

Retina-Inspired and Physically Based Image Enhancement

Ange Abdelhamid AbdelMalek Aly

A thesis submitted in partial fulfilment of the requirements of
Staffordshire University for the degree of
Doctor of Philosophy

Staffordshire University
School of Digital, Technologies and Arts
United Kingdom

A thesis submitted in partial fulfilment of the requirements of
Staffordshire University for the degree of
Doctor of Philosophy

May 2020

*To my dad, Professor A. AbdelMalek Aly . . .
Following your footsteps*

*"Now, do you not see that the eye embraces the beauty of the whole world.
It counsels and corrects all the arts of mankind... it is the prince of
mathematics, and the sciences founded on it are absolutely certain."*

Leonardo da Vinci

Acknowledgements

First and foremost, I would like to express my gratitude to my principal supervisor Dr Tomasz Bosakowski for his ultimate encouragement and guidance. I would like to thank my second supervisor Dr Justin Champion for being part of my supervision team.

I am also grateful to my examiners Prof Ella Pereira and Dr Carolin Bauer for reviewing my thesis and for their useful discussion and guidance.

I would like to extend my thanks to Staffordshire University for offering me the opportunity to pursue my PhD ambition.

My deepest gratitude goes to my mum and dad , for their love and support, they gave me the opportunities and experiences that have made me who I am today, my dear sister May; her dedication and enthusiasm have always been an inspiration for me. Thank you for always being there for me.

For my sons Hamed and Hamid, You have made me stronger, better and more fulfilled than I could have imagined. I love you to the moon and back.

Finally, to my dear husband Mohamed, for his love and support through my PhD and throughout my life, who is without him, it would have been impossible to reach this stage.

Abstract

Images and videos with good lightness and contrast are vital in several applications, where human experts make an important decision based on the imaging information, such as medical, security, and remote sensing applications.

The well-known image enhancement methods include spatial and frequency enhancement techniques such as linear transformation, gamma correction, contrast stretching, histogram equalization and homomorphic filtering. Those conventional techniques are easy to implement but do not recover the exact colour of the images; hence they have limited application areas. Conventional image/video enhancement methods have been widely used with their different advantages and drawbacks; since the last century, there has been increased interest in retina-inspired techniques, e.g., Retinex and Cellular Neural Networks (CNN) as they attempt to mimic the human retina. Despite considerable advances in computer vision techniques, the human eye and visual cortex by far supersede the performance of state-of-the-art algorithms.

This research aims to propose a retinal network computational model for image enhancement that mimics retinal layers, targeting the interconnectivity between the Bipolar receptive field and the Ganglion receptive field.

The research started by enhancing two state-of-the-art image enhancement methods through their integration with image formation models. In particular, physics-based features (e.g. Spectral Power Distribution of the dominant illuminate in the scene and the Surface Spectral Reflectance of the objects contained in the image are estimated and used as inputs for the enhanced methods). The results show that the proposed technique can adapt to scene variations such as a change in illumination, scene structure, camera position and shadowing. It gives superior performance over the original model.

The research has successfully proposed a novel Ganglion Receptive Field (GRF) computational model for image enhancement. Instead of considering only the interactions between each pixel and its surroundings within a single colour layer, the proposed

framework introduces the interaction between different colour layers to mimic the retinal neural process; to better mimic the centre-surround retinal receptive field concept, different photoreceptors' outputs are combined. Additionally, this thesis proposed a new contrast enhancement method based on Weber's Law.

The objective evaluation shows the superiority of the proposed Ganglion Receptive Field (GRF) method over state-of-the-art methods. The contrast restored image generated by the GRF method achieved the highest performance in contrast enhancement and luminance restoration; however, it achieved less performance in structure preservation, which confirms the physiological studies that observe the same behaviour from the human visual system.

List of Contents

Acknowledgements	iv
Abstract	v
List of Contents	vii
List of Figures	xi
List of Tables	xiv
List of Abbreviations	xv
Chapter 1	1
Context and Scope of Work	1
1.1 <i>Context</i>	1
1.1.1 Image Enhancement	2
1.1.2 Image Formation Models	3
1.1.3 Bio-Inspired Image Enhancement Models	4
1.2 <i>Scope and Objectives of the Investigation</i>	5
1.3 <i>Main Contributions</i>	6
1.4 <i>Research Approach</i>	7
1.5 <i>Thesis organization</i>	8
Chapter 2	10
Image Enhancement – Review	10
2.1 <i>Image Enhancement Problem</i>	10
2.2 <i>Image enhancement architectures</i>	12
2.3 <i>Image Decomposition</i>	13
2.3.2 Non-Physics-Based Methods.....	15
2.4 <i>Image Filtering</i>	16

2.4.1	Frequency-Domain Methods	16
2.4.2	Spatial-Domain Methods.....	17
2.5	<i>Bio-inspired methods</i>	18
2.5.1	Retinex Algorithms	18
2.5.2	Multiscale Retinex Colour Restoration (MSRCR).....	19
2.6	<i>State-of-the-Art Image Enhancement Methods</i>	20
2.6.1	Convolutional Neural Networks (CNN).....	20
2.6.2	Low Light Image Enhancement with Illumination Map Estimation (LIME)	21
2.6.3	Multi-Deviation Fusion-Based Method (MF).....	22
2.6.4	Naturalness Preserved Enhancement Algorithm (NPEA)	22
2.6.5	Low-light Enhancement CAmera Response Model (LECARM)	23
2.6.6	Multiscale Adaptation Model (MAM)	24
2.6.7	Naturalness Preservation Model.....	25
2.7	<i>Summary</i>	26
Chapter 3		27
The Retina.....		27
3.1	<i>Retina layers</i>	27
3.2	<i>Receptive Field</i>	29
3.3	<i>Colour Vision</i>	31
3.4	<i>Physical vs perceived light intensity contrast: Weber- Fechner’s law</i>	35
3.5	<i>Retina Circuitry Computational Model</i>	36
3.6	<i>Summary</i>	38
Chapter 4		39
Image Formation Models		39
4.1	<i>Vision System</i>	39
4.2	<i>The Dichromatic Model</i>	41
4.3	<i>Surface Spectral Reflectance Estimation</i>	42
4.3.1	Linear Models	42

4.4	<i>Illuminant estimation</i>	43
4.5	<i>Colour space</i>	45
4.5.1	RGB.....	45
4.5.2	Normalised <i>rgb</i>	46
4.5.3	HSV.....	46
4.5.4	Ycbr.....	47
4.6	<i>Summary</i>	48
Chapter 5		49
Proposed Physics-Based Retina-Inspired Models		49
5.1	<i>Introduction</i>	49
5.2	<i>Multiscale Adaptation Model MAM</i>	49
5.2.1	MAM computational implementation.....	49
5.2.2	Physics-Based-Multiscale Adaptation Model (MAM _P): Enhanced model.....	57
5.3	<i>Naturalness preservation Model</i>	61
5.3.1	Computational Model.....	61
	Step Two: Insertion of brightness.....	63
5.3.2	Physics-Based – Naturalness preservation model: Enhanced model.....	65
5.4	<i>Ganglion Receptive Field Model (GRF): Novel model</i>	67
5.4.1	The computational model.....	70
5.5	<i>Summary</i>	79
Chapter 6		80
Performance Evaluation		80
6.1	<i>Introduction</i>	80
6.2	<i>Performance metrics</i>	81
6.2.1	Naturalness preservation assessment.....	81
6.2.2	Illumination-based quality estimation.....	83
6.2.3	Structure consistency and contrast restoration.....	85
6.3	<i>Proposed contrast enhancement assessment criterion</i>	88

6.4	<i>Datasets</i>	91
6.5	<i>Illustrative visual comparison</i>	91
6.6	<i>Objective evaluation</i>	98
6.6.1	Quantitative testing – video signal	102
6.6.2	Evaluation.....	103
6.7	<i>Summary</i>	104
Chapter 7		106
Conclusions and Further Work		106
7.1	<i>Overview</i>	106
7.2	<i>Contribution to Knowledge</i>	109
7.3	<i>Work Limitation</i>	112
7.4	<i>Recommendations for Further Work</i>	113
Appendix A		115
The Retinal Anatomy		115
A.1.	<i>Introduction</i>	115
A.2.	<i>Anatomic Layers of the Retina</i>	117
A.3.	<i>Neurological network function</i>	117
A.1.1.	Photoreception.....	118
A.1.2.	Transmission to Bipolar Cells Through Synapses	119
A.1.3.	Transmission to Ganglion Cells.....	120
A.1.4.	Transmission along the Optic Nerve	120
A.1.5.	Receptive field	121
References		129

