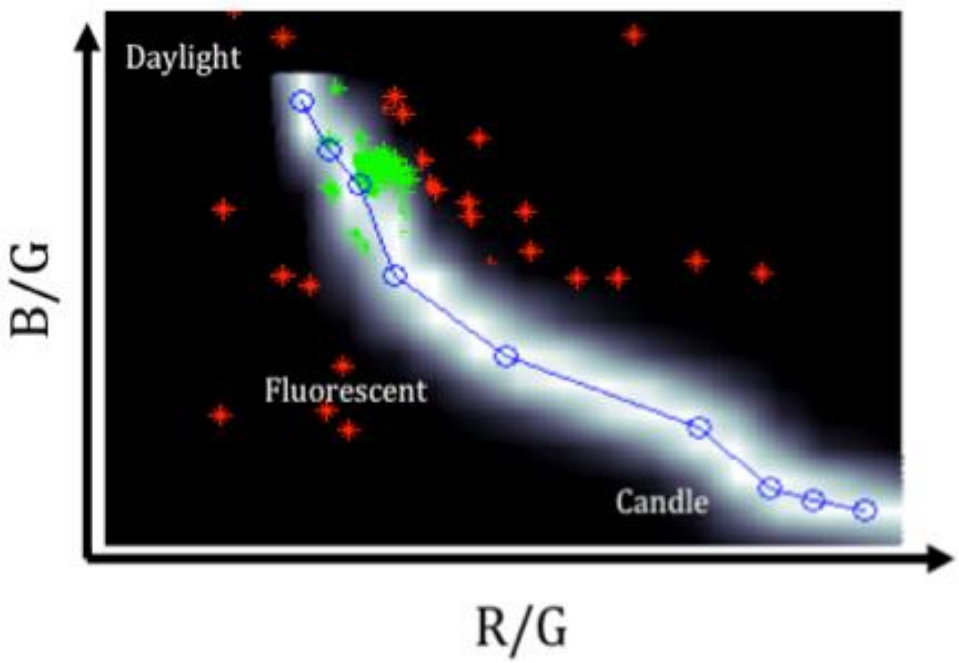
6.3.2 Colour temperature estimation

The illumination of colour temperature estimation is a probabilistic data driven method, derived from localised MxN statistics zones in the image generated by an ISP. This method uses the concept of the constrained colour temperature curve, and it looks for various colours in the image by using negative and positive deviation with a Gaussian distribution along the characterised Planckian curve (Figure 6.4 (a)). Thus, the correct red and blue gains

of the sensor are confined to a small subset of all possible gains as shown in Figure 6.4 (b). The lower and upper limits of the curve can be set to the desired range of illuminants that the camera system should encounter. Negative and positive deviation along the Planckian curve also allows green channel variations because in reality the gains do not lie exactly on the curve as such, but near the curve. In order to follow a more comprehensive probabilistic approach, a machine learning classifier technique is used, the Naïve Bayes Classifier, to determining the probability of high or low temperature in the scene.



a)



b)

Figure 6.4: a) Negative and positive deviations along the Planckian curve. b) Plot of valid (green) and non-valid (red) zones R:G, B:G ratios for estimation of colour temperature.

The Naïve Bayes algorithm is a classification algorithm based on Bayes rule that assumes the attributes *X1… Xn* are all conditionally independent of one another, given *Y* [106]. The value of this assumption is that it simplifies the representation of *P(X|Y)*, and the problem of estimating probabilities from statistical data.

The Naïve Bayes Classification theory is derived as follows: Let *X = {X1, X2}*, which leads to *P(X|Y) = P(X1, X2|Y)*. From the general property of probabilities it can be rewritten *P(X|Y)* as *P(X|Y) = P(X1| X2,Y)P(X2|Y)*, and by following the conditional independence of a random variable *X*, *Y* and *Z*, defined as: *X* is conditionally independent of *Y* given *Z*, if and only if the probability distribution governing *X* is independent of the value of *Y* given *Z*; then *P(X|Y)* can be rewritten as in Equation 6.20:

$$p(X_1 \dots X_n|Y) = \prod_{i=1}^{n} P(X_i|Y) \tag{6.20}$$

Assuming that *Y* is any discrete-valued variable, and the attributes *X1… Xn* are any discrete or real-valued attributes. Then, the desired classifier will output the probability distribution over the possible values of *Y*, for each new instance of *X*. The expression of the probability that *Y* will take on its K^{th} possible value, according to Bayes rule, and by the assumption the *Xi* are conditionally independent given *Y* is given in equation 6.21:

$$p(Y = y_k|X_1 \dots X_n) = \frac{p(Y = y_k)\prod_i p(X_i|Y = y_k)}{\sum_j p\big(Y = y_j\big)\prod_i p\big(X_i|Y = y_j\big)} \tag{6.21}$$

Equation 6.21, the fundamental equation of the Naive Bayes Classifier, shows how to calculate the probability that *Y* will take on any given value, given the observed attribute values of X_{new}*={ X1,…,Xn}* and given the distributions *P(Y)* and *P(Xi|Y)* estimated from the training data. In order to determine the most probable value of *Y*, the following classification rule given by equation 6.22 is introduced:

$$Y \leftarrow argmax_{y_k} \frac{p(Y = y_k)\prod_i p(X_i|Y = y_k)}{\sum_j p\big(Y = y_j\big)\prod_i p\big(X_i|Y = y_j\big)} \tag{6.22}$$

Because the denominator does not depend on y_k, equation 6.22 simplifies to:

$$Y \leftarrow argmax_{y_k} P(Y = y_k) \prod_i P(X_i|Y = y_k) \tag{6.23}$$

Thus, in order to use the Naïve Bayes classifier to determine the probability of a scene being high or low temperature, the assumptions that high colour temperatures are only present at high brightness levels with low red gains in the image, and low colour temperatures are present in low light conditions with high red gains in the image is made. That is, a high temperature scene, or a low temperature scene depends on the red gain values in an image and lux levels attributes.

To ensure that that the probability of an illuminant is correct, the Bayes Theorem is used in conjunction with a Naïve Bayes Classifier as follows.

The class relationships of lux levels to high and low temperature, and red gains to high and low temperature are calculated by computing the mean μ_y, variance σ_y^2 of *X = {lux, red gains}* associated with class *Y = {low temperature, high temperature}*. The probability of Y_i given some value *v*, the *R:G* gain for each statistic zone and lux value of the scene, is given by equation 6.24:

$$P(X = v|Y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(u\mu_y)^2}{2\sigma_y^2}} \tag{6.24}$$

The Naïve Bayes classifier is calculated by getting the maximum posterior probability of the *m* number of *Y* classes and *n* number of *X* attributes as:

$$Y_{iMax} \leftarrow \text{argmax}_m \frac{\prod_{j=1}^{n} P(Y_i)P(X_j|Y_i)}{\sum_{i=1}^{m}\prod_{j=1}^{n} P(Y_i)P(X_j|Y_i)} \tag{6.25}$$

In order to satisfy the Bayes' theorem, it is required to estimate the probability of the localised 1 to *w*, ISP statistics zones *Z*, defined as the Gaussian distribution drop-off style from the Planckian curve as:

$$p(Z_{1:w}) = e^{-\frac{(\text{distance of } Z_{1:w} \text{ from planckian curve})^2}{\text{deviationfrom planckian}}} \tag{6.26}$$

By Using Bayes' theorem, each of the statistical zones is weighted according to $P(Z_{1:w})$ and $P(YMax_{1:w})$.

$$P(\text{Colour Temperature} \mid Z_{1:w}) = P(Z_{1:w}) Y_{iMax_{1:w}} \tag{6.27}$$

Hence, the illumination vector *e* is given by the sum over all probability weighted *R:G*, and *B:G* zones as:

$$e_i = \frac{\sum_{i=1}^{w} Z_i \, P(\text{colour temperature} \mid Z_i)}{\sum_{i=1}^{w} P(\text{colour temperature} \mid Z_i)} \tag{6.28}$$

$$e_i \epsilon [R/G, 1, B/G]$$

The image can be corrected by composing the vector *e* with the gains calculated from Equation 6.28 and then normalising using Equation 6.3 and correcting the illuminant using equation 6.1. In order to estimate the colour temperature in Kelvin, the *R/G, B/G* gains will indicate the position in the characterised Planckian curve as in Figure 6.3 (a), the colour temperature of the scene

6.3.3 Colour balancing driven by AWB

Colour accuracy is an important but at the same time ambiguous image quality factor. Accurate colours may be critical in technical or medical photography, but less important in consumer photography, where many camera users prefer enhanced colour saturation, particularly in memory colours such as foliage, sky, skin, etc. Thus, this research puts the emphasis on that fact, accurate colours are not the same as pleasing colours.

Colour accuracy can be affected by Bayer colour filter arrays in the imaging sensor, signal processing, optical characteristics, and lighting conditions. Thus, in addition to making the image's grey colours neutral by the AWB module, all spectral colours need to be compensated to achieve colour constancy under any illumination and camera characteristics. Achieving colour accuracy in all spectral colours, also known as colour correction matrix (CCM), can be accomplished by applying a 3x3 correction matrix to the image using equation 6.2, a linear non-restrictive transformation.

The challenge of colour balancing an image by the CCM is to find the 9 coefficients that will produce the most accurate representation of colours. In order to estimate the 9 coefficients of the CCM, the colour response is measured by using the widely available 24-patch Macbeth ColourChecker under different illuminations. The CCM colour accuracy is obtained by minimising the colour *ΔE*ab* and *ΔC** errors in the CIELAB (L*a*b*) colour space where *L** is luminance, *a** is colour on a green-red scale, and *b** colour in a blue-yellow scale. *ΔE*ab* and *ΔC** are the Euclidian distances between the L*a*b* values. *ΔE*ab* and *ΔC** errors are calculated as follows:

$$\Delta C^* = \sqrt{((a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2)} \qquad (6.29)[107]$$

$$\Delta E_{ab}^* = \sqrt{((L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2)} \qquad (6.30)[107]$$

Note that the *ΔC** value ignores both exposure and saturation error and only looks at the deviation in hue of the Macbeth colour patches from the ideal target. Thus, some colour error *ΔC** may be acceptable since ideal colours are not necessarily right for consumer products, thus, some subjective judgment is needed when generating the CCM. In this research, a popular commercial image quality standard software named Imatest is used to generate the 3 CCMs, which are measured at incandescent, TL84, and D65 light colour temperatures. In order to generate the 3 CCMs, Imatest uses a mean squares error metric in an optimiser to reduce the *ΔE*ab* and *ΔC** errors in the CCM [107]. The optimiser constraint is that the row and columns have to sum to 1. It was also used a patch weighting function to assign a higher or lower weight to different patches to avoid oversaturation,

colour clipping, or to improve the colour Δ^* error. Figure 6.5 (a) shows an example of the Macbeth chart and in Figure 6.5 (b) the plot of the colour errors in CIELAB space.

a) b)

Figure 6.5: a) 24-patch Macbeth chart. b) Plot of the colour errors in CIELAB space

The reason for having three CCMs is because colours vary under different illuminations. Thus, if it is known what the colour temperature is, it is possible to apply the correct CCM to achieve colour accuracy under any illumination. The correction is done using equation 6.2, with the addition that the CCMs are alpha blended from CCM_A to CCM_{TL84}, and CCM_{TL84} to CCM_{D65} depending on the colour temperature of the scene, estimated in the AWB module.

6.3.4 Cooling-warming factor

The last step of the system is to make colours more pleasant to camera users by adding a cooling-warming effect to the image depending on the colour temperature estimated in the AWB module. To this point, the image processed should be colour balanced in all spectral and neutral colours, so that any extra colour shifts are performed in a controlled manner.

In this research it was adopted Kruithof theory on pleasing colours which can be interpreted as: *colours in an image appear to be bluer as colour temperature increases, and redder as colour temperatures decrease*. The cooling-warming effect is applied using equation 6.31, where R_{wc}, G_{wc}, B_{wc}, are the cooling-warming factors calculated from a piecewise linear

equation depending on the colour temperature of the scene *T*, and ω a weighting factor applied to each of the colour channels.

$$[R' \quad G' \quad B']^T = \begin{bmatrix} 1 + R_{wc}\omega_r & 0 & 0 \\ 0 & 1 + G_{wc}\omega_g & 0 \\ 0 & 0 & 1 + B_{wc}\omega_b \end{bmatrix} . [R \quad G \quad B]^T \tag{6.30}$$

$$C_{wc} \epsilon \{R_{wc} G_{wc} B_{wc}\} \leftarrow C_{wc}\, f(T)$$

$$f(T) = \begin{cases} -400000\, x + 0.2\,, & T < 5000 \\ 400000\, x + 0.2\,, & T \geq 5000 \end{cases}$$

$$f(\omega, T\,) = \begin{cases} \omega_r = 1, \omega_g = 0.33, \omega_b = 0, & T < 5000 \\ \omega_r = 0, \omega_g = 0.33, \omega_b = 1\,, & T \geq 5000 \end{cases} \tag{6.31}$$

Since the above is a subjective matter, the *f(T)*, and *f(ω, T)* could slightly change according to the manufactures' preference, when characterising the canonical camera system. In this research, the tuning of the cooling-warming piecewise linear equation is estimated by visually matching the colours displayed in an EIZO ColorEdge calibrated monitor to the colours perceived by the human eye in the light box at different colour temperature illuminants.

6.4 Evaluation of the Proposed Colour Constancy Method

In this section, the experimental setup, and results obtained comparing the performances of the state-of-the-art colour constancy methods described in Section 6.2 and the proposed colour constancy method described in Section 6.3.2 are presented

Each of the colour constancy methods previously described have been benchmarked in previous research, but not yet identified as a comprehensive solution to the problem at hand. It is believed this is because researchers have tested and developed colour constancy

methods that differ in purpose and functionality. Moreover, the ultimate aim in resolving the colour constancy problem in digital photography has not been yet well defined. For instance, is the aim to reproduce an image with pleasing memory colours, or to reproduce accurate colours as in a photo lab regardless of the illuminant in scene? Furthermore, how do the methods achieve colour constancy? Is it i) by removing the colour cast in already processed images as a post-processing technique that needs a previous knowledge set by humans, or, ii) by defining an automated approach for estimating/compensating the illuminant's colour temperature inside an ISP, as a pre-processing technique?

Authors believe that a further reason for why a solution to the colour constancy problem has not been yet identified, is because experiments in previous research have not been consistent. This may be due to the inability to accurately identify all image-processing algorithms that take palace inside commercial ISPs being used. For example, researchers have claimed that the images gathered for the experiments were in RAW format. However, a closer look revealed that the images were read back within the computer by using image-processing software such as "Dcraw" to produce the corresponding RGB images [104, 111]. The type of software that read RAW files and convert them into RGB, undergo a computer simulation of a typical hardware ISP pipeline and thus all typical colour correction processes of an ISP such as static white/colour balancing CCM, gamma correction, black level correction, demosaicing, hot pixel elimination among other processes are applied according to the metadata imbedded in the RAW image file. Thus, the RAW images converted to RGB were not colour correction free and data was not linear. In other research, images were captured in RGB disabling the camera's AWB module and setting it to manual outdoors white balance mode [99, 112]. Thus, the images gathered yet again were not in the linear form, and had colour corrections applied.

Consequently, if the aim of a colour constancy method is to estimate/compensate colour shifts under any illuminant instead of removing colour casts from an already balanced image, the experiments performed in previous research had many uncontrolled factors that could have added uncertainty to the ability of the proposed methods to work as an ISP's AWB module. All available databases that showed the above can be found in http://colorconstancy.com [105].

In order to illustrate the discrepancy of the inputs to the experiments performed in previous research, Figure 6.6 (a) shows an image captured in RAW Bayer format at 2600K without any colour correction applied (only demosaiced for illustration purposes); this would normally be the input to the AWB module in an ISP. However, when capturing the same image using an ISP with manual WB set to indoors-white balance, the RGB output image is colour corrected and presents an orange colour cast as shown in Figure 6.6 (c). Figure 6.6 (b) shows the same image obtained using the same ISP but with the AWB module enabled.

Hence, it is important to define the key focus of the colour constancy method as either: a) remove colours casts from an already balanced image, or b) estimate and correct the illuminant colour shifts from an image coming straight from the imaging sensor. This research focuses on the approach defined under (b) above.

In this research, it is believed that it is important that the images tested should not have any colour correction algorithm applied prior to AWB calculations. Otherwise, colours in the image will significantly change and thus the focus of the colour constancy method will not only be balancing neutral and spectral colours in a controlled environment, but also to achieve colour constancy in an uncontrolled environment. This is because of the lack of knowledge of how the colour correction processes were applied inside the ISP used, and consequently the data in the image may not be linearised again after gamma correction, and include compression artefacts.

a) b) c)

Figure 6.6: Comparison of an image captured in different formats and white balance modes at 2600K. a) Raw image without any colour correction applied and only demosaiced for illustration purposes. b) Image processed by an ISP with AWB module enabled. c) Image processed by an ISP with AWB module disabled and set to indoors manual white balance.

6.4.1 Experiment setup

In the research presented in this work, the colour constancy methods previously described in section 6.2, are placed and tested inside an ISP pipeline, where the estimation of illuminant is achieved in RAW Bayer format, and without any colour correction applied to the image, except for the subtraction of black levels offset of the Bayer sensor. The block diagram of the ISP pipeline used in the experiments to be conducted is shown in Figure 6.7. By using a generic, simplified, and a known ISP for testing colour constancy methods, it is possible to benchmark in a controlled environment the ability of the proposed methods to produce a 100% neutral image by estimating and correcting the illuminant's colour temperature shifts inside an ISP pipeline. That is, the main function of an auto-white balance module of a typical camera system.