List of Figures

Figure 1-1 Scope of the work	6
Figure 1-2 Research methodology sequence diagram	8
Figure 2-1 Image enhancement architecture	13
Figure 2-2 Image enhancement principle in the frequency domain (Bedi and Khandelwal, 2013).	16
Figure 2-3 Left: A 3-layer, standard Neural Network. Right: A ConvNet arranges three-dimensional neurons (width, height, depth) as seen in one of the layers (Stanford, 2021).	21
Figure 3-1 A Top: a human eye. Bottom: the cross-section of a human retina (Carreras <i>et al.</i> , 2017).	28
Figure 3-2 Receptive field = centre + surround (A Dahl, 2015).....	29
Figure 3-3 Illuminated or stay neutral if both centre and surroundings are illuminated. Likewise, other two-colour	31
Figure 3-4 Normalized spectral sensitivity functions of the LMS cones, (Carreras, et al., 2017).	32
Figure 3-5 Left: Additive mixing of primary colours (RGB) gives all the complementary colours and white. Right: Subtractive mixing of the complementary colours gives the primary colour (but darker) and black.....	33
Figure 3-6 The opposing mechanism of the colour of LGN-Ganglion.....	34
Figure 3-7 On the left: The anatomy of the retinal layers This figure shows the connectivity and hierarchical structure of the retinal cells. On the right: Diagram of the bio-inspired colour enhancement model. Two key processes are bipolar and ganglion, centre/surround, colour-opponent processes and ganglion spatiotemporal responses. The double arrows show feedback and control between horizontals and cones, and amacrine and bipolar. The LGN and cortical processes are not investigated in this thesis (Yufeng Zheng, 2009)	37
Figure 4-1 Schematic diagram of image formation	40
Figure 4-2 Spectral sensitivity characteristics of the Sony ICX098BQ camera (Sony ICX098BQ).....	43
Figure 4-3 Parkkinen’s first three basis-functions	43
Figure 5-1 Multi-scale adaptation model flow-chart.....	52
Figure 5-2 In the new enhanced model, the cones and rod spectral characteristics are multiplied by the SSR, and the result is integrated to obtain the cone and rod responses, four images representing the calibrated photoreceptor responses (L, M, S, R)	58
Figure 5-3 Long(red), Medium (Green) and Short(Blue) Cones Sensitivities and Rod sensitivities(Black).....	59
Figure 5-4 <i>Framework of the naturalness preservation model technique</i>	61
Figure	5-5
<i>Framework of the enhanced naturalness preservation model technique</i>	66
Figure 5-6 Proposed method block diagram.....	69

Figure 5-7 Reconstructed SSR	74
Figure 5-8 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods	77
Figure 5-9 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods	78
Figure 6-1 Naturalness-preservation examples	83
Figure 6-2 Illumination-based quality examples	84
Figure 6-3 Structure consistency quality examples	88
Figure 6-4 Contrast enhancement quality example and their associated maps	90
Figure 6-5 Illustrative visual comparison of the results obtained, the original image (left), the output of the multiscale model of adaptation (middle) and physics-based retina-inspired model (right).	92
Figure 6-6 Illustrative visual comparison of the results obtained, the original image (left), the output of the multi-scale model of adaptation (middle) and physics-based retina-inspired model (right).	93
Figure 6-7 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods	95
Figure 6-8 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods	96
Figure 6-9 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods	97
Figure 6-10 Frame number 1500, library video, thermal dataset	103
Figure 6-11 Illustrative visual comparison of the results obtained, the original image (left), the output of the multi-scale model of adaptation (middle) and physics-based retina-inspired model (right).....	103
Figure 6-12 Average Precision, Recall and F-measure for all videos of the Thermal category.	104
Figure A-1 Top: a human eye. Bottom: the cross-section of a human retina(Carreras, et al., 2017).....	116
Figure A-2 Cross section of the retina (Dubuc, 2002).....	118
Figure A-3 Rod-Cones distribution in the fovea (A Dahl, 2015).....	119
Figure A-4 Bipolar cells (A Dahl, 2015).....	120
Figure A-5 Receptive field = centre + surround (A Dahl, 2015).	121
Figure A-6 Different Rods paths to the Ganglion cells.....	122
Figure A-7 Bipolar-cells-Ganglion cells interactions (Oguztoreli et al., 1985).	123
Figure A-8 Illuminated or stay neutral if both center and surroundings are illuminated. Likewise, other two-color	123
Figure A-9 Voltage firing rate of ganglion cell types ON-centre and Off-centre depending on a light spot (yellow circle) at different areas of the cell's receptive field. When the light spot is on, a yellow bar shows the voltage-spike indicator. (a) Rest (no light) situation	

with spontaneous activity; (b – f) ON-center type of ganglion cells; (g – i) off-center cell type126
Figure A-10 Receptive fields of two ganglion cells127
Figure A-11 Centre-surround organization using horizontal cells circuitry127

List of Tables

Table 3-1: A taxonomy of retina cell functions and image processing.....	30
Table 5-1 p-values for different spatial frequencies.....	53
Table 6-1 Quantitative measurement results of LOE	98
Table 6-2 Quantitative measurement results of CE.....	99
Table 6-3 Quantitative measurement results of PCQI	99
Table 6-4 Quantitative measurement results of QI	100
Table 6-5 Quantitative measurement results of e	100
Table 6-6 Quantitative measurement results of r	101
Table 6-7 Quantitative measurement results of σ	101
Table 6-8 Overall Quantitative measurement results	102

List of Abbreviations

CE	<i>Contrast Enhancement</i>
CNN	Cellular Neural Networks
DoRF	Diverse database of Real-world camera response Functions
HVS	<i>Human Visual System</i>
GRF	<i>Ganglion Receptive Field</i>
PCQI	<i>Patch-based Contrast Quality Index</i>
LECARM	<i>Low-light Enhancement CAmera Response Model</i>
LGN	<i>Geniculates Lateral Nucleus</i>
LIME	<i>Low-Light Image Enhancement via Illumination Map Estimation</i>
LOE	<i>Lightness Order Error</i>
MAM	<i>Multiscale Adaptation Model</i>
MF	<i>Multideviation Fusion</i>
MRI	<i>Magnetic Resonance Imaging</i>
MSRCR	<i>Multi Scale Retinex Colour Restoration</i>
NPEA	<i>Naturalness Preservation Enhancement Algorithm</i>
SPD	<i>Spectral Power Distribution</i>
SSR	<i>Surface Spectral Reflectance</i>
RRF	<i>Retinal Receptive Field</i>

Chapter 1

Context and Scope of Work

1.1 Context

Image enhancement plays a fundamental role in computer vision. Images and videos with good lightness and contrast are vital in several areas of applications, where human experts make an important decision based on the imaging information, such as medical, security, forensic and remote sensing applications (Saichandana & Ramesh, 2014).

An image enhancement algorithm aims to improve the information perception in images/videos for a human viewer and support an image/video processing algorithm with a better input. Such algorithms are processed image that is more suitable than the original image for a specific application (Maini & Aggarwal, 2010).

Typically, the enhancement is performed to improve certain desired features, e.g., image contrast, lightness or grey-levels distribution. The choice of these features and the way they are modified are specific to a given task.

The well-known image enhancement methods include spatial and frequency enhancement techniques such as linear transformation, gamma correction, contrast stretching, histogram equalisation, homomorphic filtering. Those conventional techniques are easy to implement but generally do not recover the exact colour of the images; hence they have limited application areas. For example, one of the drawbacks of histogram equalisation is generating colour distortion in images during enhancement.

Conventional image enhancement methods have been widely used with their different advantages and drawbacks; recently, there has been increased interest in retina-inspired image enhancement techniques, e.g., Retinex and Cellular Neural Networks (CNN), as they attempt to mimic the human visual system.

For many computer vision applications, the possibility of deriving colour models which are less sensitive to changes in the imaging parameters, e.g. illumination conditions, scene structure and different viewpoints, than raw camera outputs is highly desirable.

To overcome this limitation, physics-based image formation models that can represent the effects of changes in imaging factors have been proposed by several researchers. Indexing on such models has shown to deliver better recognition results. Recently, physics-based image formation models, e.g., the Dichromatic model, have gained the attention of the computer vision research community as they describe mathematically how a digital image is formed inside a camera, and they attempt to estimate the Spectral Power Distribution (SPD) of the dominant illuminate as well as the Surface Spectral Reflectance (SSR) of objects.

The current computational Retinal network representations are limited to the superficial layers of the Retina and do not represent the interconnectivity between different layers. Recent advancement in Neuroscience has allowed a better understanding of such relationships between different Retinal layers.

1.1.1 Image Enhancement

Image enhancement approaches may be classified depending on the image representation used as non-physics-based or physics-based. Non-physics-based image enhancement methods use one of the known colour spaces as a cue to model the scene. At the same time, the word physics refers to the extraction of intrinsic features about the materials contained in the scene based on an understanding of the underlying physics that govern the image formation. This process is achieved by applying physics-based image formation models to estimate or eliminate the illumination, and geometric parameters extract information about the SSR (Sedky, *et al.*, 2014).

Recently, physics-based image formation models, e.g., the Dichromatic model, have gained the attention of the computer vision research community as they describe mathematically how a digital image is formed inside a camera, and they attempt to estimate the SPD of the dominant illuminate as well as the SSR of objects.

This research investigates the integration of physics-based image formation models with different retina-inspired computational models for image enhancement applications. In particular, physics-based features (e.g., SPD of the dominant illuminant and the SSR of

the objects contained in the image) was estimated and used as inputs to one of the retina-inspired algorithms to overcome the limitations of the image enhancement. The estimated physics-based spectral image representations deduced from the dichromatic model would represent a more realistic input to the retina-inspired models and would mimic the signal received by the human eye.

1.1.2 Image Formation Models

Digital images depend on four main elements: Illuminate, medium, material and vision system. The illuminat represents the source of the electromagnetic energy of illumination, and its SPD could characterise it. The medium is where the electromagnetic waves travel, while the surface of the material modulates the incident electromagnetic energy represented by its SSR. The vision system is represented by the spectral sensitivities of its sensors which detect the electromagnetic light waves, i.e., the spectral sensitivities of the rods and cones for the human eye or the red, green and blue filters in a digital camera.