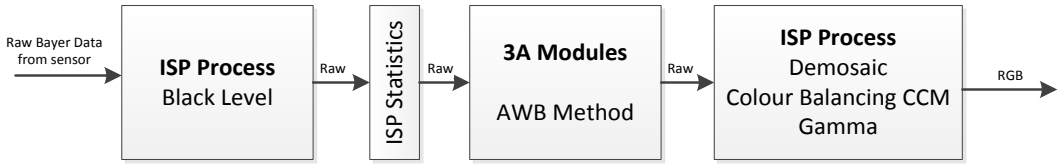


Figure 6.7: Simplified ISP pipeline used in experiments.

In order to test the colour constancy methods, a camera characterisation is needed for knowing the black level offset of the Bayer sensor, and a CCM generation under D65 lighting.

The black level offset was obtained by capturing a black RAW image at analogue gain of 1. In order to estimate the black offset, the mean values of each of the four-colour channels were measured. The CCM under D65 lighting was generated as in section 6.3.3, the type of Gamma correction used was sRGB, and a bilinear demosaicing algorithm was used to convert RAW Bayer image to RGB. The simplified ISP was implemented using MATLAB and the source code [109] of the benchmark algorithms described in Section 6.2.

The sensor characterisation was performed in three different camera systems: i) SONY NEX-5, ii) Canon Power-shot G10, and iii) using an OmniVision OV8835 image sensor. In order to perform the experiments, a total of 1000 RAW images using the three camera systems mentioned above were captured. This database contains a substantial number of scenes with different properties such as: indoor and outdoor conditions, monotonic colours, green grass colour versus CWF lighting, opposite colour objects' surface reflectance to the light colour temperature in the scene, sunset, low light conditions, absence of grey, noisy images, mixed lighting, cloudy scenes, and scenes with blue-sky.

In order to obtain objective measurements, an Opteka18% Digital Grey Card, which has showed a high degree of spectral neutrality under varying lighting conditions [110] was placed in the scene.

After processing all sets of images, the R, G, and B pixel values from the grey chart in the images were measured in order to estimate the mean, maximum, and standard deviation, of the *R/G* and *B/G* ratios. In addition to the metrics mention above, the deviation from neutral grey ΔE_x was also calculated by estimating the Euclidian distance from the measured R/G, and B/G ratios to the reference white point, *CR/G = 1*, and *CB/G = 1*. ΔE_x is defined as in equation 6.32.

$$\Delta E_x = \sqrt{\frac{\left(C_{R/G,ref} - C_{R/G,x}\right)^2}{C_{R/G,ref}} + \frac{\left(C_{B/G,ref} - C_{B/G,x}\right)^2}{C_{B/G,ref}}} \tag{6.32}$$

By setting up this experiment, it was possible to benchmark the ability of colour the constancy methods to perform as an AWB module of an ISP. Note that for this experiment,

the CCM, and cooling-warming factor driven by AWB described in Section 6.3.3 and 6.3.4 were disabled. At this point, the particular focus was aimed to know the methods' robustness and ability to produce 100% neutral images under different illuminations.

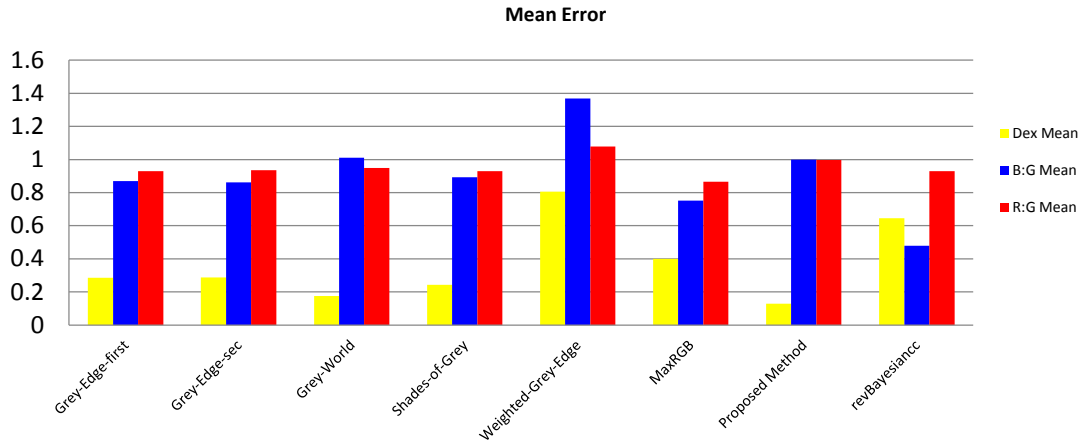
6.4.2 Experimental results

Table 6.1 shows the mean, standard deviation, and maximum error of the *R/G*, *B/G* ratios, and ΔE_x objective measurements obtained from the experiments. The mean value of the *R/G*, *B/G* ratios, and ΔE_x shows the general behaviour of the methods to either, overestimate or underestimate the amount of red or blue gains in order to balance the image. The maximum indicates by how much the method overestimated or underestimated the colour temperature of the scene, and the standard deviation shows the variation of gains applied over the whole set of images tested. Ideal values for the measurements taken are if the *R*, *B*, and *G* values measured in the grey chart have the same intensity value. That is, $R/G = B/G$, which leads to a $\Delta E_x = 0$, and maximum error = 0.

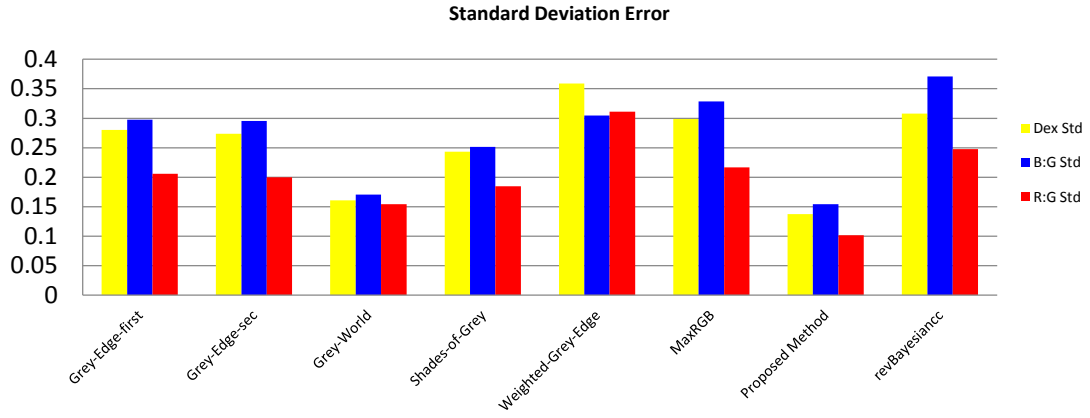
	R:G		B:G		ΔEx	
Grey-Edge-first order	Mean	0.93	Mean	0.87	Mean	0.29
	Std	0.21	Std	0.30	Std	0.28
	Max	2.03	Max	2.18	Max	1.36
Grey-Edge-second order	Mean	0.93	Mean	0.86	Mean	0.29
	Std	0.20	Std	0.30	Std	0.27
	Max	2.09	Max	1.92	Max	1.43
Grey-World	Mean	0.95	Mean	1.01	Mean	0.18
	Std	0.15	Std	0.17	Std	0.16
	Max	1.50	Max	1.72	Max	0.93
Shades-of-Grey	Mean	0.93	Mean	0.89	Mean	0.24
	Std	0.18	Std	0.25	Std	0.24
	Max	1.96	Max	1.85	Max	1.33
Weighted-Grey-Edge	Mean	1.08	Mean	1.37	Mean	0.81
	Std	0.31	Std	0.30	Std	0.36
	Max	4.33	Max	2.26	Max	3.38
MaxRGB	Mean	0.87	Mean	0.75	Mean	0.40
	Std	0.22	Std	0.33	Std	0.30
	Max	1.84	Max	2.36	Max	1.65
Proposed Method	Mean	1.00	Mean	1.00	Mean	0.13
	Std	0.10	Std	0.15	Std	0.14
	Max	1.73	Max	1.75	Max	0.90
revBayesiancc	Mean	0.93	Mean	0.48	Mean	0.65
	Std	0.25	Std	0.37	Std	0.31
	Max	2.77	Max	1.50	Max	2.06

Table 6.1: Objective measurements obtained from the benchmarked colour constancy performance within an ISP.

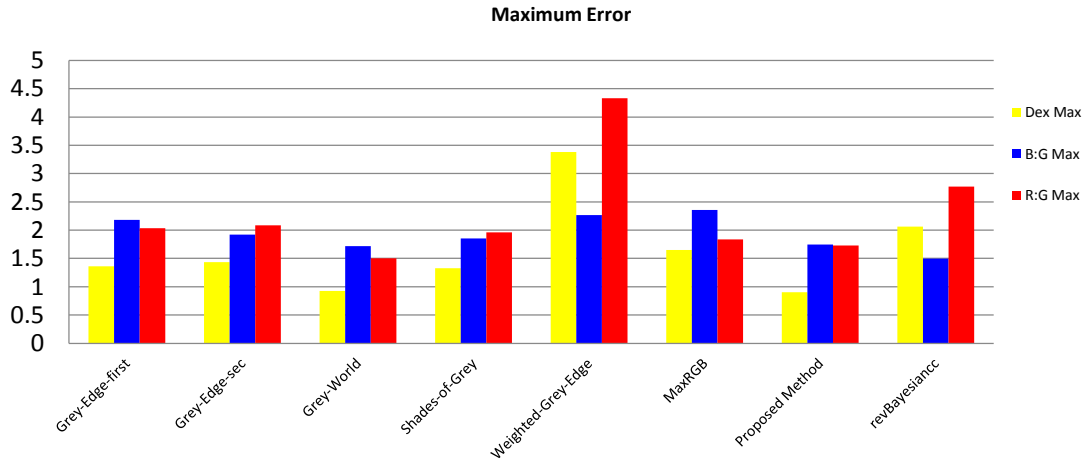
The plots of the objective measurements are shown in Figure 6.8. These plots clearly show the superiority of the proposed method in terms of robustness, Figure 6.8 (a) and (b)), and maximum gains errors in Figure 6.8 (c).



a)



b)



c)

Figure 6.8: Plots analysis of objective measurements. a) Plot of the mean *R:G*, *B:G*, and ΔE_x. b) Plot of the standard deviation of *R:G*, *B:G*, and ΔE_x. c) Plot of the maximum error of the *R:G*, *B:G*, and ΔE_x.

According to the data analysed, Grey-world based methods have a similar behaviour and are capable of balancing images fairly well in the presence of grey in the scene. However, these methods are very sensitive to saturated colours, full monotonic colours scenes, mix lighting, absence of grey, and opposite colour objects' surface reflectance to the light colour temperature in the scene.

In order to illustrate the accuracy of the proposed method and the importance of AWB algorithms to accurately balance an image, Figure 6.9(a) illustrates a colour checker image perfectly balanced, and as a reference, the same image is shown in figures 6.9(b - i) with ±5% offset in the red and blue channel.

Similar to Figure 6.9, Figure 6.10(a) illustrates a colour checker image perfectly balanced, and as a reference, the same image is shown in figures 6.9(b - i) with ±10% offset in the red and blue channel.

Figure 6.9 and Figure 6.10 illustrate the average performance of the colour constancy algorithms tested, where the mean of R:G and B:G ratios were between ±10% from the neutral grey. This kind of colour shift in images would not be acceptable in any industry standards. In contrast, the proposed algorithm R:G and B:G ratios mean values were between ±1%. This variation does not introduce visible colour casts as shown in Figure 6.11.

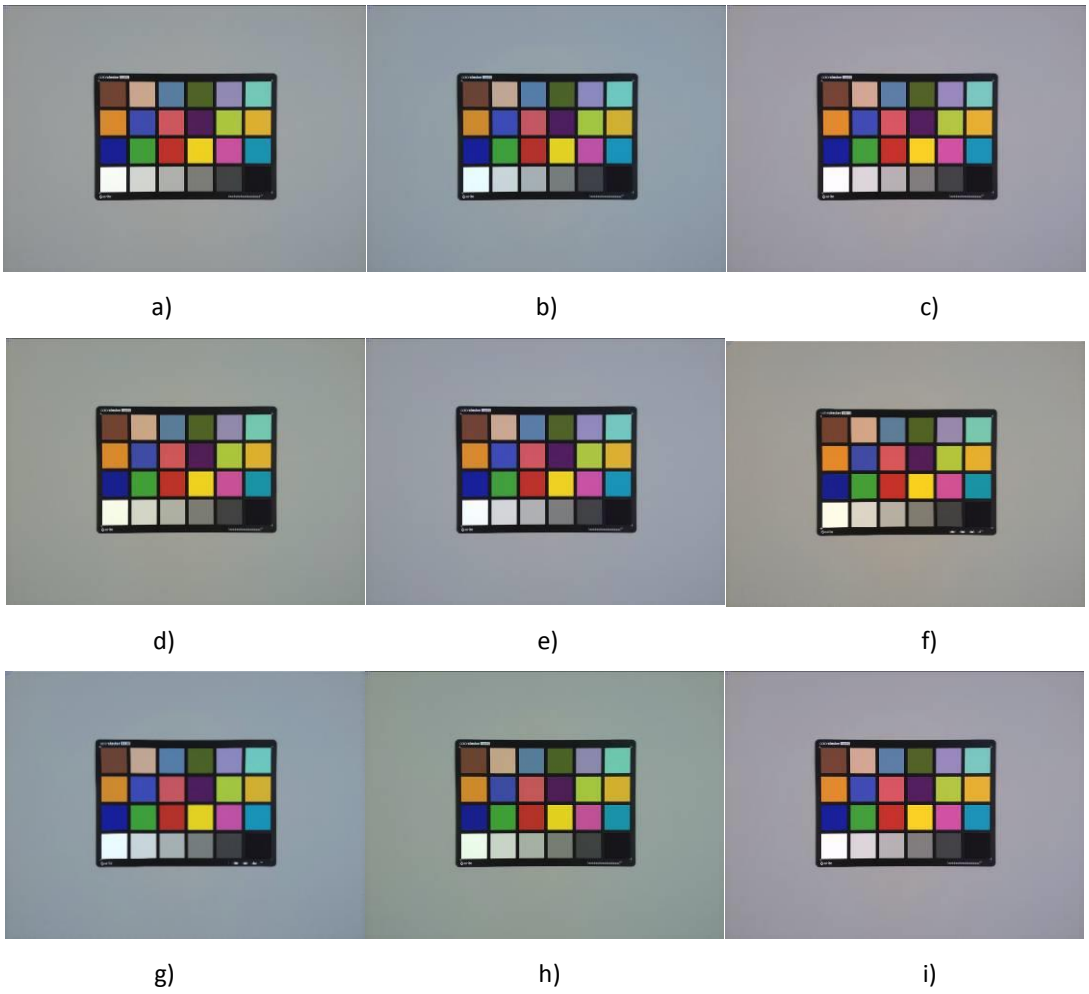


Figure 6.9: a) Original image, b) -5% offset in the red channel, c) +5% offset in the red channel, d) -5% offset in the blue channel, e) +5% offset in the blue channel, f)-5% offset in the red channel and +5% offset in the blue channel, g) +5% offset in the red channel and -5% offset in the blue channel, h) -5% offset in the red channel and -5% offset in the blue channel, i) +5% offset in the red channel and +5% offset in the blue channel