Conventional video cameras, analogous to a retina, sense reflected light so that colour values integrate the product of incident illumination power spectral distribution, object's surface spectral reflectance and camera sensors sensitivities. Humans can separate the illumination power spectral distribution from the surface spectral reflectance when judging object appearance, and such ability is called colour constancy (Ho, *et al.*, 1990). Researchers in the field of colour constancy (Ho, *et al.*, 1990; Klinker, *et al.*, 1990; Cardei, *et al.*, 2002; Sapiro, 1999) have done great efforts to extract new physics-based features from the image giving more care to the elements which govern the camera output, and new models are introduced. This physical approach studies colour image formation and estimates or eliminates scene illumination to extract the object's surface spectral reflectance, considering camera sensors sensitivities.

The adopted physics-based image formation models denote the image representation in image enhancement algorithms and their links to different colour spaces used in state-of-the-art approaches.

Image Enhancement literature shows little work done toward using physics-based image representations; only simple approximated models, discussed in 4.3.1, are investigated and implemented. In the literature of image enhancement, the main gap is that the use of such physical models has not been fully investigated.

1.1.3 Bio-Inspired Image Enhancement Models

Despite considerable advances in computer vision techniques, the human eye and visual cortex by far supersede the performance of state-of-the-art algorithms. Marr's approach to studying computer vision starts by following representational theories of minds (Marr & Vision, 1982). He argues that to fully understand a particular machine carrying out a particular information-processing task, we have to study that machine and the information-processing task. He saw the computational study of vision as tightly linked to psychophysics and neurophysiology. But the last three decades have seen some weakening of that integration.

Bio-inspired computational models are derived from the neurological anatomy of the Human Visual System (HVS). Retinex, a compound word comprised of retina and cortex, is a computational theory derived from anatomy and neuroscience; it is based on a simple image formation model, shading model, integrated with both brightness and colour perception model human visual system. Retinex is a prevalent and effective method to remove environmental light interferences. The two main concepts behind the theory are: (1) colours of objects are determined by the capability to reflect rays of light, and (2) colours are not influenced by uneven illumination. The main advantages of Retinex are image sharpening and colour constancy (Land, 1986).

Georgeson and Sullivan (Georgeson & Sullivan, 1975) have studied apparent contrast at different spatial frequencies, in foveal and peripheral vision, in photopic and scotopic conditions, and to varying orientations in astigmatic subjects. Their experiments found that despite gross differences in the contrast thresholds for two patterns, they will match in apparent contrast when their physical contrasts are equal. They have concluded that, in

man, visual information is blurred by optical and neural processes restored by an active process of compensation; they have defined this ability as ‘contrast constancy’.

Similarly, humans can separate the illumination power spectral distribution from the surface spectral reflectance when judging object appearance; such ability is called ‘colour constancy’ (Ho, *et al.*, 1990).

1.2 Scope and Objectives of the Investigation

This research aims to propose a Retinal network computational model for image enhancement that mimics the Retinal layers, namely targeting the interconnectivity between the bipolar receptive field and the Ganglion receptive field.

To achieve the aim of this research, the following objectives were to be met:

1. To review the literature on image enhancement techniques,
2. To review the literature on physics-based image formation models, including the dichromatic model, shading model, illumination estimation, and surface spectral reflectance estimation,
3. To study in detail the highly structured neuron network of the retina and existing computational models for the retina,
4. To propose and develop a new physics-based image enhancement framework, which mimics biological retina functioning layers, integrated with the dichromatic model,
5. To propose a novel image enhancement computational model based on the proposed physics-based and retina-inspired model,
6. To test the proposed image/video enhancement technique, using appropriate image data sets, which include different illumination, scene structure, camera position and shadowing,
7. To evaluate the test results with state-of-the-art image/video enhancement techniques.

1.3 Main Contributions

The main concern of image enhancement is to improve image quality. The image improvement can be achieved through various enhancement techniques where each depends on the required task.

Image enhancement applications are endless, such as photography, printing, machine vision, satellite image processing, microscope image processing, medical image processing, face recognition, video analytic, etc.

Image enhancement improves information perception in images for human viewers and image/video re-processing step for a computer vision software; those enhancements are performed to modify the image contrast, brightness or grey-levels distribution. Image enhancement passes through two main stages: image transformation to achieve a suitable image representation and filtering, as shown in Figure 1-1.

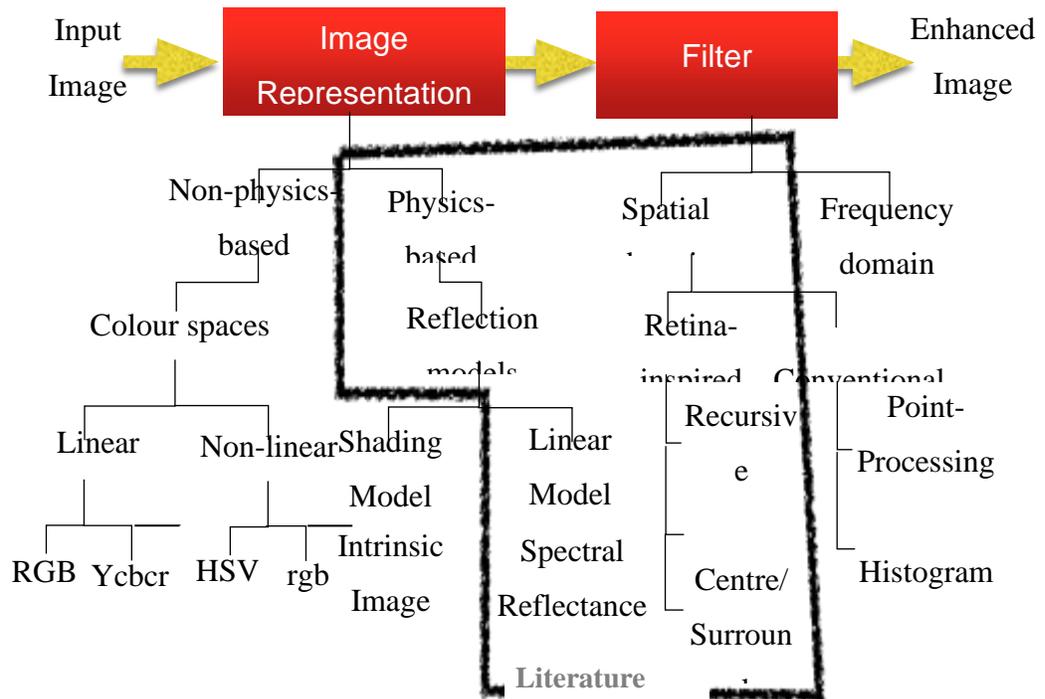


Figure 1-1 Scope of the work

1.3 Main Contributions

The contributions to knowledge are:

1. A novel physics-based computational model of the retina mimics biological retina functioning layers, integrated with the dichromatic image formation model. (Primary contribution of this research project)
2. Two modified image enhancement models, which incorporate a physics-based image formation model and the dichromatic model. In particular, physics-based features (e.g., Power Spectral Distribution of the dominant illuminate in the scene and the Surface Spectral Reflectance of the objects contained in the image are estimated and are used as inputs to the model for adaptation). The results show that the proposed technique can adapt to scene variations such as a change in illumination, scene structure, camera position and shadowing and gives superior performance over the original model.
3. A new contrast enhancement assessment criterion based on Weber's law image. (Primary contribution of this research project)

1.4 Research Approach

The methods of investigation for this research include an interdisciplinary literature review to assist the design of a novel framework to simulate the primate remarkable ability to achieve the challenging task of image enhancement. This project started by reviewing the literature of image enhancement techniques; this research informed the capture of the requirements, the understanding of the challenges, and the identification of key techniques and their advantages, disadvantages and limitations. Then the neurological functions of the human retina will be studied.; this was followed by conducting secondary research to evaluate state-of-the-art retinal computational models and image formation models to identify the models that suit the application in hand image/video enhancement. A novel physics-based and retina-inspired technique were proposed using the information gathered from primary research.

The scene variation is a significant challenge for any image enhancement technique; hence the choice of the dataset used to evaluate the performance of the proposed technique. A

real-world data set was used, where the proposed methods were objectively evaluated before and after the integration.

To develop a robust image enhancement technique, it is required to accurately quantify the performance over a range of data representative of operational conditions. The available tools to evaluate the recognition rate and visualize the results are the image quality performance metrics. Figure 1-2 shows the sequence diagram of the adopted research methodology.

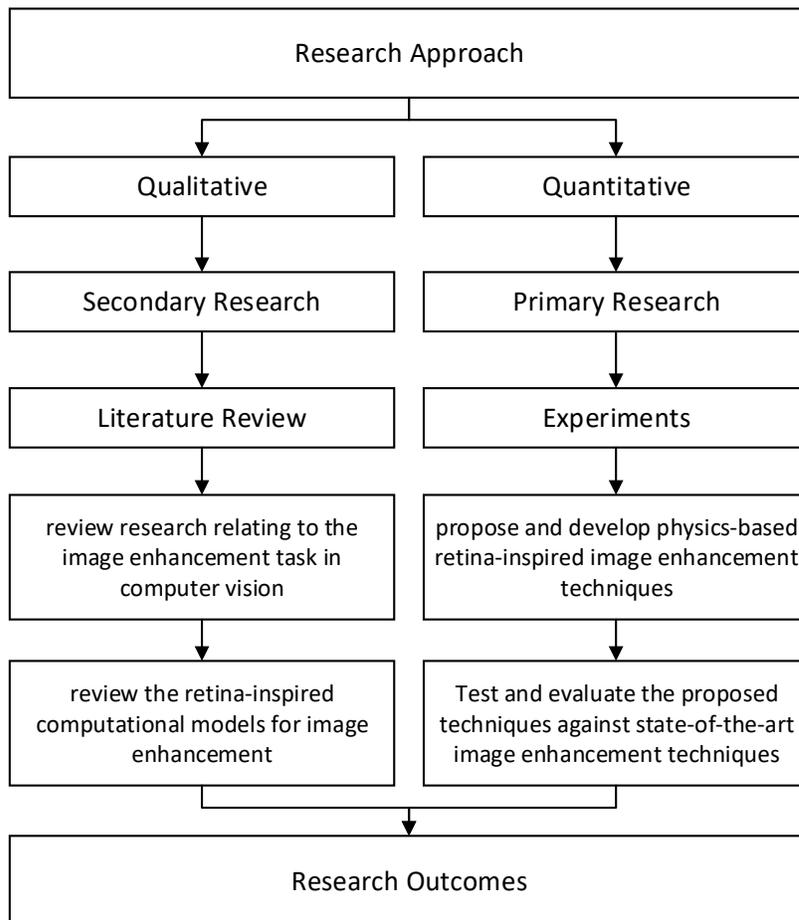


Figure 1-2 Research methodology sequence diagram

1.5 Thesis organization

This thesis is organized as follows:

Chapter 1 gives an overview of image enhancement applications in computer vision. The objectives, as well as the contribution to the knowledge of this research work, are stated. Finally, the chapter concludes with a brief overview of the Chapters to follow.

Chapter 2 reviews the state-of-the-art of image enhancement techniques. The image enhancement problem is first defined. The main computational steps involved in image enhancement are introduced. The various physics-based and non-physics-based image representations are reviewed. Methods for image representation and filtering are explained.

Chapter 3 Introduces a background on the retinal neural network and explain the concept of the receptive field.

Chapter 4 Introduces fundamental physics-based image formation models and their relation with different colour spaces.

Chapter 5 presents new models that illustrate the integration of the image formation model with different bio-inspired models. The models introduced are the physics-based - Multisacale adaptation model, the physics-based- Naturalness model and the Ganglion Receptive Field (**GRF**) Model. GRF is a novel technique that mimics the retinal neural network in processing the visual scene.

Chapter 6 illustrates the adopted testing and evaluation methods for each of the three proposed models, including an illustrative visual comparison, objective, and subjective evaluation.

Chapter 7 presents conclusions and recommendations for future work. It discusses the contributions of this research work, the practical results obtained, and the future work required for improving the proposed image enhancement techniques.

Chapter 2

Image Enhancement – Review

Image enhancement has been subject of many studies in several application areas including video analytics (Sedky, *et al.*, 2014), medical diagnosis (Saba, *et al.*, 2018) (Ortiz, *et al.*, 2012) (Ortiz, *et al.*, 2012), industrial inspection (Rahman, *et al.*, 2018), remote sensing (Saichandana, *et al.*, 2014), object-oriented video coding, interactive gaming and forensic (Cao, *et al.*, 2014). The task of image enhancement is the improvement of information perception in images for human viewers and/or as image/video re-processing step for a computer vision software; those enhancements are performed to modify the image contrast, brightness or grey-levels distribution.

Image enhancement passes through four main stages, image decomposition, filtering, tone mapping and image reconstruction. Such techniques have become important in the past two decades since it represents a key computer vision component that contributes to a broad spectrum of applications in diverse disciplines.

This chapter presents a review of current advances within the field of Retina-inspired image enhancement. The chapter begins by formally defining the image enhancement problem, illustrating different factors that can affect image appearance and discussing the main computational steps involved in image enhancement. The various physics-based and non-physics-based image representations are reviewed to achieve an image representation that is useful for image enhancement. Image filtering techniques that exploit this information for image enhancement are then described.

2.1 Image Enhancement Problem

Conventional cameras, similar to our Retina, sense colours from a scene. The colour values observed are the product of the incident illumination power spectral distribution, the object's surface spectral reflectance and the cameras own sensing abilities. The human body deals with sensing the same information (through the retina) by separating the illumination from the surface spectral reflectance; this allows humans to perceive the

colour of objects to remain generally the same in different illumination conditions known as colour constancy. Research into machine vision has led to extracting new physics-based features from the image (Fu, *et al.*, 2014), enabling more helpful information about the scene to be understood.

Not all digital images have a visual appearance that satisfies the human eye. Dissatisfaction can arise due to noise, lack of illumination quality in the images where it is either too dark or too bright. So, there is a need for methods to enhance the quality of digital images to display the detail on the image for easy observation. These techniques aim to improve the contrast of the input image while preserving its brightness and details. The image enhancement problem could be articulated as to how to manipulate the characteristics and features of an image to improve it for the desired application. Many image enhancement techniques have been presented for each specific purpose, where they could be divided into two main tasks contrast enhancement and naturalness preservation, which conflict with each other.

The spatially varying contrast of the human eye is well-studied in the perception literature (Majumder & Irani, 2006). The human eye is more sensitive to contrast than pixel intensities, and it perceives less information from an image with insufficient intensity distributions than from the same image with better intensity distributions (Valliammal & Geethalakshmi, 2011). Accordingly, in scientific imaging application, spatial correlation is more important than absolute pixel intensity.

Enhancement of many natural images poses special challenges because there may be less spectral variation from pixel to pixel, and most importantly, there may be a high degree of correlation among the data from different sources.

Contrast enhancement is a necessary aiding tool for medical imaging for reducing noise and sharpening details to improve the visual representation of the image in Magnetic Resonance Imaging (MRI) (Rundo *et al.*, 2019) and ultrasound (Ortiz, *et al.*, 2012).

Contrast enhancement comforts us to upturn the vividness of an image. Medical type images always suffer from low contrast (Jan, *et al.*, 2018).

2.2 Image enhancement architectures

Image enhancement improves information perception in images for human viewers and as an image/video re-processing step for a computer vision software; those enhancements are performed to modify the image contrast, brightness or grey-levels distribution. The architecture of image enhancement typically consists of four cascaded modules, image decomposition, image filtering, tone mapping and image reconstruction, and a feedback module, image quality assessment (see Figure 2-1).

The tone mapping module aims to model different local illumination variation and varieties of scene structures. An image reconstruction module is crucial to filter out irrelevant changes, such as camera noise, holes in the enhanced image before generating the final output. An image quality assessment module is required to update the tone mapping parameters to adapt to global illumination conditions.

The image decomposition module comprises the choice of the image representation feature(s), where an ideal representation can convert the input image into information that describes the material and is invariant to the illumination. The image filtering module aims to statistically manipulate the chosen feature(s) to discount colour patches, illumination variation and non-uniformly illuminated images.

The tone mapping module aims to define the decision rule, which will be implemented to maintain colour and contrast constancies.

In this chapter, existing image enhancement methods are classified as either non-physics-based or physics-based depending on the chosen type of image representation. Non-physics-based methods use one of the known colour spaces as a cue to model the scene. At the same time, the word physics refers to the extraction of intrinsic features about the materials contained in the scene based on an understanding of the underlying physics that govern the image formation.

This process is achieved by applying image formation models that attempt to estimate or eliminate the illumination to extract information about the surface spectral reflectance.

The rest of the chapter is dedicated to reviewing the two main domains of image filtering and getting into more details about the bio-inspired spatial domain filters, as the

integration of both the physics-based image decomposition method and the bio-inspired spatial filters is the focus of this research.

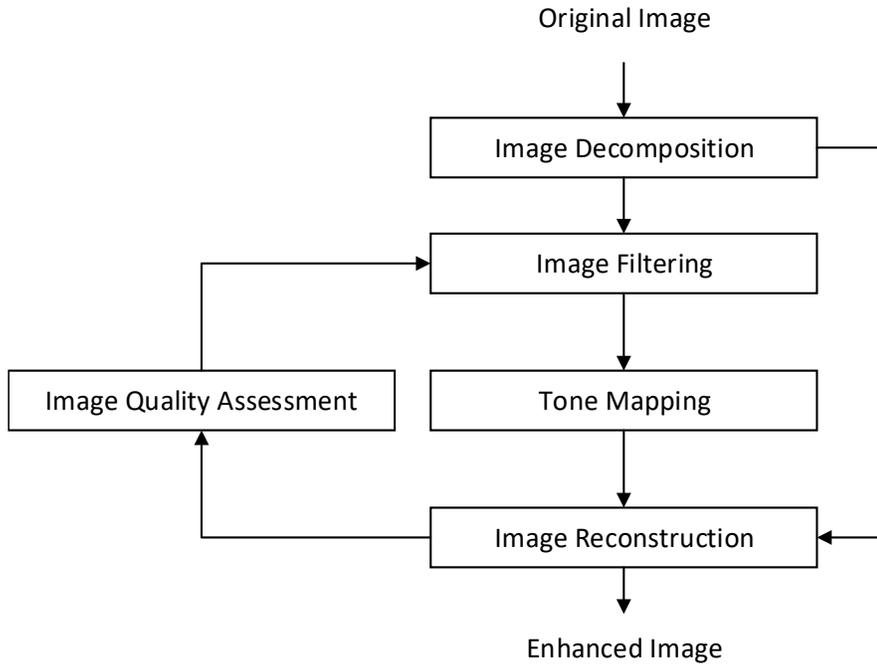


Figure 2-1 Image enhancement architecture

2.3 Image Decomposition

Colour-based image enhancement applications face the challenge that colours are variant to illumination. Each image representation has its strength and weakness and is particularly applicable for handling a particular type of variation. Photometric invariant features describe the colour configuration of each image pixel to compute a part of the input images that are invariant to confounding scene properties but are discriminative on desired scene information.