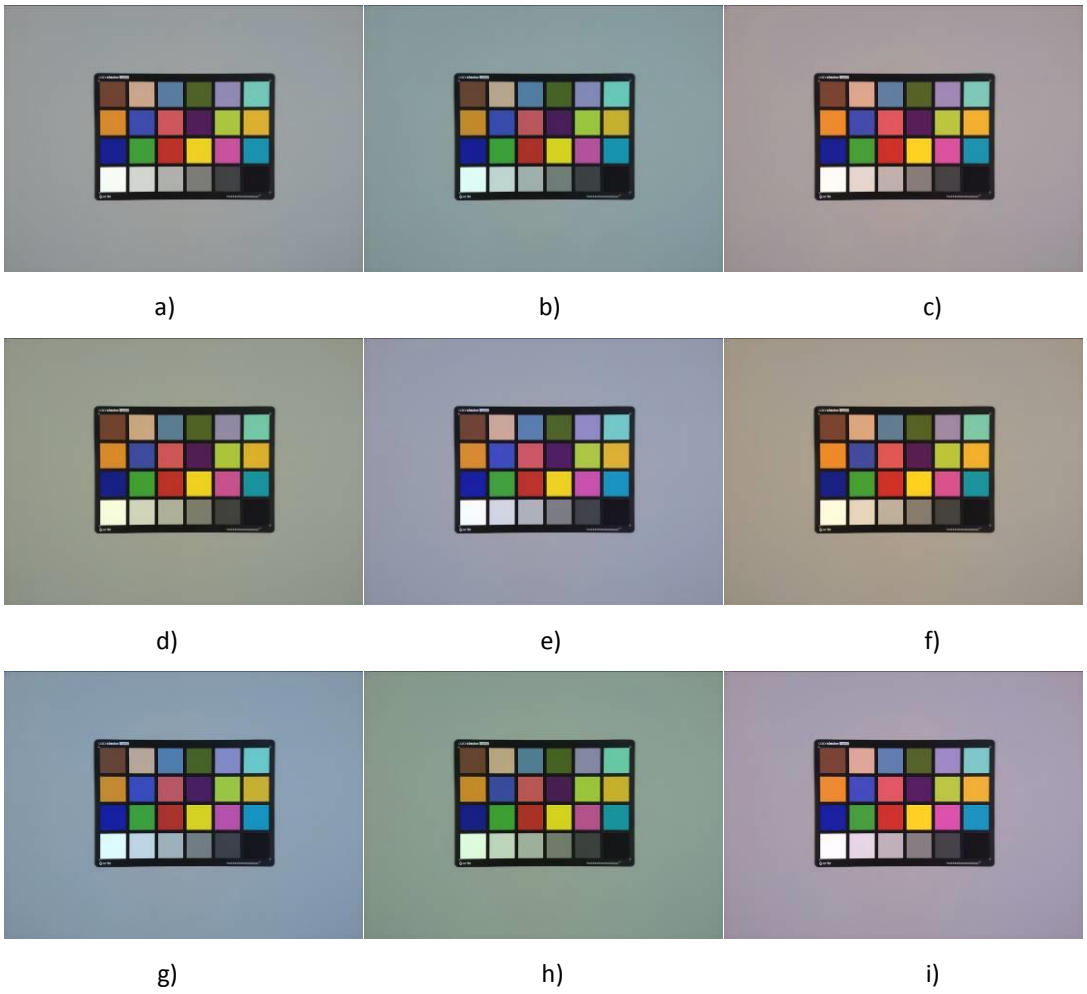


Figure 6.10: a) Original image, b) -10% offset in the red channel, c) +10% offset in the red channel, d) -10% offset in the blue channel, e) +10% offset in the blue channel, f)-10% offset in the red channel and +10% offset in the blue channel, g) +10% offset in the red channel and -10% offset in the blue channel, h) -10% offset in the red channel and -10% offset in the blue channel, i) +10% offset in the red channel and +10% offset in the blue channel

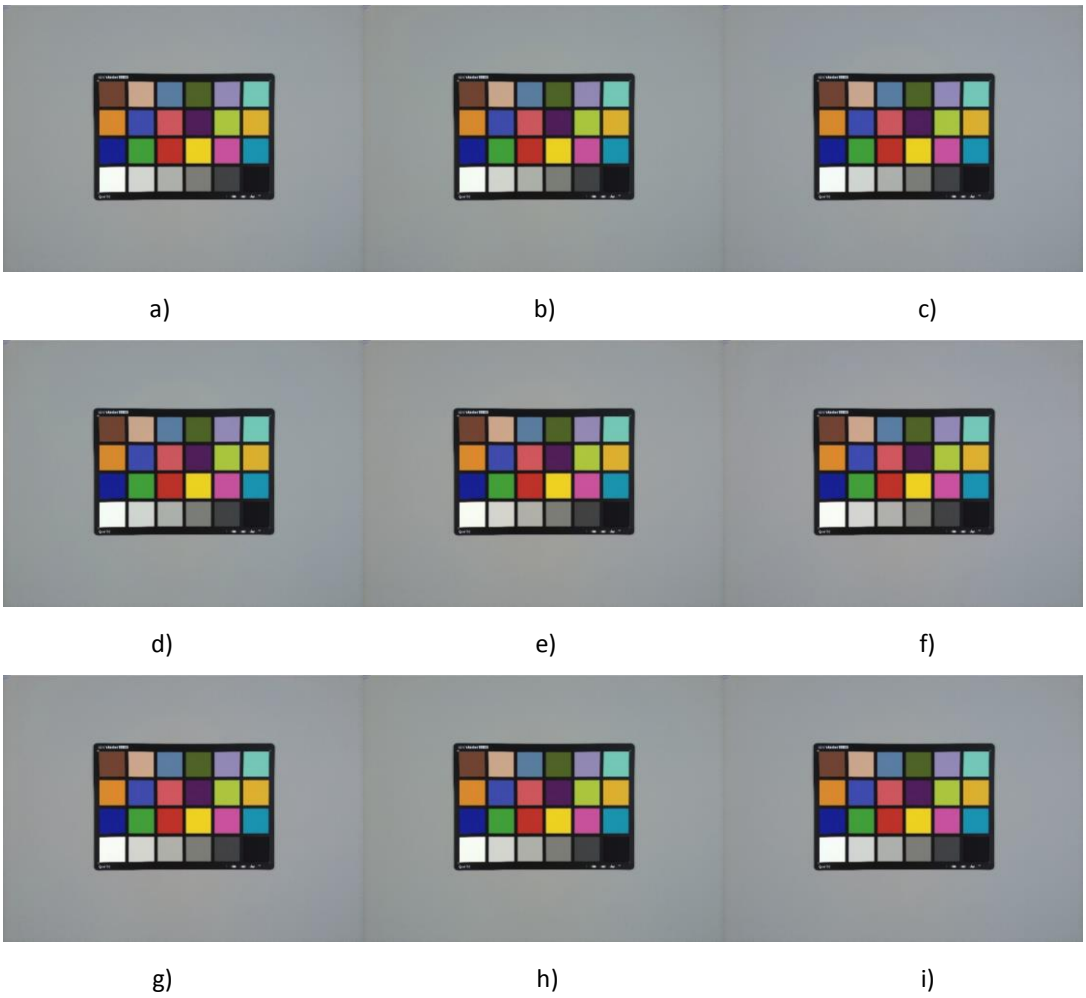


Figure 6.11: a) Original image, b) -1% offset in the red channel, c) +1% offset in the red channel, d) -1% offset in the blue channel, e) +1% offset in the blue channel, f)-1% offset in the red channel and +1% offset in the blue channel, g) +1% offset in the red channel and -1% offset in the blue channel, h) -1% offset in the red channel and -1% offset in the blue channel, i) +1% offset in the red channel and +1% offset in the blue channel

Figures 6.12 – 6.18 show examples of failures produced by the colour constancy methods proposed in literature when they are tested in challenging situations. Figures 6.19 and 6.20 show the ability of the proposed method to process the challenging scenes that made the other colour constancy methods fail. Note that the failure images in Figures 6.12 – 6.18 show images that are completely wrongly white balanced, with a significant colour cast. The aim in this research is not to assess the human preference in colour reproduction, thus, it is considered a fail case when the algorithms failed completely to white balance for different illuminants and different scenes.

Figure 6.12: Examples of failures of MaxRGB when they are tested in challenging situations.

Figure 6.13: Examples of failures of Weighted Grey Edge when they are tested in challenging situations.

Figure 6.14: Examples of failures of Revisited Bayesian when they are tested in challenging situations.

Figure 6.15: Examples of failures of Shades of Grey when they are tested in challenging situations.

Figure 6.16: Examples of failures of Grey-World when they are tested in challenging situations.

Figure 6.17: Examples of failures of Grey-edge 1st order when they are tested in challenging situations.

Figure 6.18: Examples of failures of Grey-edge 2nd order when they are tested in challenging situations.

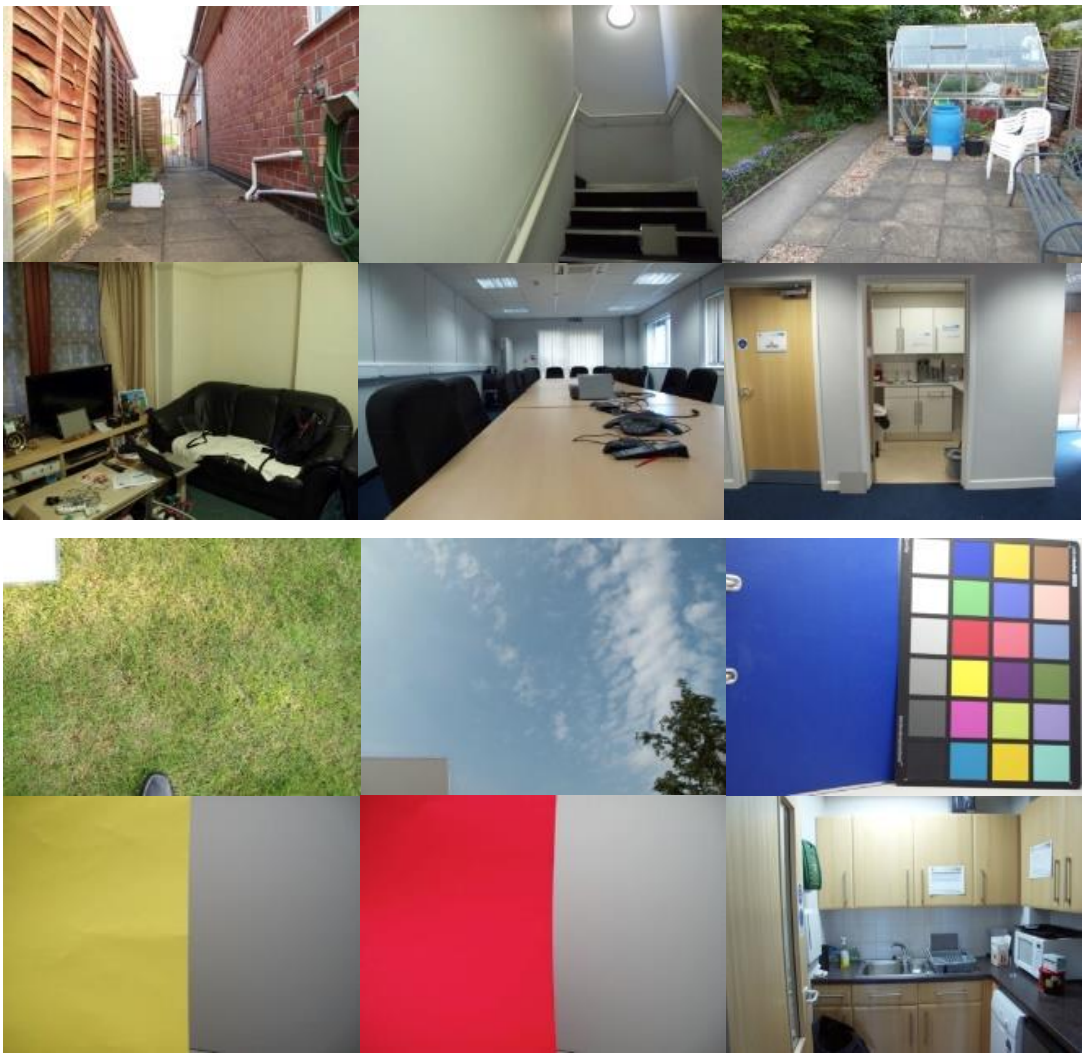


Figure 6.19: Examples of the ability of the proposed method to balance the challenging scenes that made the other colour constancy methods fail.

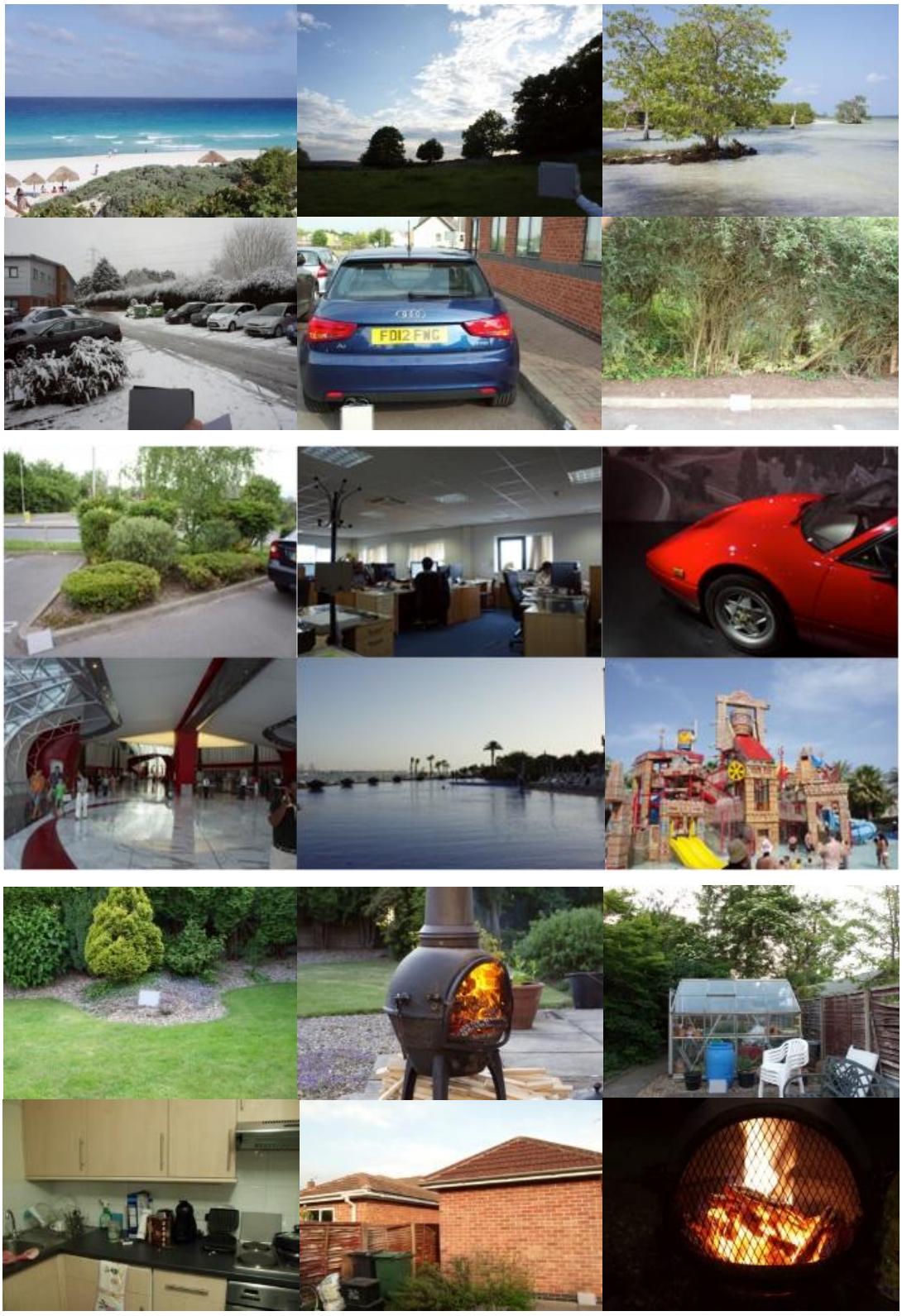
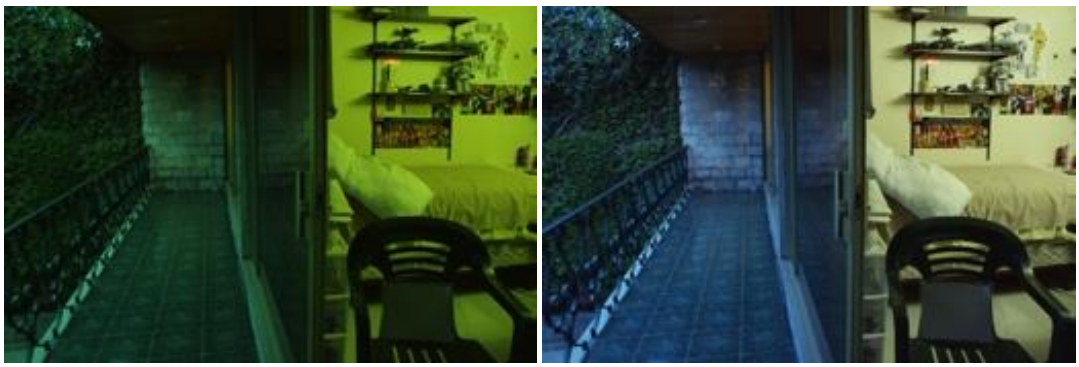


Figure 6.20: Examples of the ability of the proposed method to balance the challenging scenes that made the other colour constancy methods fail.

According to the results obtained, the proposed approach to colour constancy outperformed all other algorithms in terms of robustness, accuracy, consistency to correctly white balance all images, as can be seen from the standard deviation, mean, and maximum error measurements.