An important factor in image/video enhancement is the choice of the transformation applied to the raw data to obtain the information relevant to the specific application domain. The basic idea behind image representation is how to select the photometric invariant features, which are the best invariance to illumination changes while maintaining an accurate representation of the object(s) contained in the scene. Several image

representations have been proposed, classified into three types: spectral, spatial, and temporal features.

Spectral features could be associated with grey scale (intensity, brightness, luminance), raw colour, normalized colour, colour information (saturation, hue, chromaticity), intrinsic image and reflectance ratio. Some spatial features are also exploited, such as spatial gradients and texture, and temporal features could be associated with temporal gradients at the pixel, inter-frame changes and optical flow.

Spectral features are sensitive to camera noise and cannot be computed accurately in regions with low intensities. Spatial features can obtain invariance to local illumination and camera noise but are not suitable for an area with low spatial gradients. General robustness to lighting changes can be achieved using edge-based or block-correlation methods. However, the enhancement obtained by these methods is not precise, as the generated image is either made of blocks or edges and not of accurate regions. The temporal features suffer from the same problem, where they cannot be calculated accurately in areas with low textures. However, it helps handle dynamic scenes and camera motion.

In most image/video enhancement algorithms, the features are chosen arbitrarily, and the same features are used globally over the whole scene.

2.3.1 Physics-Based Methods

The colour is not a physical property of an object (Maloney, 1986); instead, it is a perceptual phenomenon and a subjective human concept. The perceptual ability that allows humans to discount spectral variation in the illumination and assign stable colours to objects is called colour constancy.

The camera output is dependent upon the camera characteristics, the incident illuminant and the scene content. The physical approach, used in colour constancy (Parkkinen *et al.*, 1989), (Klinker *et al.*, 1990) and (Westland *et al.*, 2012) studies colour image formation to estimate scene illumination spectral power distribution to extract the full objects surface spectral reflectance taking into consideration camera sensors sensitivities. Physics-based

representations of objects using images under many lighting conditions have been proposed for illumination invariant image enhancement (Tang *et al.*, 2011) and shadow removal.

Image enhancement based on physical models would permit image representation via the study of image formation. Algorithmically, the basics of these approaches are often very similar to those of other image/video enhancement methods and only differ from them because these algorithms explicitly use the reflectance models of surfaces to represent a colour image.

2.3.2 Non-Physics-Based Methods

The colour distribution in *RGB* space has a wide range of evolution because of the illumination variation at different parts of the field. Lu and Zhang (2002) show that the colour distribution in *UV* space is comparatively stable. They conclude from their initial empirical data that a colour classifier defined on *the UV plane can accurately classify* colours despite the mild illumination variation on the field under a stable lighting condition.

Developers have used different colour spaces, e.g. *RGB*, *rgb*, *HSV*, *YUV* and *CIELAB*. However, normalized *rgb* and *HSV* is the most common colour spaces used. It has been shown that these colour spaces are more tolerant to minor variations in the illumination (Kumar *et al.*, 2002).

The *rgb* colour space is one of the photometric invariants widely used in image enhancement (Salvador *et al.*, 2004). However, a rationale is usually not given; *rgb* representation is invariant to illuminant intensity; it depends on illumination colour and surface reflectance but not on its orientation, making it useful for material-based segmentation. The *rgb* colour space is widely used for its fast computation, and it is more effective than *RGB* coordinates for suppressing unwanted changes due to shadows of objects. However, the *rgb* space suffers from a problem inherent to the normalization, where low intensities result in unstable chromatic components.

Like *rgb*, many existing invariants seek to isolate information about the material properties in a scene and are therefore designed to be invariant to local illumination and viewing geometry. The hue H and saturation S of an *HSV* colour space are used to obtain a limited level of intensity invariance in an indoor scene (François *et al.*, 1999) and (Raja *et al.*, 1998). However, the *HSV* colour space faces unstable Hue values at low Saturation (Blauensteiner *et al.*, 2006).

2.4 Image Filtering

The Filters in image enhancement can be used in two key domains: the spatial domain, which directly deals with the image pixels, where the pixels are manipulated to achieve the desired enhancement or in the frequency domain after the image is transferred to the frequency domain; this means that all the enhancements are done on the Fourier transform of the image in the frequency domain; then an inverse Fourier transform is performed to get the final enhanced image (Saini and Narang, 2013).

2.4.1 Frequency-Domain Methods

The principle of frequency-domain methods is to compute the Fourier transform of an image and multiply the result by a filter. After that, the inverse of this filtered image is taken to produce the enhanced image (see Figure 2-2)(Bedi and Khandelwal, 2013).

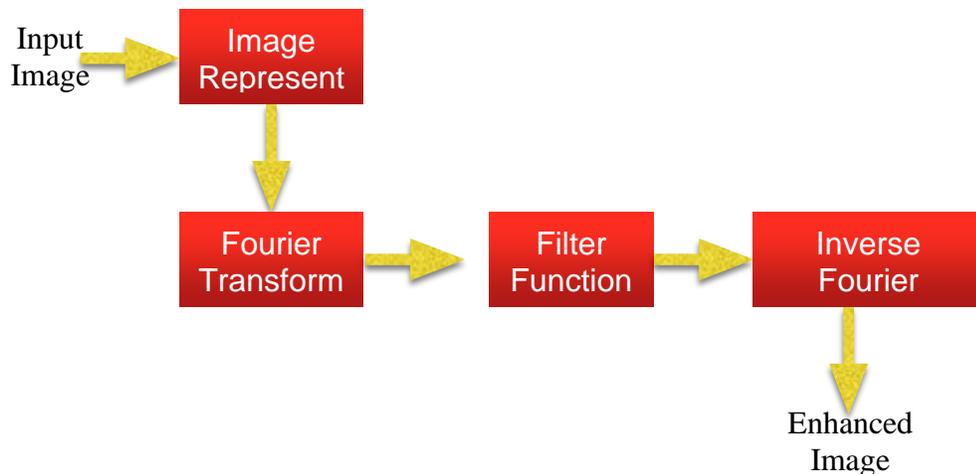


Figure 2-2 Image enhancement principle in the frequency domain (Bedi and Khandelwal, 2013).

The primary filters used in the frequency domain are low pass filters, high pass filters and homomorphic filters.

1. Low-pass filters eliminate high-frequency components from the image resulting in the reduction of the sharp transition associated with noise. Their two main problems are blurring and ringing.
2. High-pass filters are used to sharpen the image. However, overdoing the sharpening can degrade image quality.
3. Homomorphic filters are used for correcting non-uniform illumination in images. Homomorphic filtering increases the image contrast and normalizes the brightness. They are mainly used to remove the multiplicative and additive noise. One of these filter drawbacks is that the illumination and reflectance are not separable (Agrawal *et al.*, 2014).

2.4.2 Spatial-Domain Methods

In the spatial-domain method, the work is done on the pixel values directly; there is a wide variety of techniques to achieve visually accepted images. The choice of such techniques depends on the image content, observer characteristics, and viewing condition.

2.4.2.1 Conventional methods

Point processing methods

The point processing methods are the most primitive yet essential image processing operations in the spatial domain and are used primarily for contrast enhancement. Some examples of point processing techniques are digital negative, contrast stretching, thresholding, grey-level slicing, bit-plane slicing, and dynamic-range compression (Saini and Narang, 2013).

Each of those techniques is suited for a specific task; for example, the negative image technique is suited for enhancing white detail embedded in dark regions, and it has applications in medical imaging. Contrast stretching is done when the captured image has poor illumination; the idea behind contrast stretching is to make dark portions darker and

bright portions brighter. For a dark image, an expansion of grey levels is accomplished using a power-law transformation with a fractional exponent. Power-law transformations are helpful for general-purpose contrast manipulation. Point processing techniques can improve visual appearance or make features easier to detect, but one of the drawbacks is that they may cause loss of information.

Histogram Methods

The histogram provides relevant information regarding the contrast of an image. Histogram equalisation is a transformation that stretches the contrast by redistributing the grey level values uniformly across the histogram (Rahman *et al.*, 1997). The histogram equalisation method is considered broadly applicable and efficient, mainly with images represented by close/low contrast values. For example, the histogram equalisation method is best used with radiographic and thermal images. However, slow speed and enhancement of noise are two of their main drawbacks.

A signal transform equivalent to histogram equalization also seems to happen in biological neural networks to maximize the output firing rate of the neuron as a function of the input statistics; this has been proven in particular in the fly retina (Laughlin, 1981)

2.5 Bio-inspired methods

The mammalian retina is believed to be far more intelligent than scientists have thought. Inspired by the visual processing of Receptive Fields concept in the retina, from the layer of photoreceptors to the layer of retinal ganglion cells (RGCs). Most of the early problems in understanding vision arise from the difficulty of undoing what happens when light projects from a three-dimensional world onto the two-dimensional world onto the back of our eye.

2.5.1 Retinex Algorithms

Retinex can be categorised into three classes: path-based, recursive, and centre/surround Retinex algorithms (Le *et al.*, 2014).

1. **Path-based algorithm:** The original work of Land (Land and McCann, 1971) and (Land 1986) belong to this category. In this algorithm, the lightness of each pixel depends on the multiplication of ratios along random walks. The drawbacks of this algorithm are the dependency of the path geometry and the computational complexity, which has many parameters.
2. **Recursive algorithm:** This was developed by (Frankle and McCann, 1983). This algorithm works on long-distance iterations between pixels, then moves to short-distance interactions. The spacing between pixels being compared decreases at each step in clockwise order. This algorithm is computationally more efficient than the path-based one. The main drawback is that it depends on the number of times a pixel's neighbours are visited. This parameter is named some iteration. The optimal value for this parameter is not clear and can negatively influence the final result (Ciurea and Funt, 2004).
3. **Centre/Surround algorithm** (Morel, 2010): This technique introduces a weight in the reciprocal influence of two pixels (Finlayson, 1995), which is inversely proportional to their square distance.

The lightness values are computed by subtracting a blurred version of the input image from the original image. This algorithm is faster than the path-based one, with fewer parameters. The main drawback is the “grey world” assumption, which was addressed (Berns *et al.*, 1993) by introducing the multi-scale Retinex with colour restoration technique (Reddy, 2013).

2.5.2 Multiscale Retinex Colour Restoration (MSRCR)

The Retinex theory, aimed at understanding the perception of human colour, its derivations have led to efficient algorithms improving local image contrast, thus allowing "see in the shadows" among other features. Multiscale Retinex is possibly the most effective centre-surround image philtre amongst those derived algorithms. The MSCRP offers Multiscale Retinex analysis and implementation by demonstrating that the critical final colour correction step can be seriously improved. If a colour balance is needed, the

MSRCR approach must be applied separately to each colour (Jobson *et al.*, 1996) or the luminance if the goal is to achieve local contrast enhancement (MSCRP). In (Petro, et al., 2014), a new approach was proposed using two different algorithms to deal with both cases.

2.6 State-of-the-Art Image Enhancement Methods

2.6.1 Convolutional Neural Networks (CNN)

CNNs are specially architected Neural Networks. The architecture of the CNN contains a convolution that is followed by a pooling operation. In the input, each neuron in a convolutionary layer is connected to a specific region called a local receptive field. In CNN, all weights are shared within a receptive field based on the position. Filters are also called the common weights (Huang & Lin, 2007).

Standard Neural Networks accept input and move through several layers of hiding. Each of those secret layers consists of some neurons, where each neuron in the previous layer is fully connected to the rest of the neurons, neurons in a single layer will also work in a completely separate way and would not share any connexions. The "performance layer" is considered the new and connected layer.

Convolutional Neural Networks uses the fact that there are images in the input, and they begin to create the architecture more sensibly. Compared to a standard Neural Network, a ConvNet's layers have neurons organised in three dimensions: distance, height, depth. The word depth refers to the third dimension of the activation volume, not the total number of layers within a network. Neurons in a particular layer would be connected in a completely connected manner to a tiny region from the layer before them, instead of the entire neuron. Figure 2-3 shows a simple ConvNet, a series of layers, each layer of ConvNet switches one type of activation by a different feature to another.

Convolutional Layer, Pooling Layer, and Fully-Connected Layer are the three layers used to render ConvNet architectures. These layers could be stacked to form a complete architecture for ConvNets.

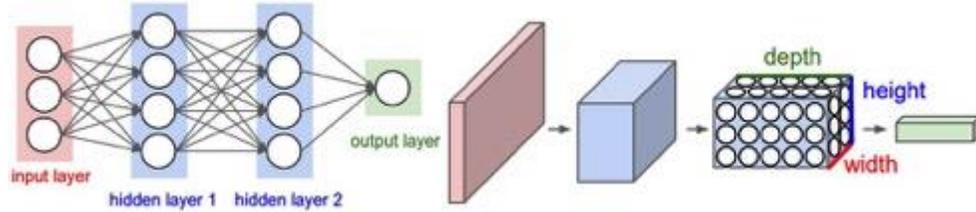


Figure 2-3 Left: A 3-layer, standard Neural Network. Right: A ConvNet arranges three-dimensional neurons (width, height, depth) as seen in one of the layers (Stanford, 2021).

Each layer of a ConvNet converts the volume of 3D input into a volume of neuron activations in 3D output. In this example, the image is held by the red input layer, so its width and height would be the image dimensions, and the depth would be 3 (red, green, blue channels).

2.6.2 Low Light Image Enhancement with Illumination Map Estimation (LIME)

Images frequently suffer from poor visibility in low light conditions. This poor quality can also significantly degenerate the output of many computer vision and multimedia algorithms specifically designed for high-quality inputs. The key to improving low light is how well the map of the illumination is estimated. LIME is proposed in (Guo & Li, 2016), a simple but successful method for enhancing low-light images (LIME). The concept of structure-aware smoothing was developed to improve the accuracy of the lighting. The LIME method uses two algorithms: one to achieve the exact optimal solution by independently calculating the illumination of each pixel to obtain the maximum value in the channels R , G and B . While the other refines the initial illumination map by placing a corresponding structure as a final illumination map. The enhancement can be achieved after having a well-constructed map for illumination.

2.6.3 Multi-Deviation Fusion-Based Method (MF)

The Multi-Deviation Fusion (MF) approach was suggested in (Fu, et al., 2016), enhancing poorly illuminated images using many existing image processing methods. The first stage is calculating the illumination, based on the decomposition of an observed image into a reflection image and an illumination image. Then derive two inputs that use the sigmoid function and adaptive histogram equalisation to reflect luminance-improved and contrast-enhanced versions of the first decomposed lighting—selecting suitable inputs and weights from the estimated illumination. The MF method generates an adapted illumination by fusing the derived inputs in a multi-scale manner with the corresponding weights. With a proper method of weighting and fusion. The final enhanced image is obtained by compensating back to reflectance with the adjusted illumination.

2.6.4 Naturalness Preserved Enhancement Algorithm (NPEA)

The Retinex-based algorithms find illumination elimination as a default choice and fail to restrict the reflectance range; the naturalness of non-uniform illumination images can not be maintained effectively. Naturalness is essential for enhancing the picture in order to achieve pleasing perception consistency. The naturalness Model not only improves picture specifics but retains the naturalness as well. In order to maintain naturalness while reinforcing information, (Wang S *et al.*, 2013), who are the creators of the Naturalness Model, made a measure of Lightness Order Error (LOE) to objectively access preservation of naturalness. The Model applies a bright-pass filter to decompose an image into reflection and illumination, determining the details and naturalness of the image, respectively. Then a bi-log transformation used to map the illumination to balance details and naturalness. Experimental results (Wang, 2013) show that the proposed algorithm can improve the specifics and retain the naturalness for non-uniform illumination images. This model does not take the relation of illumination into account in different scenes; one of its drawbacks is introducing slight flickering for video applications.

The model does not take the illumination relationship into account in various scenes; one of its disadvantages is that it can cause minor flickering for video applications. For non-uniformly illuminated images, the model has a multi-layer enhancement process. Unlike other enhancement methods that empirically increase contrast, this approach modifies multiple components of an image based on high-quality outdoor images' a priori multi-layer lightness statistics. The multi-layer model effectively extracts details from different frequency bands, and the multi-layer lightness prior helps render components with low frequencies.

The basic principle behind the NPEA is the assumption that in multiple frequency bands, image information (reflectance) can exist. Information in all various conditions is unlikely to be extracted easily by a single-layer model. Researchers like (Guo & Li, 2016) and (Fu, et al., 2016), typically fine-tune their single-layer models for different applications.

The multi-layer NPEA model has no clear cut between illumination, reflection and extracts data from different frequency bands; as more layers are used, certain low-frequency information is returned to images of the result. The NPEA does not always perform better than all other methods, and it may seem that the image is barely enhanced when the average value of the estimated illumination for each layer is similar to the previous lightness statistics.

2.6.5 Low-light Enhancement CAmera Response Model (LECARM)

Low-light image enhancement algorithms can increase low-light image visual quality and enable practical information extraction for specific computer vision techniques. Therefore, existing techniques ultimately introduce distortions of colour and lightness when enhancing the images (Ren, et al., 2018). A novel enhancement method using cameras' reaction characteristics to lower the distortions is introduced to evaluate a fair model of camera response and its parameters.

Then the estimation techniques for illumination to estimate the ratio of exposure for each pixel is used. Finally, according to the approximate map of the exposure ratio, the selected

camera response model changes each pixel to the desired exposition. Experiments show that the approach can enhance enhancement with fewer distortions of colour and lightness compared to the other state-of-the-art methods.

The LECARM approach for efficient naturalness-preserved low-light image enhancement blends the conventional Retinex model with the camera response models. The enhanced images are achieved by changing the exposure of low-light images locally. Since the estimation of the precise three-channel curves from a single low-light image is less robust, the Diverse average database of Real-world camera Response Functions (DoRF) database curve is used as an alternative. Using the DoRF database average curve to approximate the accurate three-channel curves for most cameras can only introduce limited distortions to the outcomes. The main advantage of this model is that it can use various camera response models and map estimation strategies for exposure.

2.6.6 Multiscale Adaptation Model (MAM)

The Multi-scale Adaption Model (MAM) (Pattanaik, et al., 1998), attempts to mimic the human visual system through a multi-scale representation of the pattern, luminance, and colour processing in the Retina. In the human vision system, the neural response produced by photoreceptor cells (Rods and cones) depends on chemical reactions produced by light on the cell's photo-pigments. The cell's response to light is limited by the maximum rate and intensity of these chemical reactions. If the reactions are occurring near their maximum levels, and the amount of light striking the photo-pigments increases, the cell may not entirely signal the increase. This situation is known as saturation. The result of saturation is response compression: incremental increases in light intensity will produce smaller and smaller changes in the cell's response rate above a certain level.

The Rods and Cones photoreceptors connect through a network of neurons in the retina to Bipolar cells and Ganglion cells whose axons form the optic nerve. Adaptive processes located in this neural network adjust the base activity and gain of the early visual system to response compression in the photoreceptors. A multiplicative process adjusts the system's gain by selectively scaling the input by a constant related to the background

luminance. This mechanism controls the relationship between the photoreceptor signals and the amplitude of spectral stimulus by turning down the gain when the stimulus amplitude is high and turning up the gain when the stimulus amplitude is low. The gain control is independent of the different photoreceptors and explains the visual system's capability to adjust to both wide ranges of illumination and varying colours of illumination to approximately preserve the appearance of the object colours (Pattanaik, et al., 1998) .