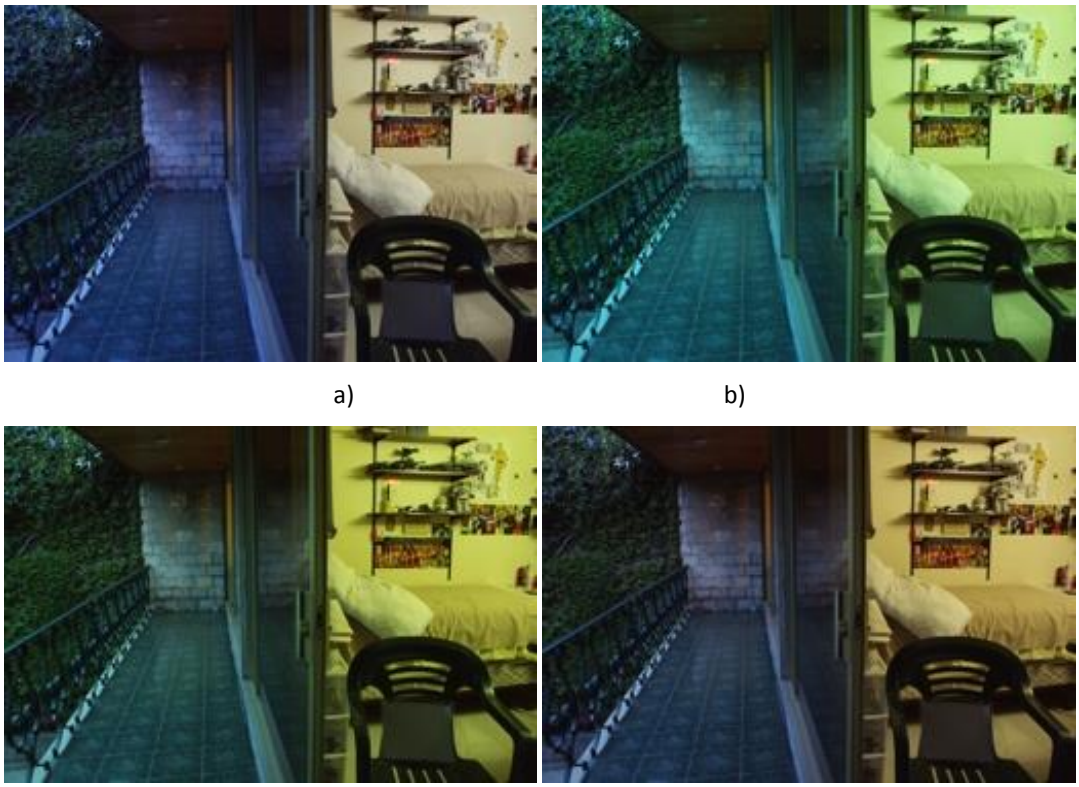
All the colour constancy methods evaluated had in common that none of them could accurately compensate in extreme, mixed lighting conditions. This is a known and perhaps a fundamental problem in all colour-balancing algorithms. In general there are few options a colour constancy algorithm could adopt in mixed lighting conditions: compensate towards one of the illuminants at the expense of an extreme colour cast in the other illuminant, or converge halfway the two illuminants white points and decide on a compromise. The above extreme scenes had an impact on the objective measurements since the grey points measured from the grey chart had a colour cast if the algorithm decided to compromise for the illuminants white points, or depended on the grey chart placement on the scene and on the illuminant that the colour constancy method converged. Figures 6.21 and 6.22 show results of all tested methods under this type of scene.

a) b)



c) d)

Figure 6.21: Example of mix lighting scene, indoor (4000K CWF bulb) and outdoor sunset in shadow (7000 K) output of all colour constancy methods tested. a) MaxRGB output. b) Weighted Grey Edge output. c) Revisited Bayesian output.



a) b)

c) d)

Figure 6.22: Example of mix lighting scene, indoor (4000K CWF bulb) and outdoor sunset in shadow (7000 K) output of all colour constancy methods tested. a) Grey-World output. b) Grey-edge 1st order output. c) Grey-edge 2nd order output. d) Proposed method output

6.4.3 Results of the proposed system

In Section 6.4, the ability of the proposed colour constancy method to act as an AWB method in an ISP and its superior performance when compared to other colour constancy methods previously proposed in literature were presented. This section demonstrates the ability of the proposed system to produce pleasant reproduction of colours based on Kruithof curve theory.

In order to reproduce pleasant colours based on Kruithof curve theory, the image has to be white balanced so that the neutral colours in the image are 100% neutral. Then, by applying one of the non-diagonal CCMs according to the colour temperature estimated in the AWB module, the spectral colours of the image are corrected. At this point the image colours should be as if captured in a photo lab controlled environment, where the ΔE_x, ΔE^*ab and

*ΔC** colour errors are low and consistent under any illuminant. Figure 6.23 shows examples of images processed by the simplified ISP, using the proposed AWB method and the 3-CCM blending method previously described in section 6.3.3. These images were captured in a light box at different colour temperatures ranging from 16000K to 2200K. It is possible to observe how colour constancy is achieved and how the ΔE_x, *ΔE*ab* and *ΔC** colour errors are very low and constant as summarised in table 6.2.

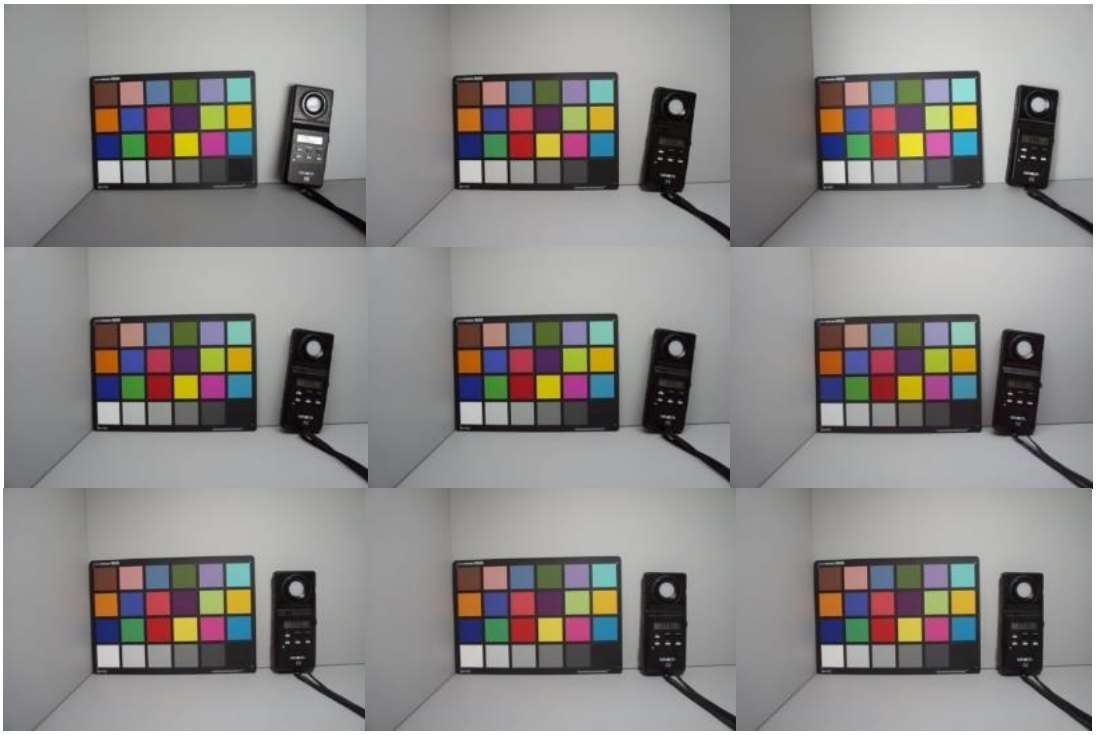


Figure 6.23: Example of images processed by the simplified ISP, using the proposed AWB method and the 3-CCM blending method proposed. The images were captured in a light box using light bulbs with colour temperatures ranging from 16000K (top-left) to 2200K (bottom-right).

	16000K	D75	D65	D50	D40	TL84	D30	2800K	2200K
Δex	0.02	0	0.01	0	0.1	0.02	0	0.003	0
ΔE*ab	7.33	6.44	4.33	4.55	5.79	7.66	6.33	4.11	6.48
ΔC*	6.77	4.66	3.54	4.23	6.77	9.21	8.32	5.66	8.44

Table 6.2: Proposed system performance summary ΔEx, ΔE*ab and ΔC* colour errors obtained from the images captured in a light box using light bulbs with colour temperature ranging from 16000K to 2200K.

In Figure 6.24, it is illustrated how the Kruithof curve theory is applied to the colour checker images from Figure 6.23 with the cooling-warming effect as described in Section 6.3.4. Moreover, Figure 6.24 clearly shows how colours at high colour temperatures tend to have a blue cast, and low colour temperatures a redder/orangery colour cast.

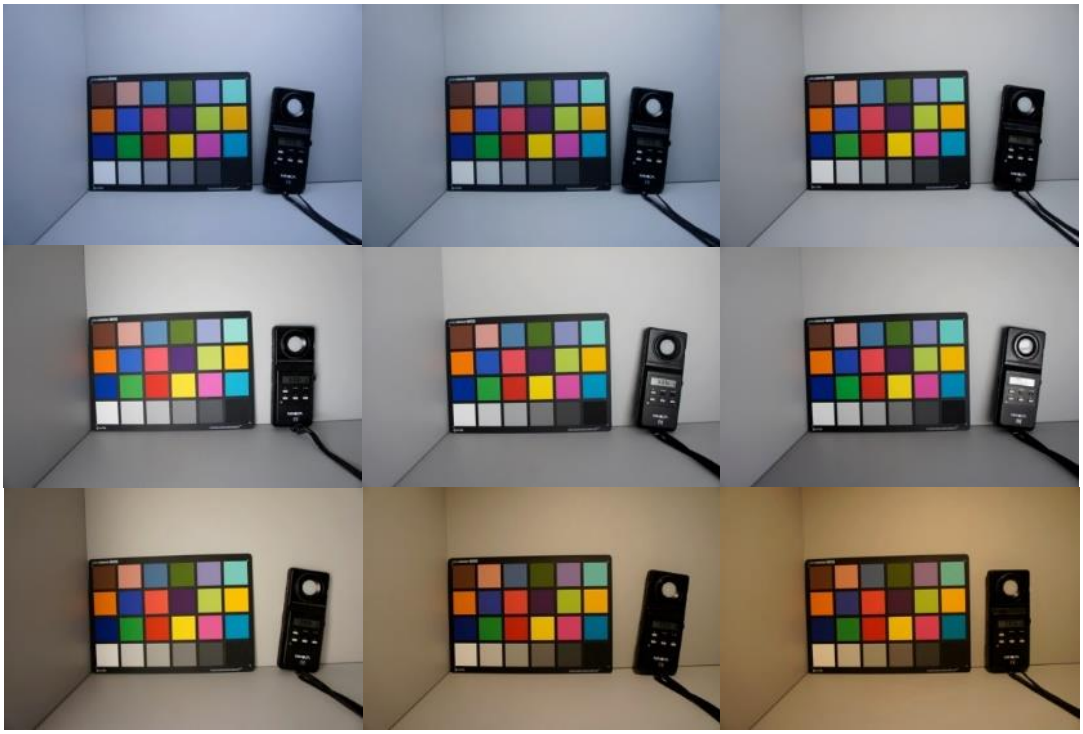


Figure 6.24: Example of images processed by the simplified ISP, using the proposed AWB method, the 3-CCM blending method, and a cooling-warming effect according to Kruithof curve theory. The images were captured in a light box using light bulbs with colour temperature ranging from 16000K (top-left) to 2200K (bottom-right).

Examples of the images processed using the colour management system proposed, are shown in Figures 6.25 and 6.26, where the cooling effect is applied to scenes with higher colour temperatures, no effect to the middle range of the colour temperatures scenes, and the warming effect to the low colour temperature scenes. In practice, the cooling-warming factor effect allows camera manufactures to have a colour management control tool to produce a more natural and pleasant image.

Figure 6.25: Example of images processed by the simplified ISP, using the proposed colour management system. Where the cooling effect is applied to scenes with higher colour temperatures (top-left), no effect to the middle range of the colour temperatures scenes.

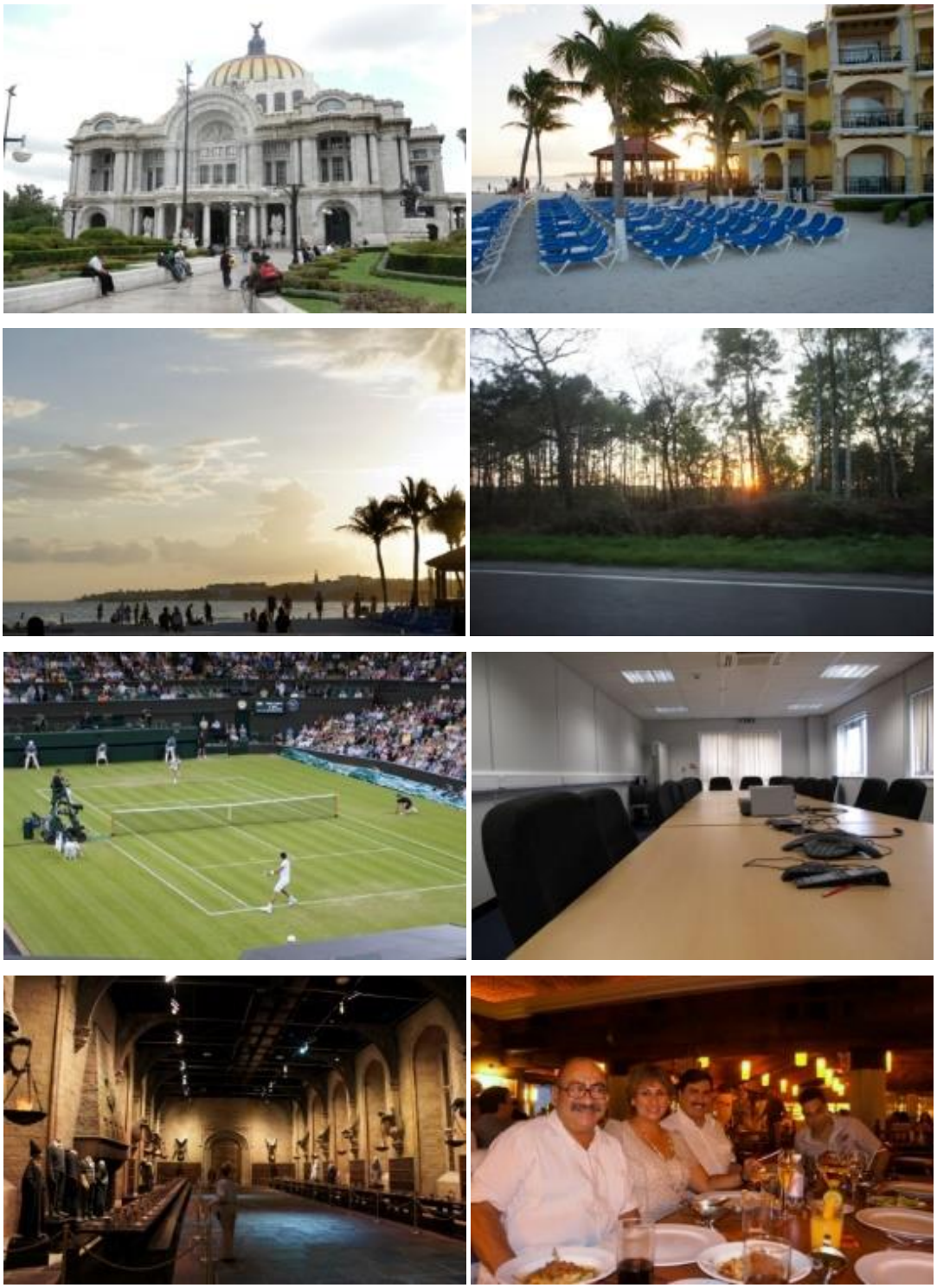


Figure 6.26: Example of images processed by the simplified ISP, using the proposed colour management system. Where no cooling-warming effect is applied to scenes in middle range of the colour temperatures (top-left), and the warming effect applied to the low colour temperature scenes (bottom-right).

6.5 Conclusions and Future Work

Colour accuracy is an important but at the same time ambiguous image quality factor. Accurate colours may be critical in technical or medical photography, but less important in consumer photography, where many camera users prefer images to be close to the real scene captured and with enhanced colour saturation particularly in "memory colours" such as foliage, sky, skin, etc. For this reason, in this research a colour management system that works inside an ISP and is able to compensate for different illuminants and at the same time reproduce pleasant colours was investigated and proposed.

The proposed colour management system within an ISP pipeline encompasses three stages. First, the colour temperature of the illuminant is estimated (AWB module); secondly, a colour correction/balancing matrix is applied (CCM); and thirdly a cooling-warming effect is applied to achieve camera end-users preferred colours, and therefore, a solution to the colour constancy problem. The proposed approach was designed in such a way that it is robust enough to cope with sensor-to-sensor variations, accurate enough to pass industry image quality standards for mobile devices, industrial, security, video conferencing, and automotive cameras. Further the proposed approach is flexible enough to be able to produce pleasant colours in images by following Kruithof curve theory, which aim is to reproduce and mimic the human eye's colour constancy.

In order to verify the performance of the proposed AWB module a large RAW image database was captured using two different DSLR commercial cameras, and an image sensor commonly used in smart phones (i.e. OV8835). With this data, the proposed AWB method was benchmarked against the most known colour constancy methods in literature. The results of the experiment showed that proposed method outperformed all others in terms of accuracy and robustness. I.e. The proposed system proved to be able to balance neutral and spectral colours under all types of illuminants, and was able to handle challenging scenes such as: monotonic colours, green grass colour versus CWF lighting, opposite colour objects' surface reflectance to the light colour temperature in the scene, sunset, low light conditions, absence of grey, noisy images, mixed lighting, cloudy scenes, and scenes with blue-sky

images. It is noted that all other colour constancy methods failed under these extreme circumstances.

The database of images tested can be accessed at:
https://drive.google.com/folderview?id=0B2xIm02tjf9kTTFZSTB4UXB2UDg&usp=sharing,

After analysing the results from the experiments performed, it was clearly shown that all colour constancy methods, including the proposed, cannot 100% compensate images in mixed lighting situations without compromising the final colour in the image. Thus future research in colour constancy should focus on solving this problem.