2.6.7 Naturalness Preservation Model

Many image enhancement techniques consider only the image reflectance component and neglect the component for illumination. However, illumination plays a significant role in achieving a satisfactory quality of perception in the picture. The naturalness of an image cannot be maintained without preserving its illumination (Joshi & Prakash, 2020).

The technique suggested supports both grey-scale and colour images. In order to enhance the colour image provided in RGB space, it is first converted to HSV (Hue, Saturation, Value) colour space, and then the "Value" portion in HSV space is enhanced using the proposed technique. The Enhanced Value component is combined with the original Hue and Saturation components to obtain a final enhanced colour image. The technique enhances grey-scale images without any transformation.

The first step, a Bad Illumination Pass Filter (BIPF), is used to locate poorly illuminated areas within the image. The second step uses an Illumination-based image quality estimation technique to analyse both the quality of the input and the enhanced images. The third step applies an Adaptive Logarithm Enhancement (ALE) technique to insert the required amount of brightness in areas with non-uniform illumination. The proposed technique primarily performs two parallel tasks. With the help of BIPF, it produces an enhanced intermediate image in the first mission, which improves by identifying image areas with information and good illumination. In the second task, image quality and applying the ALE logarithm calculate another enhanced image from the input image. Ultimately, to get the final enhanced image, output images of these two tasks are fused into one.

2.7 Summary

Chapter two introduces the image enhancement problem where contrast enhancement and naturalness preservation conflict with each other. The architecture of image enhancement, which consists of four modules, are discussed, where the image representation and image filtering are reviewed in more depth. The different methods to represent an image, based on being physics-based and non-physics based, are reviewed

In image filtering, the two key domains are the frequency-domain and the spatial-domain methods; the principle of frequency-domain methods is to compute the Fourier transform of an image, then multiply the result by a filter. At the same time, the spatial domain deals directly with the pixel value, and there is a wide variety of techniques to achieve visually accepted images. Conventional image enhancement methods have been widely used with their different advantages and drawbacks. Recently, there has been increased interest in Bio-inspired image enhancement techniques, e.g. Retinex, Convolutional Neural Networks (CNN), Multiscale Adaptation Model MAM, and Naturalness, as they attempt to mimic the human visual system. The main advantages of Retinex are image sharpening and colour constancy. The key benefit of CNNs is their accuracy in object recognition. In contrast, the MAM and Naturalness will be studied and implemented in more details in chapter five.

The retina inspired computational models represents the first stage of visual information processing in vertebrates. Despite discoveries regarding the anatomy and physiology of neural structures, we still have an incomplete understanding of the visual pathway; this is mainly because of the high number, various functionality, and complex connection patterns of neurons, which opens the door for a lot of development and enhancements regarding the retina inspired computational models. To model the retinal visual pathway, a thorough understanding of the retinal anatomy and circuitry is required. The retina neurological network is studied in more details in chapter three.

Chapter 3

The Retina

This chapter reviews the different types of retinal neurons and their information processing circuitry. It is a challenge to apply knowledge from neurology research to create an image processing model to mimic the Human Visual System (HVS); a more detailed description of the retinal anatomy is presented in Appendix A.

3.1 Retina layers

The HVS is a retinal organization-dominated network; parallel processing, feed forward, feedback, and lateral connexons (Tuscany *et al.*, 2014). The image processing in the HVS does not start in the brain but rather in the retina, which begins immediately. In addition, the retina and the optic nerve originate as an outgrowth of vertebrate embryonic development. Figure 3.1 portrays a human eye and the cross-section of a retina. The retina is formed by many layers of neural cells, starting with about 130 million photoreceptors (rods and cones) and finishing with about 1 million Ganglion cells. The primary processing that takes place in each cell type is complex and still not fully understood.

The retina's inner surface is parallel to the light that opens the eye. The photo sensors of the retina are reversed, not in the direction of light. The Retinal cells are divided into photoreceptor cells and neuronal cells (Andrew A Dahl, 2015), as shown in Figure 3-2.

The retina is classified into four main layers, (i) The photoreceptors, (ii) the Outer Plexiform Layer (OPL), where the synapses between photoreceptors, Horizontals, and Bipolar take place, (iii) the Inner Plexiform Layer (IPL), which contains the cells bodies of horizontal, bipolar and amacrine and finally (iv) the ganglion cell layer which contains the Retinal Ganglion Cells (RGC). The first layer of the retina consists of rods and cones, they work like the camera sensors, rods are scotopic, and they detect mainly luminance, cones that are photopic detect the colours, cones are mainly three types: long (L), medium (M) and short (S) cones, just like the camera sensors R, G, B.

3.1 Retina layers

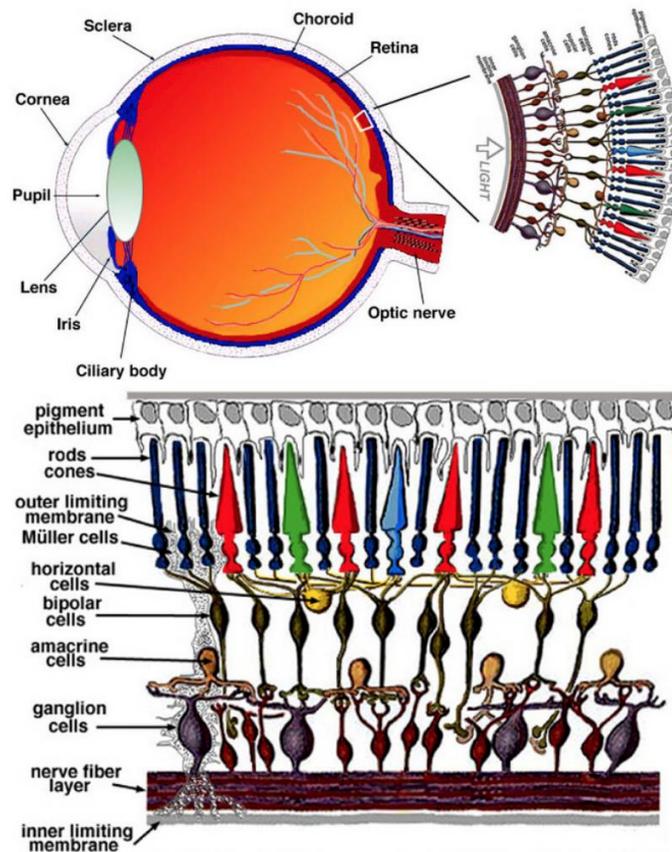


Figure 3-1 A Top: a human eye. Bottom: the cross-section of a human retina (Carreras *et al.*, 2017).

The second layer of the retina consists of the Horizontal cells and the Gap cells, which operate horizontally; the first type binds a group of photoreceptors to the Bipolar cells' surround receptive field and usually ensures the highest non-accuracy of light intensity, this type of cells function as a data compressor. Horizontals enhance the concentric sensitivity of Bipolar and RGCs, exciting or inhibiting surrounding photoreceptors. For high signal speed, the Gap cells connect the rods to the bipolar cone cells.

The third layer of the retina consists of Bipolar cells and Amacrine cells; bipolar cells have a receptive field consisting of a centre and surround; this type of cells function as edge detectors, intensity information, and motion detectors of objects. In comparison, the Amacrine cells function as detectors of background motion.

3.2 Receptive Field

The fourth layer is the Ganglion cell layer, and cells receive their input from Bipolar and Amacrine cells. This type of cells functions as edge detectors, object motion detectors, especially when the object is close, those cells transform information in the forms of spikes. Those electric spikes are sent to three major parallel optic nerve pathways that are linked to the visual cortex in the brain; the P-pathway carries signals from the L and M cones, the K-pathway carries signals from the yellow (L+M) and S cones, and the M-pathway carrying the signals of luminance. Table 3-1 presents a taxonomy of retina cell functions and image processing.

3.2 Receptive Field

In the visual system, volumes in visual space are receptive fields, as shown in. The area of space (e.g., the visual field) through which a stimulus influences neuron firing, e.g., cone, bipolar cell, or ganglion cell) is called the photoreceptor's or neuron's receptive field. The receptive field is always determined relative to the given receptor or neuron. In other words, the stimulus is the activity we observe which affects the neuron's activity or state. If we test individual photoreceptor receptive fields, they will all be identical in that they are determined by whether or not light hits the receptor. However, since many rods converge on their bipolar and ganglion cells while cones project in a 1:1 fashion, the receptive fields of rods will be larger than those of cones, consisting of the sum of the receptive fields of all the rods supplying input, as shown in Figure 3-2

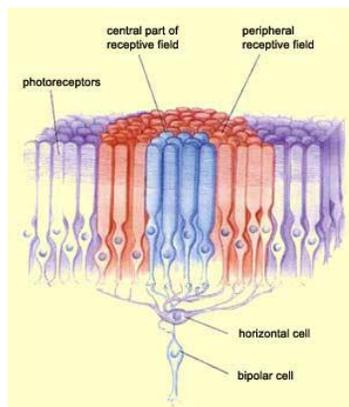


Figure 3-2 Receptive field = centre + surround (A Dahl, 2015).