Chapter 7

Colour Constancy in HDR Imaging

This chapter proposes a solution to the colour constancy problem in HDR imaging. Generally, HDR imaging is composed of different illuminants that not only differ in their spectral characteristics, but also in their luminance levels. This observation is employed to segment images into areas lit by different illuminants and thus being able to achieve colour constancy, when different illuminants in the scene are captured by the camera system.

7.1 Introduction

HDR imaging has been attracting the attention of both photographers and researchers for a long time. However, it was not until recent times that the ever growing computational power, availability of microprocessors, and imaging technologies, have made HDR imaging a reality for consumer grade electronic devices.

HDR imaging is a technique for improved recordings of scene radiances, whereas colour constancy is referred to as a variation of a camera's automatic white balance algorithm [123]. On closer inspection, luminance levels limit the range of light that imaging sensors and human retinas can detect; in the case of HDR imaging, it works well as it preserves the details in the scene's spatial contrast. However, human colour constancy depends on spatial comparisons that synthesise appearances from all the scene areas, something that cannot be achieved with spatially invariant colour constancy algorithms [114].

Moreover, colour constancy is usually considered to be unrelated to HDR scene rendition. Unlike films or imaging sensors that have fixed spectral sensitivities, humans achieve colour constancy through variable spectral stimuli of colour reflectances. Many assume that retinal

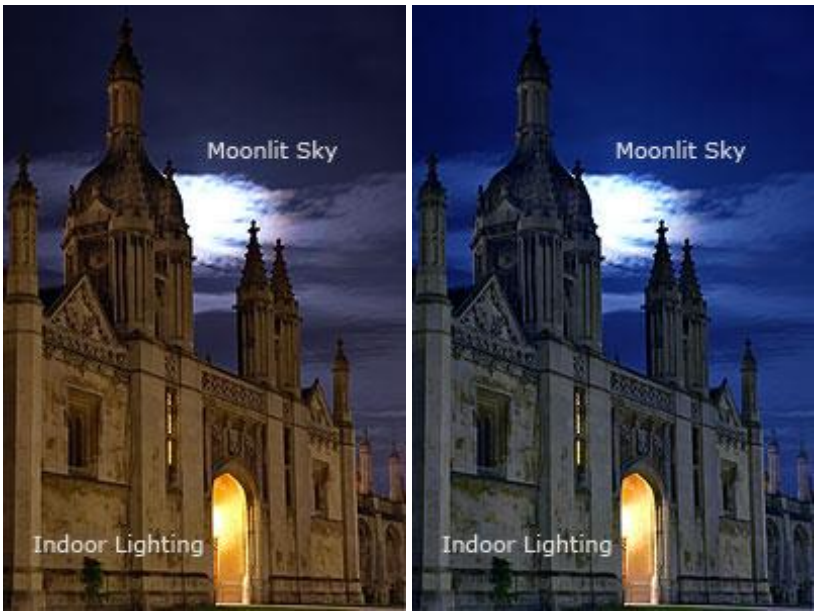
cones adapt to changes in illuminations, so that appearances correlate with reflectances [114].

As a difference to the human eye, most colour constancy algorithms assume that the spectral distribution of light is uniform across the image. Although this assumption works well in most cases and is widely used in commercial cameras, there exist common cases in which this assumption is violated. For example, a room illuminated by daylight coming from a window in combination with artificial light; scenes where some areas of the image are in-shadow and some others are out-of-shadow areas; images taken at ambient light levels with the use of a flash; and two different artificial light sources illuminating an indoor room.

HDR scenes lit by multiple illuminants, such as a room illuminated with artificial light and the sun shining through the window is an often-encountered real-life scenario that offers the opportunity for the HDR feature of a camera to show its relative advantages. However, since traditional colour constancy algorithms are spatially invariant and estimate a global white point of the scene, colour constancy algorithms are left with few options to adopt in mix lighting conditions: compensate towards one of the illuminants at the expense of an extreme colour cast in the other illuminant, or converge halfway the two illuminants' white points and decide on a compromise.

Colour constancy related research for HDR imaging can mainly be found in the psychophysical field [114, 115] or computer graphics related applications [29, 115, 120] and yet, there is only limited work contributing to the literature, which aims to resolve the colour constancy challenge when multi illuminants are present in a scene captured by a digital camera [119].

Figure 7.1 clearly shows an example of the colour constancy challenge when multiple illuminants are present in the scene [118].



a) b)

Figure 7.1: a) White balance for the moonlit sky; b) White balance for the church [118].

In digital imaging, illuminants with different colour temperatures are not only distinguished by their chromatic attributes or spectral power distribution, but also by the luminance levels as shown in Chapter 6, where the luminance of natural daylight can be a couple of magnitudes higher [9] than that of artificial lighting. Thus, it is intuitive to realise that if multiple illuminants were to be included in the same scene, parts illuminated by different light sources would occupy different areas of the dynamic range of the scene. This can be further interpreted as the data captured at different intensities, or exposures if exposure bracketing is used. This leads to the assumption that luminance information can be analysed in order to determine the likelihood of different illuminants in the scene.

The colour constancy problem when multiple illuminants are present in a scene can therefore be categorised as follows [121]: i) multiple illuminants at different luminance levels such as: day light from windows plus indoor lighting, in-shadow areas plus out-of-shadow areas, and camera flash in combination of ambient light; and ii) multiple illuminants with same illuminance levels such as: two different light sources with similar luminance levels present in an indoor room. This research focuses on the problem defined under (i) above.

In this Chapter, a spatially variant colour constancy algorithm within an ISP pipeline to solve the colour constancy problem in HDR imaging with multiple illuminants or reflectances in scene, is proposed.

For clarity of presentation, this chapter is organised as follows: Section 7.2 presents a detailed description of the proposed spatially variant colour constancy method within an ISP pipeline. In Section 7.3 the experimental setup and results obtained in this research are presented. Finally section 7.4 concludes by giving an insight to further work and possible improvements.

7.2 Proposed Spatially Variant Colour Constancy Method

As stated above, a spatially variant colour constancy algorithm is proposed within an ISP pipeline to solve the colour constancy problem in HDR imaging with multiple illuminants or reflectances. This approach incorporates the HDR solution presented in Chapter 6; a novel image segmentation process based on the image luminance levels, followed by colour temperature estimation of each the segmented areas in the scene by using the colour constancy method presented in Chapter 6, and a non-linear white balance correction method in order to compensate for the multiple illuminants in a scene.

Figure 7.2 shows a block diagram of the proposed spatially variant colour constancy algorithm and how it would fit within an ISP.

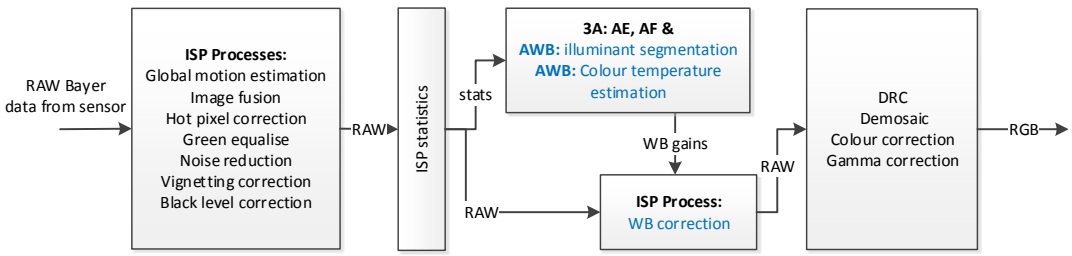


Figure 7.2: Diagram of the proposed colour constancy method inside an ISP pipeline

7.2.1 Illuminant segmentation

Different illuminants not only differ by their chromatic attributes or spectral power distribution, but also in their luminance levels especially when natural daylight is compared against artificial light sources [122].

The conditional probability model used in the Naïve Bayes Classifier in Chapter 6, which states: "*high colour temperatures are only present at high brightness levels with low red gains in the image, and low colour temperatures are present in low light conditions with high red gains in the image*", is also used in the segmentation process in order to form decision rules that allow calculating the probability of an illuminant being in the scene with correspondence to its luminance levels.

The first step of the image segmentation process is to calculate the cumulative histogram, cdf_I, of an image, I, with L number of intensity levels so that an intensity cut-off threshold, I_{cut}, can be estimated given an amount of black percentage pixels, $black_{prc}$, empirically selected. I_{cut} is used in conjunction with a second threshold, $dark_{th}$, to delimit the image signal below the sensor noise floor in order to create the segmentation mask, seg_I. This process is shown in Equations 7.1, 7.2, and 7.3, where $x = \{0,\ldots, \text{image rows}\}$, and $y = \{0, \ldots, \text{image columns}\}$.

$$cdf_I(i) = \sum_{j=0}^{i} p(I = i) = \frac{n}{n_i}, 0 \leq i < L \tag{7.1}$$

$$I_{cut} = \arg\min_{index} \left(\left| cdf_I - \left(black_{prc} \times cdf_I(L) \right) \right| \right) \tag{7.2}$$

$$seg_I(x, y) = \left(\left(I(x, y) \leq \frac{I_{cut}}{L} \right) \cup \left(I(x, y) > dark_{th} \right) \right) \tag{7.3}$$

Once the segmentation mask is calculated, the mean value, I_{avg}, of the pixels masked by seg_I is calculated using Equation 7.4. Then, by using Equation 7.5, the average value corresponding to the segmented dark pixels, I_{dk}, is estimated.

$$I_{avg}(x,y) = I(x,y) \times seg_I(x,y) \tag{7.4}$$

$$I_{dk} = \frac{\sum_{i=0}^{x}\sum_{j=0}^{y} I_{avg}(i,j)}{\sum_{i=0}^{x}\sum_{j=0}^{y} Seg_I(i,j)} \tag{7.5}$$

The colour temperature estimation method proposed in Chapter 6 uses the *rg*, and *bg* ratios calculated by the ISP statistics module. Thus, in this method, it is not necessary to segment the image itself in order to calculate multiple illuminants; it is only necessary to segment the *rg*, *bg* ratios so that two statistical sets can be analysed to determine the probability of multiple illuminants in the scene.

The ISP statistics module is modified to take as an input the rgbg coefficient, $rgbg_{coeff}$, calculated using Equation 7.6 and the RAW Bayer image, in order to generate two sets of statistics; one for the input image denoted by *Z* and another set for the input image times the rgbg coefficient denoted by *Z',* as shown in Equations 7.7 and 7.8.

$$rgbg_{coeff} = \frac{black_{prc}}{I_{dk}} \tag{7.6}$$

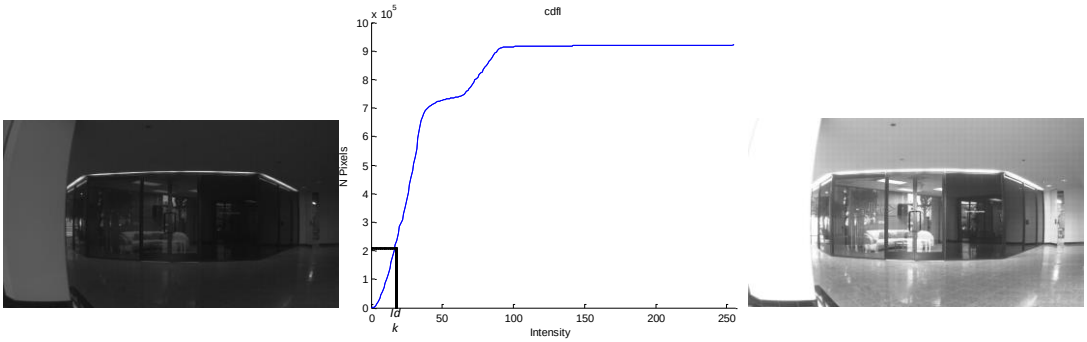
$$Z_{(1,\ldots,n)} = ISP_{stats}(RAW\ image\ ) \tag{7.7}$$

$$Z'_{(1,\ldots,n)} = ISP_{stats}\left(RAW\ image\ \times\ rgbg_{coeff}\right) \tag{7.8}$$

Where ISP_{stats} represents the ISP statistic module, which gathers the *rg* and *bg* ratios of each of the *n* number of zones.

Figure 7.3 illustrates an example of the segmentation process performed on an image with multiple illuminants in the scene. Figure 7.3(a) shows the input RAW Bayer image. Figure 7.3(b) shows *cfdI* and the estimated *Idk* threshold calculated from the input image, and Figure 7.3(c) shows a representation of the input image multiplied by $rgbg_{coeff}$. This example clearly shows how the ISP statistics module would calculate the two sets of statistics *Z* and

Z', one set for the input image and another set for the input image with the segmentation coefficient $rgbg_{coeff}$ applied.



a) b) C)

Figure 7.3: Illustration of the segmentation process.

7.2.2 Colour temperature estimation

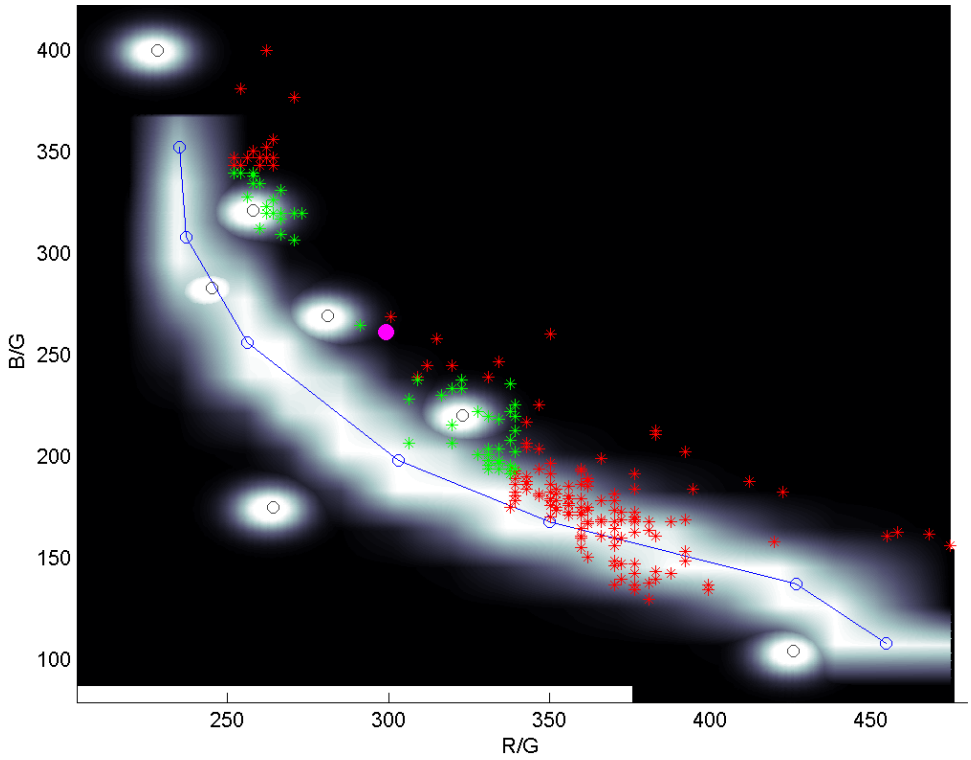
The colour temperature estimation for multiple illuminants is performed through the method proposed in Chapter 6, equations 7.24 - 7.28 with a slight modification so that it uses two sets of ISP statistics, *Z* and *Z'*, to calculate the probability of multiple white points in the scene. This process is represented by the functions cte*(Z)* and cte*(Z')* for compactness of equations 7.24 - 7.28, as shown in Equation 7.9 and 7.10. cte*(Z)* and cte*(Z')* generate the *rg* and *bg* gain coefficients denoted by the vectors *e* and *e',* which are later used in the colour correction stage.

$$e = \text{cte}(Z), \text{where } e = [rg, 1, bg]^T \tag{7.9}$$

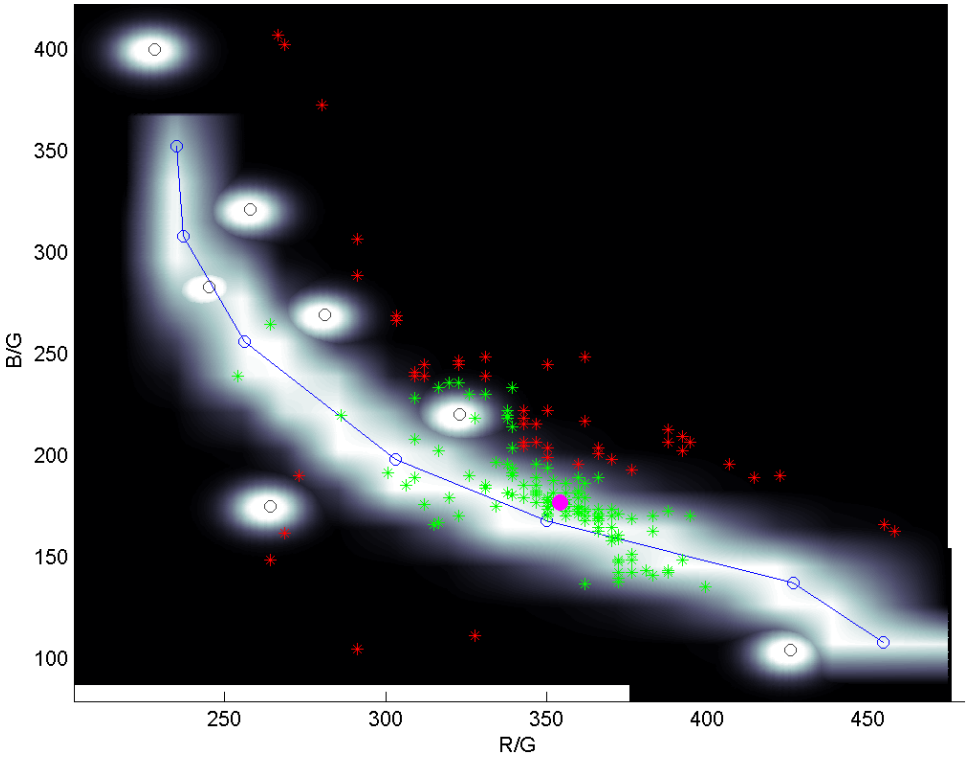
$$e' = \text{cte}(Z'), \text{where } e' = [rg', 1, bg']^T \tag{7.10}$$

Plots of the colour temperature estimation process are shown in Figure 7.4, which uses the *Z* and *Z'* ISP statistics generated after the segmentation process shown in Figure 7.3. Figure 7.4(a) is the plot generated by cte*(Z),* and Figure 7.4(b) the plot of cte*(Z')*. The plots of the ISP statistics, *Z* and *Z'*, shown in Figure 7.4 are coloured as follows: green dots are the

statistics considered by the *cte* to calculate the colour temperature; red dots are the statistics not considered by *cte* for calculations of colour temperature; blue dots show the characterised Planckian locus of the sensor, and the magenta coloured dot is the convergence point of *cte,* the illuminant detected by the algorithm.



a)



b)

Figure 7.4: Plots of *cte(Z)*, and *cte(Z')* a) Output colour temperature of *cte(Z)*= 5322K, b) Output colour temperature of *cte(Z')* = 3041K.

Similar to the white balance correction process proposed in Chapter 6, the image is colour corrected by normalising *e*, and *e'* , using equation 7.3, however, in this approach the illuminant vectors *e*, and *e'* are treated differently when colour balancing the image. This process is described below.

7.2.3 Colour correction

In order to colour balance the image, the illuminant vector *e* is applied to the image using equation 7.1, followed by a generation of a piecewise linear equations (PWL) system based on intensity levels that allows correcting for the illuminant in vector *e'*.

The generation of the piecewise linear equations, PWL_R and PWL_B is performed through equations 7.11-7.21. In this process the intensity levels considered are defined by *L*, where $L = \{0, \dots, 255\}$ intensity levels.

The first step of generating the *PWL* system of equations is to estimate the colour gain ratios from *e* to *e'* so that the image can be colour balanced according to its intensity levels and white points estimated (Equation 7.11 and 7.12).

$$rg = e(1), rg' = e'(1), bg = e(3), bg' = e'(3) \tag{7.11}$$

$$rg_{ratio} = \frac{rg}{rg'}, bg_{ratio} = \frac{bg}{bg'} \tag{7.12}$$

Where *rg, bg* and *rg', bg'* represent the gains needed for colour balancing the image and rg_{ratio} and bg_{ratio} are the ratios *rg* to *rg'* and *bg* to *bg'*. The ratios of the two white points and the $rgbg_{coeff}$ are then used for denoting the knee points of the *PWL* system using equations 7.13 and 7.14:

$$PWL_{rg}(i) = \begin{cases} i \times rg_{ratio}, i \leq (1/rgbg_{coeff}) \\ i, \qquad\qquad i > (1/rgbg_{coeff}) \end{cases} \tag{7.13}$$

$$PWL_{bg}(i) = \begin{cases} i \times bg_{ratio}, i \leq (1/rgbg_{coeff}) \\ i, \qquad\qquad i > (1/rgbg_{coeff}) \end{cases} \tag{7.14}$$

PWL_{rg} and PWL_{bg}, are the systems of equations of the *rg* and *bg* ratios based on the intensity level *i*, where $i \in L$. In order to have a continuous function so that the blending of the two white points is smooth, a blending area with coordinates $x1$, $x2$, $y_{rg}1$, $y_{rg}2$, $y_{bg}1$, $y_{bg}2$ is determined using equations 7.15-7.17:

$$x1 = \left(\frac{1}{rgbg_{coeff}}\right) \times 0.75, \qquad x2 = \left(\frac{1}{rgbg_{coeff}}\right) \times 1.25 \tag{7.15}$$

$$y_{rg}1 = PWL_{rg}(x1), \qquad y_{rg}2 = PWL_{rg}(x2) \tag{7.16}$$

$$y_{bg}1 = PWL_{bg}(x1), \qquad y_{bg}2 = PWL_{bg}(x2) \tag{7.17}$$

Once $x1$, $x2$, $y_{rg}1$, $y_{rg}2$, $y_{bg}1$, $y_{bg}2$ are known, PWL_{rg} and PWL_{bg} are recalculated using equations 7.18 and 7.19 represented by $SPWL_{rg}$ and $SPWL_{bg}$. The blending areas help removing extreme knee points in the *PWL*. However, in order to make sure the *PWLs* are continuous functions, a moving average filter is applied to $SPWL_{rg}$ and $SPWL_{bg}$, as in equations 7.20 and 7.21, to obtain the final PWL_R and PWL_B, where *M* is the number of elements of the moving average filter.

$$SPWL_{rg}(PWL_{rg}(i)) = \begin{cases} \left(\frac{y_{rg}2 - y_{rg}1}{x2 - x1}\right) \times (i - x1) + y_{rg}1, x1 \geq i \leq x2 \\ i, \qquad\qquad otherwise \end{cases} \tag{7.18}$$

$$SPWL_{bg}(PWL_{bg}(i)) = \begin{cases} \left(\frac{y_{rg}2 - y_{rg}1}{x2 - x1}\right) \times (i - x1) + y_{bg}1, x1 \geq i \leq x2 \\ i, \qquad\qquad otherwise \end{cases} \tag{7.19}$$

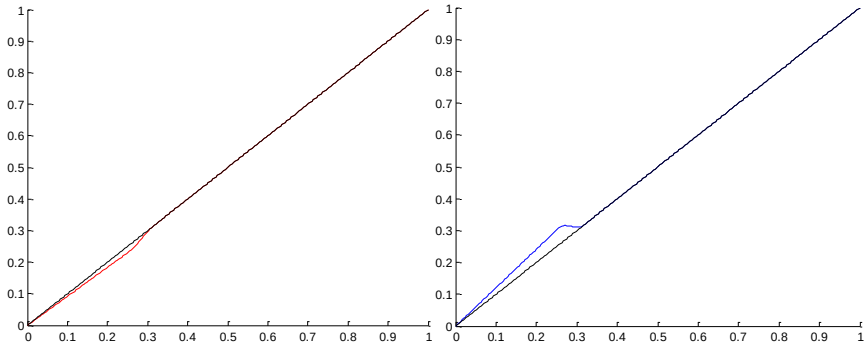
$$PWL_R(i) = \frac{1}{M} \sum_{j=0}^{M-1} SPWL_{rg}\,(i + j) \tag{7.20}$$

$$PWL_B(i) = \frac{1}{M}\sum_{j=0}^{M-1} SPWL_{bg}\,(i+j) \tag{7.21}$$

To finalise the colour correction process, PWL_R and PWL_B to the red and blue colour channels of the already colour corrected image by the illuminant in vector *e* using equation 7.1.

In order to illustrate the colour correction process, the examples shown in Figure 7.3 and 7.4 are used as a reference to calculate the *PWLs* and output images displayed in Figure 7.5. Figure 7.5(a) shows the colour corrected image using the illuminant vector *e*, which represents the output image from the colour constancy proposed in Chapter 6. In this case, the multiple illuminants in the scene can be observed: the pillar on the left of the scene is illuminated by daylight, and the rest of the scene illuminated by artificial light. Figure 7.5(b) shows the same scene, but compensated toward the daylight illuminant, Figure 7.5(c) and (d) show the plots of PWL_R and PWL_B of this specific example, and Figure 7.5(d) shows the final image with the two illuminants properly estimated and corrected in the scene.

a) b)



c) d)

e)

Figure 7.5: example of the colour correction process using the illuminant vectors *e* and *e'*. a) The colour corrected image using the illuminant vector *e*. Manually colour corrected toward the daylight illuminant. c) and d) show the plots of PWL_R and PWL_B calculated using the illuminant vector *e'*. e) Output image of the proposed colour correction process.

7.3 Experimental Results

Experimental results are presented in this section in order to evaluate the performance of the proposed spatially variant colour constancy algorithm. Results of the experiments presented in Chapter 6 demostrated that the colour temperature estimation algorithm proposed in section 6.3.2, which is also used in this research, outperformed all previously proposed state-of-the-art colour constancy methods that could be used as the AWB module of an ISP. For this reason, the experiments performed within the context of this research do not focus on the performance comparison of the proposed algorithm to other colour constancy methods, but to test the capability of proposed method to be able to compensate high dynamic range scenes lit by multiple illuminants without introducing colour artefacts, or AWB failures due to the spatially variant colour temperature estimation method added to the colour constancy solution proposed in Chapter 6.

The database gathered for the experiments presented in Chapter 6, was also used in this experiment in order to verify the performance of the proposed method in terms of robustness, repeatability, and accuracy to perfectly colour balance images captured by different imaging sensors. Moreover, for the purpose of evaluating the proposed spatially variant colour constancy algorithm, the database collected in Chapter 6 was extended so that it contained HDR scenes lit with multiple illuminants captured with OV8835, Sony Nex-5, Canon G10, and AR0132AT [124] imaging sensors.

For each of the four imaging sensors used, a camera characterisation was performed in order to know the black level offset of the Bayer sensor, the Planckian locus, and colour response. The black level offset was obtained by capturing a black RAW image at an analogue gain of 1 and estimated by calculating the mean value of each of the four-colour channels. The Planckian locus curves were characterised by capturing a large set of illuminants ranging from 1800 kelvin (K) to 17000K, including cool white fluorescent (CWF) lighting under a control environment using professional equipment such as: photo lab light box, CIE standards colour temperature light bulbs, and a 24-ColourChecker Macbeth chart for colour measurements, and the CCMs were obtained as described in section 6.3.3. For completeness of the ISP simulator, the type of Gamma correction used was sRGB, and a bilinear

demosaicing algorithm was used to convert a RAW Bayer image to a RGB. The use of the HDR feature was only used when the scene required the multi-exposure technique proposed in Chapter 5 in order to extend the dynamic range of the image captured.

Figure 7.6 shows the ISP simulation model diagram used for the above experiments.

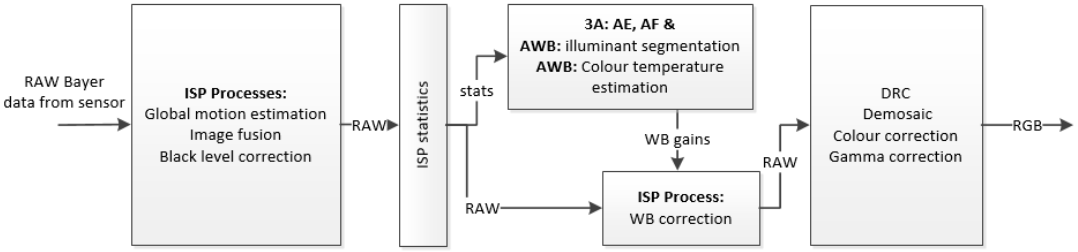


Figure 7.6: ISP pipeline used in experiments

A summary of the objective results obtained by the experiments performed on the database collected in Chapter 6 is shown in table 7.1. These results show that the spatially variant feature added to the colour temperature estimation proposed in Chapter 6 did not have a negative impact on the robustness, repeatability, and accuracy of the method to perfectly colour balance images captured by different imaging sensors under different scenes and illuminants.

	R:G		**B:G**		**ΔEx**	
Spatial invariant colour temperature estimation as in Chapter 6	Mean	1.00	Mean	1.00	Mean	0.13
	Std	0.10	Std	0.15	Std	0.14
	Max	1.73	Max	1.75	Max	0.90
Spatial variant colour temperature estimation	Mean	1.00	Mean	1.00	Mean	0.126
	Std	0.10	Std	0.14	Std	0.13
	Max	1.43	Max	1.75	Max	0.870

Table 7.1: Objective measurements obtained from the spatially invariant colour constancy method proposed in Chapter 6, and the proposed spatially variant colour temperature.

In order to show the ability of the proposed method to compensate for multiple illuminants at different luminance levels, Figure 7.7 presents a set of examples where the scenes captured represent the scenario where in-shadow areas plus out-of-shadow areas in the image have different colour response. Figure 7.8 presents another set of examples where the scenes captured represent the scenario where scenes are illuminated by daylight and artificial indoor lighting.

Figure 7.7 – 7.13(a) shows images processed by the colour estimation method proposed in Chapter 6, and Figure 7.7 – 7.13(b) shows output images of the spatially variant colour constancy algorithm proposed in this chapter. In Figure 7.7 – 7.13(a) the in-shadow areas have a visible blue colour cast, even though the out-of-shadow areas are well colour balanced. In contrast to the images in Figure 7.7 – 7.13(a), images presented in Figure 7.7 – 7.13(b) clearly show both in-shadow and out-of-shadow areas perfectly colour balanced without visible colour casts due to the multiple reflectances of the illuminant in the scene. The result of the proposed method improves the subjective image quality and makes high dynamic range images look natural and well colour balanced.

a)

b)

Figure 7.7: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)

b)

Figure 7.8: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)

b)

Figure 7.9: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)

b)

Figure 7.10: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)

b)

Figure 7.11: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)

b)

Figure 7.12: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)

b)

Figure 7.13: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

Images shown in Figure 7.14 – 7.19 represent the worst-case scenario of a colour constancy algorithm becasue the scenes are lit by opposite colour temperature illuminants with different spectral response. That is, some areas of the scenes are illuminated by daylight,

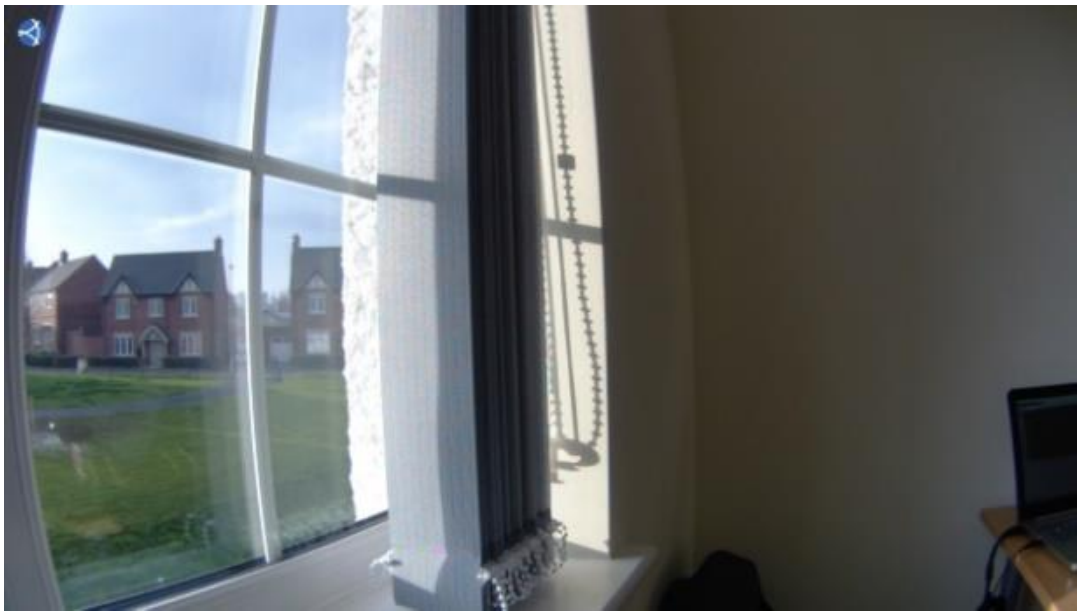
which has a colour temperature of 6500K with a blue colour cast, and other areas of the scenes are lit by indoor lighting, which has a colour temperature of 2800K with a red-yellow colour cast. Figure 7.14 – 7.19(a) shows images processed by the colour estimation method proposed in Chapter 6, and Figure 7.14 – 7.19(b) shows the output of the proposed spatially variant colour constancy algorithm. Similarly to the images shown in Figure 7.7 – 7.13, all images shown under (a) have visible colour casts due to the combination of multiple illuminants in the scene, whereas images shown under (b) are well colour balanced in the presence of multiple illuminants in the scene.

a)

b)

Figure 7.14: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)



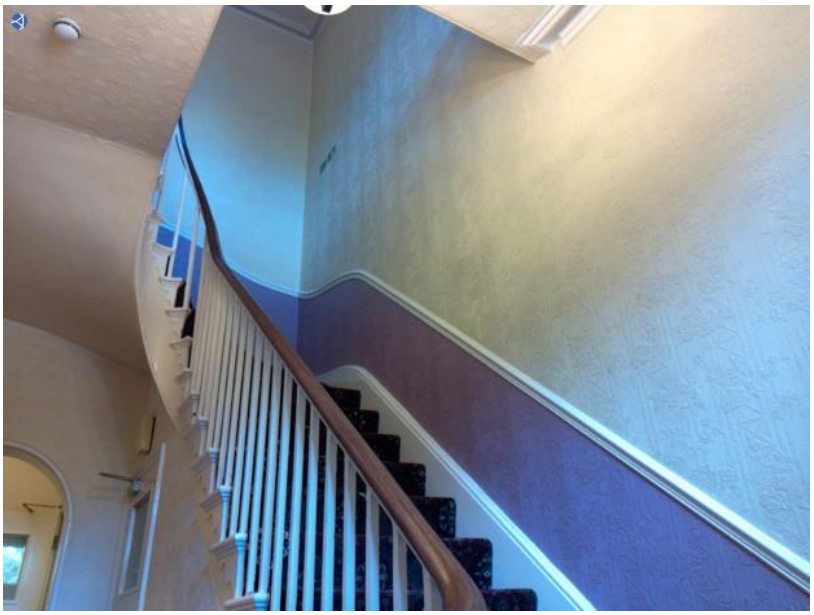
b)

Figure 7.15: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

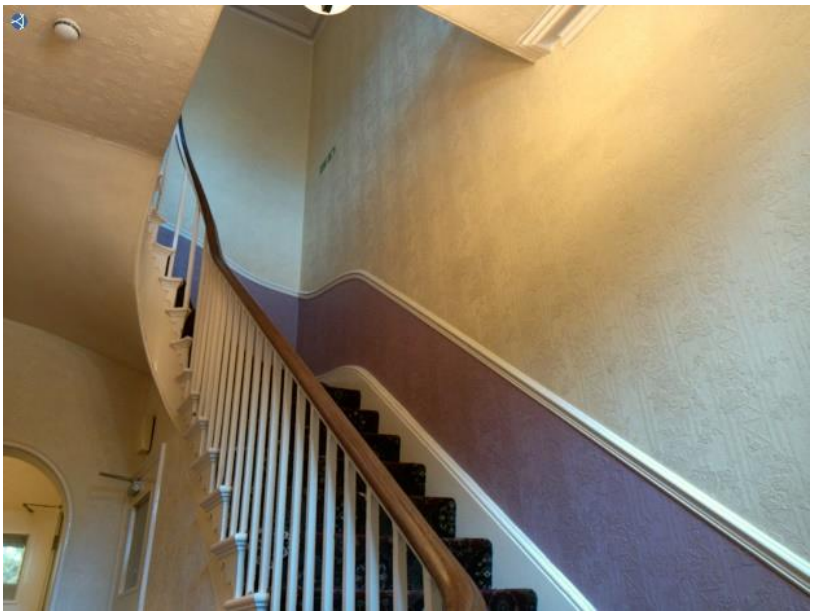
a)

b)

Figure 7.16: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

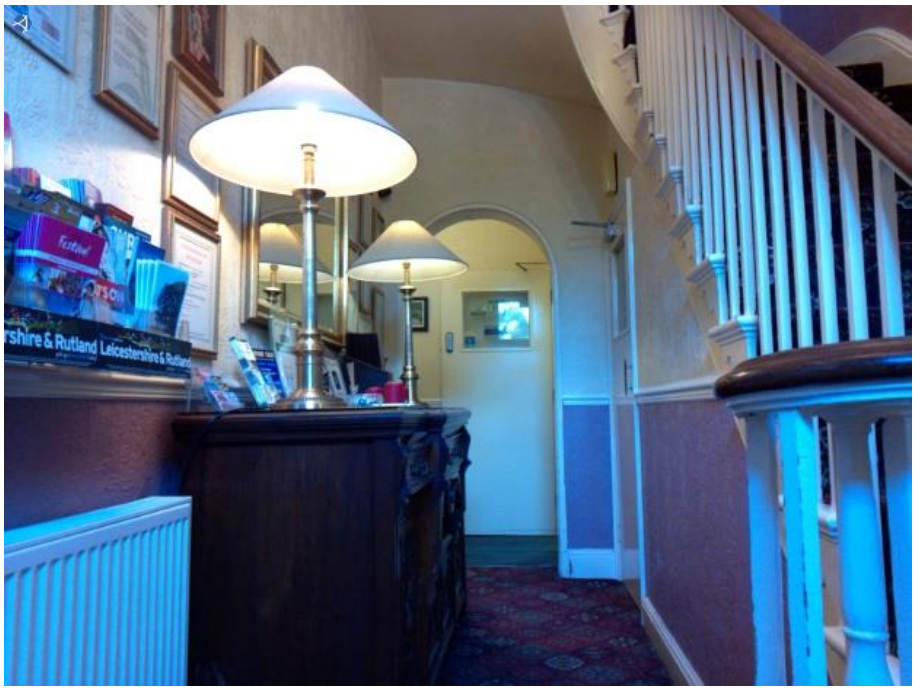


a)



b)

Figure 7.17: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.



a)

b)

Figure 7.18: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

a)

b)

Figure 7.19: Examples of a scene lit by multiple illuminants. a) Image output from the proposed colour estimation process in Chapter 6. b) Output image of the proposed spatially variant colour constancy algorithm.

The limitation of the proposed algorithm is to colour balance areas of the image where multiple illuminants are mixed in the same region, and to fully colour balance transition areas of opposite colour temperatures illuminating the scene. Nonetheless this scenario is significantly improved in comparison to the spatially invariant colour constancy methods benchmarked in Chapter 6.

7.4 Conclusions

Different illuminants not only differ by their chromatic attributes or spectral power distribution, but also in their luminance levels. For this reason, a spatially variant colour constancy algorithm within an ISP pipeline to solve the colour constancy problem in HDR imaging with multiple illuminants or reflectances in the scene was proposed. The approach incorporates the HDR solution presented in Chapter 5; a novel image segmentation process based on the image luminance levels, followed by colour temperature estimation of each the segmented areas in the scene by using the colour constancy method presented in Chapter 6, and a non-linear white balance correction method in order to compensate for the multiple illuminants in a scene.

The proposed algorithm proved to be efficient in different lighting conditions and scenes and was proven to work well with different sensors. Moreover, a summary of the objective results performed in the database collected in Chapter 6 showed that the spatially variance feature added to the colour temperature estimation proposed in Chapter 6 did not have a negative impact on the robustness, repeatability, and accuracy of the method to colour balance images captured by different imaging sensors under different scenes and illuminants. The result of the proposed method improves the subjective image quality and makes high dynamic range images look natural and well colour balanced.

After further analysing the results from the experiments performed, it was shown that the limitation of the proposed approach is its inability to perfectly colour balance areas of the scene lit by multiple illuminants with the same luminance levels. Thus future research should

focus on proposing colour correction methods that could correct for multi illuminant spectrum crosstalk illuminating the same area of a scene.

Chapter 8

Conclusions and Future Work

The main motivation to the research presented in this dissertation came from the observation that computer graphics, computer vision, and related digital imaging applications continue to demand more realistic and accurate imaging reproduction with high dynamic range scenes and colour constancy.

For this reason, the research carried out and presented in this dissertation aimed to improve the overall image quality of images captured by digital cameras by extending the dynamic range, and achieving colour constancy. These contributory topics were presented in Chapters 5, 6, and 7. Moreover, since image quality in digital imaging completely depends on the lens, imaging sensor and on the ISP's ability to produce high quality images, the novel image processing algorithms presented in this dissertation were intended so that all of them can be directly incorporated within an ISP to improve the overall image quality of a camera system.

8.1 Conclusion

The background theory and literature review presented in Chapters 2 and 3, introduced fundamental concepts and theories from different areas of mathematics, computing and engineering related to the human visual system, digital video and still photography. Moreover, these chapters provided an insight to the existing HDR imaging, and colour constancy solutions proposed in the literature.

In summary, five novel image-processing algorithms were proposed in this dissertation. These algorithms focused on two different image quality related areas: i) HDR imaging, Chapter 5 and ii) colour constancy, Chapters 6 and 7.

I) HDR imaging:

In Chapter 5, a new algorithm capable of compensating for camera shake and capable of producing HDR images and video content was proposed. This algorithm enabled users to create HDR content with a SDR camera without the aid of stabilising devices such as tripods. As a difference to HDR solutions proposed in the literature, which in their majority are only suitable for still HDR imaging capture, the proposed approach can produce video HDR content for real time imaging applications. This fact is very important since video applications such as CCTV or automotive cameras can highly benefit from HDR content reproduction.

The HDR video solution presented in Chapter 5, is a multi-exposure image fusion algorithm suitable for practical implementation in hardware. The proposed approach was implemented as a pre-processing technique that works in RAW Bayer domain and can be incorporated within an ISP as described in Chapter 4. This novel method performed global motion estimation and correction to compensate for motion due to camera shake, followed by the image fusion process, where in its operational logic, compensated for objects moving within the scene. The combination of these two steps, global motion compensation in conjunction with the motion aware fusing technique, allowed removing motion artefacts and ghosting due to camera shake and moving objects in the scene. The proposed algorithm proved to be robust and efficient in different lighting conditions, scenes, and different imaging sensors allowing digital camera devices to be able to create HDR content in real time.

III) Colour constancy

A colour management system that works inside an ISP and is able to compensate for different illuminants and at the same time reproduce pleasant colours was proposed in Chapter 6. The proposed colour management system within an ISP pipeline encompassed three stages. First, the colour temperature of the illuminant was estimated (AWB module); secondly, a colour correction/balancing matrix was applied (CCM); and thirdly a cooling-warming effect was applied to achieve camera end-users preferred colours, and therefore, a solution to the colour constancy problem.

The proposed approach was designed in such a way that it was robust enough to cope with sensor-to-sensor variations, accurate enough to pass industry image quality standards for mobile devices, industrial, security, video conferencing, and automotive cameras. Furthermore, the proposed approach was flexible enough to be able to produce pleasant colours by following Kruithof theory.

In order to verify the performance of the proposed AWB module, a large database consisting of RAW Bayer images was captured using two different DSLR commercial cameras, and an image sensor commonly used in smart phones (OV8835). With this database, which has been made available for future research, the proposed AWB method was benchmarked against the most known colour constancy methods in the literature. The results of the experiment showed that proposed method outperformed all others in terms of accuracy and robustness.

The database of RAW Bayer images has been published and can be accessed at:
https://drive.google.com/folderview?id=0B2xIm02tjf9kTTFZSTB4UXB2UDg&usp=sharing

The proposed system proved to be able to balance neutral and spectral colours under all types of illuminants, and was able to handle challenging scenes such as: monotonic colours, green grass colour versus CWF lighting, opposite colour objects' surface reflectance to the light colour temperature in the scene, sunset, low light conditions, absence of grey, noisy images, mixed lighting, cloudy scenes, and scenes with blue-sky images. It is noted that all other colour constancy methods failed under these extreme circumstances.

In Chapter 7, an extension of the proposed colour constancy method presented in Chapter 6 was proposed to solve the problem of colour constancy in high dynamic range imaging. This approach incorporated the HDR solution presented in Chapter 5, a novel image segmentation process based on the image luminance levels, followed by colour temperature estimation of each the segmented area in the scene by using the colour constancy method presented in Chapter 6, and a novel non-linear white balance correction method in order to compensate for the multiple illuminants in the scene. The proposed algorithm proved to be efficient in different lighting conditions, scenes, and was proven to work well with different sensors. The objective results obtained through the experiments performed showed that

the spatially variant feature added to the colour temperature estimation method proposed in Chapter 6 did not have a negative impact on the robustness, repeatability, and accuracy of the method to achieve colour constancy under different scenes and illuminants. Moreover, the results of the unique high dynamic range and spatially variant colour constancy algorithm enabled digital camera devices to improve the final image quality within an ISP pipeline implemented in hardware.

8.2 Future Work

In the field of high dynamic range imaging, a pre-processing HDR solution for video HDR creation was presented in Chapter 5. After further analysing the results from the experiments performed, it was found that the limitation of the proposed approaches is their inability to compensate for motion at exposure ratios bigger than 1:16. This problem can be solved if the image fusion system was extended to three-exposure image fusion instead of a two-exposure image fusion exercise. In a three-exposure image fusion system, exposure ratio between short to medium and medium to long exposures could be set to 1:16 and 1:16, and therefore extend the dynamic range of a 12 bit digital camera to a theoretical 120dB. Unfortunately, imaging sensors capable of capturing RAW Bayer data are not widely available and thus, it was not possible to test this option. Another limitation found in this research, was the ability to compensate for large rotation displacements, within the set of multi-exposure images. Thus future research should focus on: 1) proposing new HDR methods that include efficient, hardware friendly motion estimation algorithms capable of compensating for rotational displacements in order to create HDR contend in real time; and 2) improving the dynamic range imaging sensor technologies to avoid using multi-exposure techniques to extend the maximum dynamic range of a camera system.

In order to achieve colour constancy in digital imaging, two solutions were presented in this thesis. Further analysis of the results from the experiments performed in Chapters 6 and 7 showed that the limitation of the proposed colour constancy approach is its inability to perfectly colour balance areas of a scene lit by multiple illuminants with the same luminance levels, which up to date, any colour constancy method proposed in the literature or industry

have not been able to solve. Thus future research should focus on proposing a better imaging sensor spectral response, and colour correction methods that could correct for multi-illuminant spectrum crosstalk illuminating the same area of a scene.

References

[1] Gamal, Abbas El, and Helmy Eltoukhy. “Cmos Image Sensors.” Circuits and Devices Magazine, IEEE 21.3 (2005): 6-20.

[2] Theuwissen, Albert J. Solid-State Imaging With Charge-Coupled Devices. Vol. 1 Springer Science & Business Media, 2006.

[3] Sensors Explained. http://www.whatdigitalcamera.com/technical-guides/technology-guides/sensors-explained-11457

[4] Digital Camera Sensors. http://www.cambridgeincolour.com/tutorials/camera-sensors.htm

[5] Ramanath, Rajeev, Wesley E Snyder et al. “Color Image Processing Pipeline.” Signal Processing Magazine, IEEE 22.1 (2005): 34-43.

[6] Nakamura, Junichi. Image Sensors and Signal Processing for Digital Still Cameras. CRC press, 2005.

[7] Emil Martinec. “Noise, Dynamic Range and Bit Depth in Digital SLRs”. http://theory.uchicago.edu/~ejm/pix/20d/tests/noise/

[8] Richard, H. “Contrast Thresholds of the Human Eye.” (1946)

[9] Hoefflinger, Bernd. High-Dynamic-range (Hdr) Vision. Springer, 2007.

[10] Yadid-Pecht, Orly. “Wide-Dynamic-range Sensors.” Optical Engineering 38.10 (1999): 1650-60.

[11] Banterle, Francesco, Alessandro Artusi et al. Advanced High Dynamic Range Imaging: Theory and Practice. CRC Press, 2011.

[12] Ohta, Jun. Smart Cmos Image Sensors and Applications. CRC Press, 2007.

[13] Immunolight. http://www.immunolight.com/technology/

[14] IR Photography. http://dpanswers.com/content/irphoto.php

[15] Gilblom, David L, and Sang Kuen Yoo. “Infrared and Ultraviolet Imaging With a Cmos Sensor Having Layered Photodiodes.” Electronic Imaging 2004 International Society for Optics and Photonics, 2004. 186-92.

[16] Heng Z. “Applied Colour Processing”. PhD Thesis. Oregon University (2011)

[17] Committee on Colorimetry, “The Science of Colour”. Optical Society of America, Washington 1973, ISBN 0-96003-801-9

[18] d’Yves Galifret, “Les Mécanismes de la Vision”. Pour La Science 1979, ISBN 2-9029-18-43-7

[19] Babel Colour, “A Review of RGB colour Spaces” http://www.babelcolor.com

[20] Foster, David H. “Color Constancy.” *Vision research* 51.7 (2011): 674-700.

[21] Land, Edwin H. *The Retinex Theory of Color Vision.* Scientific America., 1977.

[22] Wang, Zhou, Alan Conrad Bovik et al. “Image Quality Assessment: From Error Visibility to Structural Similarity.” *Image Processing, IEEE Transactions on* 13.4 (2004): 600-12.

[23] Wang, Zhou, Alan C Bovik, and Ligang Lu. “Why is Image Quality Assessment So Difficult.” Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on 4.ICASSP. IEEE, 2002. IV-3313.

[24] “Image Quality Factors for Cameras and Lenses”. http://www.imatest.com/support/image-quality

[25] “Colour Image Formation”, http://www2.cmp.uea.ac.uk/Research/compvis/ColourIntro/

[26] Mann, S, and RW Picard. “On Being ‘undigital’with Digital Cameras: Extending Dynamic Range By Combining Differently Exposed Pictures, 7 Pages.”

[27] Kay, SM. “Fundamentals of Statistical Signal Processing: Estimation Theory Ptr Prentice Hall, Englewood Cliffs, Nj, 1993. 6 Ll Scharf.” *Statistical Signal Process: Detection, Estimation, and Time Series Analysis*

[28] Kirk, Kristian, and Hans Jørgen Andersen. “Noise Characterization of Weighting Schemes for Combination of Multiple Exposures.” BMVC Citeseer, 2006. 1129-38.

[29] Debevec, Paul E, and Jitendra Malik. “Recovering High Dynamic Range Radiance Maps From Photographs.” ACM SIGGRAPH 2008 classes ACM, 2008. 31.

[30] Mitsunaga, Tomoo, and Shree K Nayar. “Radiometric Self Calibration.” Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. 1 IEEE, 1999.

[31] Reinhard, Erik, Wolfgang Heidrich et al. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting.* Morgan Kaufmann, 2010.

[32] Tsin, Yanghai, Visvanathan Ramesh, and Takeo Kanade. “Statistical Calibration of Ccd Imaging Process.” Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on 1 IEEE, 2001. 480-87.

[33] Robertson, Mark A, Sean Borman, and Robert L Stevenson. “Estimation-Theoretic Approach to Dynamic Range Enhancement Using Multiple Exposures.” *Journal of Electronic Imaging* 12.2 (2003): 219-28.

[34] Granados, Miguel, Boris Ajdin et al. “Optimal Hdr Reconstruction With Linear Digital Cameras.” Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on.CVPR. IEEE, 2010. 215-22.

[35] Buchsbaum, Gershon. “A Spatial Processor Model for Object Colour Perception.” *journal of the Franklin institute* 310.1 (1980): 1-26.

[36] Gershon, Ron, Allan D Jepson, and John K Tsotsos. “From [R, G, B] to Surface Reflectance: Computing Color Constant Descriptors in Images.” IJCAI1987. 755-58.

[37] Maloney, Laurence T, and Brian A Wandell. “Color Constancy: A Method for Recovering

Surface Spectral Reflectance." *JOSA A* 3.1 (1986): 29-33.

[38] Forsyth, David A. "A Novel Algorithm for Color Constancy." *International Journal of Computer Vision* 5.1 (1990): 5-35.

[39] J. Hadamard, Lectures on Cauchy's Problem in Linear Partial Differential Equations, Yale University Press, New Haven, 1953.

[40] Maloney, Laurence T. "Evaluation of Linear Models of Surface Spectral Reflectance With Small Numbers of Parameters." *JOSA A* 3.10 (1986): 1673-83.

[41] D'Zmura, Michael, and Geoffrey Iverson. "Color Constancy. I. Basic Theory of Two-Stage Linear Recovery of Spectral Descriptions for Lights and Surfaces." *JOSA A* 10 (1993): 2148-65.

[42] D'Zmura, Michael, and Geoffrey Iverson. "Color Constancy. Ii. Results for Two-Stage Linear Recovery of Spectral Descriptions for Lights and Surfaces." *JOSA A* 10 (1993): 2166-80.

[43] Hordley, Steven D. "Scene Illuminant Estimation: Past, Present, and Future." *Color Research and Application* 31.4 (2006): 303-14.

[44] K. Barnard. Practical Colour Constancy. PhD thesis, Simon Fraser Univ., School of Computing Science, 2000.

[45] Finlayson, Graham D. "Color in Perspective." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 18.10 (1996): 1034-38.

[46] Finlayson, Graham D, and Ruixia Xu. "Convex Programming Color Constancy." The IEEE Workshop on Color and Photometric Methods in Computer Vision2003.

[47] Finlayson, Graham, and Steven Hordley. "Improving Gamut Mapping Color Constancy." *Image Processing, IEEE Transactions on* 9.10 (2000): 1774-83.

[48] Brainard, David H, and William T Freeman. "Bayesian Method for Recovering Surface and Illuminant Properties From Photosensor Responses." IS&T/SPIE 1994 International Symposium on Electronic Imaging: Science and Technology International Society for Optics and Photonics, 1994. 364-76.

[49] D'Zmura, Michael, Geoffrey Iverson, and Benjamin Singer. "Probabilistic Color Constancy." *Geometric representations of perceptual phenomena* (1995): 187-202.

[50] Sapiro, Guillermo. "Bilinear Voting." Computer Vision, 1998. Sixth International Conference on IEEE, 1998. 178-83.

[51] Finlayson, Graham D, Steven D Hordley, and Paul M Hubel. "Color By Correlation: A Simple, Unifying Framework for Color Constancy." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 23.11 (2001): 1209-21.

[52] Barnard, Kobus, Lindsay Martin, and Brian Funt. "Colour By Correlation in a Three-Dimensional Colour Space." *Computer Vision-Eccv 2000*. Springer, 2000. 375-89.

[53] Tominaga, Shoji, Satoru Ebisui, and Brian A Wandell. “Scene Illuminant Classification: Brighter is Better.” *JOSA A* 18.1 (2001): 55-64.

[54] Rosenberg, Charles, Alok Ladsariya, and Tom Minka. “Bayesian Color Constancy With Non-Gaussian Models.” Advances in neural information processing systems2003. None.

[55] Cardei, Vlad C, Brian Funt, and Kobus Barnard. “Estimating the Scene Illumination Chromaticity By Using a Neural Network.” *JOSA A* 19.12 (2002): 2374-86.

[56] Funt, Brian V, Mark S Drew, and Jian Ho. “Color Constancy From Mutual Reflection.” *International Journal of Computer Vision* 6.1 (1991): 5-24.

[57] Klinker, Gudrun J, Steven A Shafer, and Takeo Kanade. “A Physical Approach to Color Image Understanding.” *International Journal of Computer Vision* 4.1 (1990): 7-38.

[58] Lee, Hsien-Che. “Method for Computing the Scene-Illuminant Chromaticity From Specular Highlights.” *JOSA A* 3.10 (1986): 1694-99.

[59] Hsien-Che Lee. Method for computing scene-illuminant chromaticity from specular highlights. In Glenn E. Healey and Steven A. Shafer and Lawrence B. Wolff, editor, Color, pages 340–347. Jones and Bartlett, Boston, 1992.

[60] Tominaga, Shoji, and Brian A Wandell. “Standard Surface-Reflectance Model and Illuminant Estimation.” *JOSA A* 6.4 (1989): 576-84.

[61] Theuwissen, Albert JP. “Image Processing Chain in Digital Still Cameras.” VLSI Circuits, 2004. Digest of Technical Papers. 2004 Symposium on IEEE, 2004. 2-5.

[62] Image Quality Factors. http://www.imatest.com/docs/iqfactors/

[63] Losson, Olivier, Ludovic Macaire, and Yanqin Yang. “Comparison of Color Demosaicing Methods.” *Advances in Imaging and Electron Physics* 162 (2010): 173-265.

[64] Sergej, Tomaz. “Perceptual Evaluation of Demosaicing Techniques.” *West Pomeranian University of Technology in Szczecin, Poland* (2013)

[65] Theuwissen, Albert, John Coghill et al. “Ultra-High Resolution Image Capturing and Processing for Digital Cinematography.” *IEEE ISSCC Dig. Tech. Papers* (2003): 162-63.

[66] Lluis-Gomez, A.L Edirisinghe, EA. “Chromatic Aberration Correction in Raw Domain for Image Quality Enhancement in Image Sensor Processors.” Intelligent Computer Communication and Processing (ICCP), 2012 IEEE International Conference on.ICCP. IEEE, 2012. 241-44.

[67] Kuno, Tetsuya, Hiroaki Sugiura et al. “Aliasing Reduction Method for Color Digital Still Cameras With a Single-Chip Charge-Coupled Device.” *Journal of Electronic Imaging* 8.4 (1999): 457-66.

[68] [8] Li, Xin, Bahadir Gunturk, and Lei Zhang. “Image Demosaicing: A Systematic Survey.” Electronic Imaging 2008 International Society for Optics and Photonics, 2008. 68221J-68221J.

[69] Sharpening: unsharp mask. http://www.cambridgeincolour.com/tutorials/unsharp-mask.htm

[70] Ward, Greg. “Fast, Robust Image Registration for Compositing High Dynamic Range

Photographs From Hand-Held Exposures." *Journal of graphics tools* 8.2 (2003): 17-30.

[71] Goshtasby, A Ardeshir, and Stavri Nikolov. "Image Fusion: Advances in the State of the Art." *Information Fusion* 8.2 (2007): 114-18.

[72] Piella, Gemma. "A General Framework for Multiresolution Image Fusion: From Pixels to Regions." *Information fusion* 4 (2003): 259-80.

[73] Wang, Anna, Haijing Sun, and Yueyang Guan. "The Application of Wavelet Transform to Multi-Modality Medical Image Fusion." Networking, Sensing and Control, 2006. ICNSC'06. Proceedings of the 2006 IEEE International Conference on IEEE, 2006. 270-74.

[74] Qiguang, Miao, and Wang Baoshu. "A Novel Image Fusion Method Using Contourlet Transform." Communications, Circuits and Systems Proceedings, 2006 International Conference on 1 IEEE, 2006. 548-52.

[75] Sroubek, F, S Gabarda et al. "Multifocus Fusion With Oriented Windows." Microtechnologies for the New Millennium 2005 International Society for Optics and Photonics, 2005. 264-73.

[76] Li, Zhenhuo, Zhongliang Jing et al. "Pixel Visibility Based Multifocus Image Fusion." Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on 2 IEEE, 2003. 1050-53.

[77] Zaveri, Tanish, and Mukesh Zaveri. "A Novel Two Step Region Based Multifocus Image Fusion Method." *Int. J. Comput. Electr. Eng* 2.1 (2010): 86-91.

[78] Zhu, Qiang, Mei-Chen Yeh et al. "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients." Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on 2 IEEE, 2006. 1491-98.

[79] Lluis-Gomez, A, S Saravi, and EA Edirisinghe. "Subjectively Optimised Multi-Exposure and Multi-Focus Image Fusion With Compensation for Camera Shake." SPIE Photonics Europe International Society for Optics and Photonics, 2012. 84360Q-84360Q.

[80] Romanenko, IV, EA Edirisinghe, and D Larkin. "Block Matching Noise Reduction Method for Photographic Images Applied in Bayer Raw Domain and Optimized for Real-Time Implementation." SPIE Photonics Europe International Society for Optics and Photonics, 2012. 84370F-84370F.

[81] Liu, Ce, and William T Freeman. "A High-Quality Video Denoising Algorithm Based on Reliable Motion Estimation." *Computer Vision–eccv 2010*. Springer, 2010. 706-19.

[82] Romanenko, Ilya V, EA Edirisinghe, and Daniel Larkin. "Spatial-Temporal Noise Reduction Method Optimized for Real-Time Implementation." IS&T/SPIE Electronic Imaging International Society for Optics and Photonics, 2013. 86550L-86550L.

[83] Romanenko, IV, Alexis Lluis-Gomez, and Eran Edirisinghe. “Image Matching in Bayer Raw Domain to Remove Ghosting in Multi-Exposure Image Fusion.” Consumer Electronics (ICCE), 2014 IEEE International Conference on.ICCE. IEEE, 2014. 37-38.

[84] Romanenko. I. “Novel Image Processing algorithms and methods for improving their robustness and operational performance”. PhD thesis. Loughborough University, 2013.

[85] Romanenko, IV, EA Edirisinghe, and D Larkin. “Image Matching in Bayer Raw Domain to De-Noise Low-Light Still Images, Optimized for Real-Time Implementation.” IS&T/SPIE Electronic Imaging International Society for Optics and Photonics, 2014. 90200Y-90200Y.

[86] Foi, Alessandro, Mejdi Trimeche et al. “Practical Poissonian-Gaussian Noise Modeling and Fitting for Single-Image Raw-Data.” *Image Processing, IEEE Transactions on* 17.10 (2008): 1737-54.

[87] Liu, Ce, William T Freeman et al. “Noise Estimation From a Single Image.” Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on 1 IEEE, 2006. 901-8.

[88] Blanksby, Andrew J, Marc J Loinaz et al. “Noise Performance of a Color Cmos Photogate Image Sensor.” Electron Devices Meeting, 1997. IEDM’97. Technical Digest., International IEEE, 1997. 205-8.

[89] Petrovi, Vladimir S, and Costas S Xydeas. “Sensor Noise Effects on Signal-Level Image Fusion Performance.” *Information Fusion* 4.3 (2003): 167-83.

[90] OmniVision OV8835. http://www.ovt.com/products/sensor.php?id=127

[91] Sony NEX-5. http://www.sony.co.uk/support/en/product/NEX-5

[92] Foster, David H, Sérgio MC Nascimento et al. “Four Issues Concerning Colour Constancy and Relational Colour Constancy.” *Vision Research* 37.10 (1997): 1341-45.

[93] Zhang, Buyue, and Aziz U Batur. “A Real-Time Auto White Balance Algorithm for Mobile Phone Cameras.” Consumer Electronics (ICCE), 2012 IEEE International Conference on.ICCE. IEEE, 2012. 1-4.

[94] Nikkanen, Jarno, Timo Gerasimow, and Lingjia Kong. “Subjective Effects of White-Balancing Errors in Digital Photography.” *Optical Engineering* 47.11 (2008): 113201-113201.

[95] Kruithof, Arie Andries. “Tubular Luminescence Lamps for General Illumination.” *Philips Tech. Review* 6 (1941): 65-96.

[96] Oi, Naoyuki, and Hironobu Takahashi. “Preferred Combinations Between Illuminance and Color Temperature in Several Settings for Daily Living Activities.” Proceedings of the 2nd International Symposium on Design of Artificial Environments2007. 214-15.

[97] Agarwal, Vivek, Besma R Abidi et al. “An Overview of Color Constancy Algorithms.” *Journal of Pattern Recognition Research* 1 (2006): 42-54.

[98] Land, Edwin H. *The Retinex Theory of Color Vision.* Scientific America., 1977.

[99] Cohen, Noy. “A Color Balancing Algorithm for Cameras.” *EE368 Digital Image Processing* (2011)

[100] Finlayson, Graham D, and Elisabetta Trezzi. “Shades of Gray and Colour Constancy.” Color and Imaging Conference 2004(1).1. Society for Imaging Science and Technology, 2004. 37-41.

[101] Van De Weijer, Joost, Theo Gevers, and Arjan Gijsenij. “Edge-Based Color Constancy.” *Image Processing, IEEE Transactions on* 16.9 (2007): 2207-14.

[102] Gijsenij, Arjan, Theo Gevers, and Joost Van De Weijer. “Improving Color Constancy By Photometric Edge Weighting.” *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 34.5 (2012): 918-29.

[103] Rosenberg, Charles, Alok Ladsariya, and Tom Minka. “Bayesian Color Constancy With Non-Gaussian Models.” Advances in neural information processing systems2003. None.

[104] Gehler, Peter Vincent, Carsten Rother et al. “Bayesian Color Constancy Revisited.” Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on IEEE, 2008. 1-8.

[105] Datasets, Colour constancy, http://colorconstancy.com/?page_id=21

[106] Mitchell, Tom M. “Machine Learning. Wcb.” (1997)

[107] Colour correction matrix. http://www.imatest.com/docs/colormatrix/ , May 2014

[108] Worthey, James A, and Michael H Brill. “Heuristic Analysis of Von Kries Color Constancy.” *JOSA A* 3.10 (1986): 1708-12.

[109] Available Source code, Colour constancy http://colorconstancy.com/?page_id=19#wge, May 2014

[110] Opteka 8" X 10" Digital Color & White Balance Grey Card Set for Digital Photography http://opteka.com/dgcl.aspx, May 2014

[111] D. Coffin. Raw photo decoder ”dcraw” v8.77. http://cybercom.net/ dcoffin/dcraw/.

[112] Shi, Lilong, and Brian Funt. “Re-Processed Version of the Gehler Color Constancy Dataset of 568 Images.” *Simon Fraser University* (2010)

[113] Freeman, WT, and DH Brainard. “Bayesian Decision Theory, the Maximum Local Mass Estimate, and Color Constancy.” *Computer Vision* (1995)

[114] McCann, JJ. “Hdr Imaging and Color Constancy: Two Sides of the Same Coin.” *IS&T/SPIE Electronic Imaging* (2011)

[115] True HDR for iPhone. http://www.pictional.com/TrueHDR/Overview.html

[116] Arnold, Claus. “Surface Color Perception Under Different Illuminants and Surface Collections.” 2009.

[117] McCann, J. “Hdr and Color Constancy-New Psychophysical Results.” *Google tech talks* (2009).

[118] Cambridge in Colour, Tutorials: white balance http://www.cambridgeincolour.com/tutorials/white-balance.htm

[119] Kuang, Jiangtao, and Weihua Xiong. “Color Constancy for Multi-Illuminants High-Dynamic-range Scenes.” Color and Imaging Conference 2008(1).1. Society for Imaging Science and Technology, 2008. 56-60.

[120] Wang, Yang, and Dimitris Samaras. “Estimation of Multiple Illuminants From a Single Image of Arbitrary Known Geometry.” *Computer Vision—eccv 2002*. Springer, 2002. 272-88.

[121] Joze, Hamid Reza Vaezi, and Mark S Drew. “Exemplar-Based Color Constancy and Multiple Illumination.” *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 36.5 (2014): 860-73.

[122] Zhang, Heng, Huaping Liu, and Shuxue Quan. “Multi-Illuminant Color Constancy for Hdr Images Through Exposure Segmentation.” IS&T/SPIE Electronic Imaging International Society for Optics and Photonics, 2012. 829204-829204.

[123] Aptina AR132AT imaging sensor https://www.aptina.com/products/image_sensors/ar0132at/

[124] Akenine-Möller, Tomas, Eric Haines, and Naty Hoffman. *Real-Time Rendering.* CRC Press, 2008.

[125] Colour filter array. https://en.wikipedia.org/wiki/Color_filter_array

[126] Planckian locus https://en.wikipedia.org/wiki/Planckian_locus

[127] The Kruithof Curve – Color Temperature VS Illuminance. http://www.jimonlight.com/2009/09/25/the-kruithof-curve-color-temperature-vs-illuminance/

[128] Saw Tooth Signal Test Chart. http://www.image-engineering.de/products/charts/all/517-te192