Human Eye Retina										
Consists of four main layers										
Layer name	Photoreceptors		Outer plexiform layer (OPL)		Inner plexiform layer (IPL)			Ganglion cell layer		
Layer content	Rods	Cones	Horizontal cells	Gap function cells	Bipolar cells		Amacrian cells	Ganglion cells		
					On Bipolar Cells	Off Bipolar cells		On Ganglion Cells	Off Ganglion Cells	
									P -pathway	M-pathway
Cells functions	<ul style="list-style-type: none"> • Scotopic • detect luminance (light intensity) 	<ul style="list-style-type: none"> • Photopic • Detecting color • Three types (L'Red', S'blue', M'green') • L and M centered in the fovea 	<ul style="list-style-type: none"> • Connect the most lit photoreceptors in the surround receptive field to ensure highest light intensity not accuracy. 	<ul style="list-style-type: none"> • connects the rods to cone bipolar cells in mesopic condition for faster transmission to the P- path instead of going through the rod BP cells to M path 	<ul style="list-style-type: none"> • when light hits its centre receptive field it depolarised and hyper polarise when it reaches its surround receptive field 	<ul style="list-style-type: none"> • When light hits its centre it hyper polarise and depolarise when it reaches its surround receptive field. 	<ul style="list-style-type: none"> • 20 types only three are identified 1. Movement sensitive 2. enhances the Centre surround effect in ganglion cell connects rod bipolar cells to cone bipolar cells 3. 	<ul style="list-style-type: none"> • Small G cells 90% • Red/green opponency • signals from L/M • More sensitive to shape and detail. 	<ul style="list-style-type: none"> • large G cells 5% • Luminance (no colours) • More influenced by rods • sensitive to motion • Cause of receptive field size fastest response 	<ul style="list-style-type: none"> • large G cells 5 % • Blue/yellow • Mix signals from L/M & S cons
Image processing	Camera sensor		Compression		<ul style="list-style-type: none"> • Edge detection • Intensity information • Object motion detection 	<ul style="list-style-type: none"> • Background motion detection 	<ul style="list-style-type: none"> • Edge detection • Object motion detection image getting closer (increase in size) • intensity information • uniformly lighted image segments. 			
Type of Signalling	Gradual potential (Hyper polarisation/depolarisation)				Gradual potential (Hyper polarisation/depolarisation)		Action potential (spiks)			

Table 3-1: A taxonomy of retina cell functions and image processing

There are two Bipolar / Horizontal and Ganglion / Amacrine centre / surround layers; and four-colour opponents, Red (R), Green (G), Blue (B), and Yellow (Y). The central cell (Bipolar or Ganglion) takes the surrounding information from one or more Horizontal or Amacrine cells, and both Bipolar and Ganglion have sub-types ON and OFF shown in Figure 3-3. For example, a +R/-G Bipolar (red-centre- ON/green-surround-OFF) will be excited if only the centre is illuminated or inhibited if only the surroundings (Bipolar) are opponents with ON-centre/OFF-surround, +G/-R and +B/-Y, follow the same rules. The yellow (Y) channel can be obtained by averaging red and green channels. On the other hand, Off-centre/ON-surround Bipolar (i.e., -R/+G and -G/+R, but no - B/+Y) are inhibited when the centre is illuminated. An ON-bipolar (or OFF-bipolar) only transfers signals to an ON Ganglion (or OFF-Ganglion), where Amacrine provide surrounding information. (Lamb *et al.*, 2007).

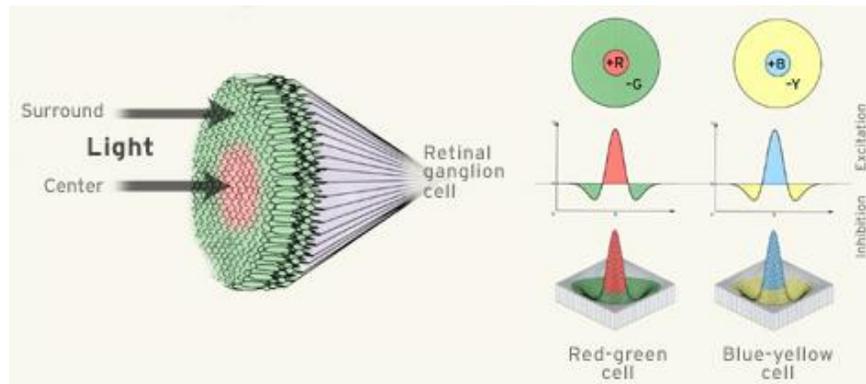


Figure 3-3 Illuminated or stay neutral if both centre and surroundings are illuminated. Likewise, other two-colour

3.3 Colour Vision

The fact that the retina fovea comprises three different types of cones, each with different sensitivity in the wavelength spectrum of Short (S), Medium (M) and Long (L), lets us see colours. The spectral sensitivities are shown in Figure 3-4. The trichromatic theory is called the three-dimensional nature of colour vision. After the theory's nineteenth-century developers, it is also referred to as the "Young-Helmholtz theory." It has shown that each

colour can be produced by mixing different light quantities of the three primary colours: Red, Green and Blue (RGB); this is difficult with two colours, although more than three colours are unnecessary.

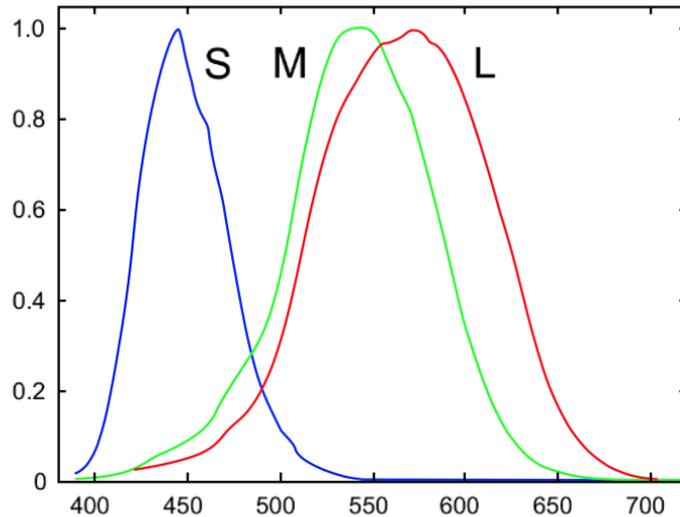


Figure 3-4 Normalized spectral sensitivity functions of the LMS cones¹, (Carreras, et al., 2017).

Compared to mixing paint or subtractive colour combination, the combination of different light colours is called additive colour mixing. The additive mixing effect is lighter than the individual components, while white is obtainable (Figure 3-5, left). The result is darker than the individual subtractive mixing components (the paints absorb light) and eventually black (Figure 3-5, right). Each cone in the fovea connects to its Ganglion cell in the retina. Cones integrate further down the visual path into other forms of Ganglion cells at a substation located in the central part of the brain (the thalamus). This substation is called Geniculates Lateral Nucleus (LGN). LGN-ganglion cells have a receptive field in the retina. Only the cone cells make up this region. De Valois et al. (1966) were the first to use microelectrodes to map the connexion of LGN-ganglion cells within the brain and cones in the fovea. With an ON-OFF function, an LGN-ganglion cell processes the cone signals.

¹ www.iste.co.uk/provenzi/color.zip

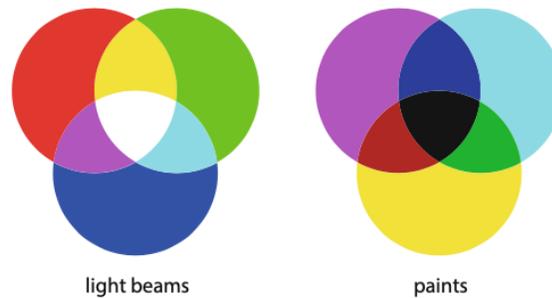


Figure 3-5 Left: Additive mixing of primary colours (RGB) gives all the complementary colours and white. Right: Subtractive mixing of the complementary colours gives the primary colour (but darker) and black

Retinal Ganglion cells process conflicting brightness signals in the receptive field coming from the rods (light "on" and light "off," respectively). In the receptive field, LGN-ganglion cells process opposing colour signals coming from the cones. Hering (1878) proposed that colour vision was based on two pairs of opposing colours even in the late nineteenth century: blue and yellow on one side and red and green. This explanation comes from the fact that there are no bluish-yellow and reddish-green colours, although there are bluish-green (purple) and reddish yellow (orange) colours. Hering's argument was also reinforced by the fact that colour blind for blue is also colour blind for yellow and that the same holds for colour blindness to red and green. Indeed, by comparing the contrasting colours blue and yellow and red and green, the LGN-ganglion cells process the signals from the cones (Hubel, 1995).

Yellow is available for processing by introducing red-sensitive cone (L-type) and green-sensitive cone (M-type) signals. A third class exists, that of opposing black and white, in which the three signals of the responsive red, green and blue cones are incorporated. As with the ON-centre and Off-centre retinal ganglion cells, there are also two groups of LGN-ganglion cells in each opposing colour. The right colour light in the receptive field centre (+) increases the cell output (increases the voltage spikes rate). The right colour

light in the receptive field's surroundings) (decreases (inhibits) the output signal (reduces the spikes rate).

There are also quite a few physiological dimensions of the opposing colour process (Conway 2014; Wuerger and Xiao 2016). Significant observations are being made in the visual cortex about more colour processing processes (Johnson *et al.*, 2008; Harada *et al.*, 2009; Wandell & Chichilnisky, 2012; Mély & Serre, 2017).

The opposing mechanism for the transmission of the colour of LGN-ganglion cells receiving signals from cones. The right colour light in the receptive field centre (+) increases the tension firing rate of the ganglion cell, while the right colour light in the surroundings decreases the tension firing rate).

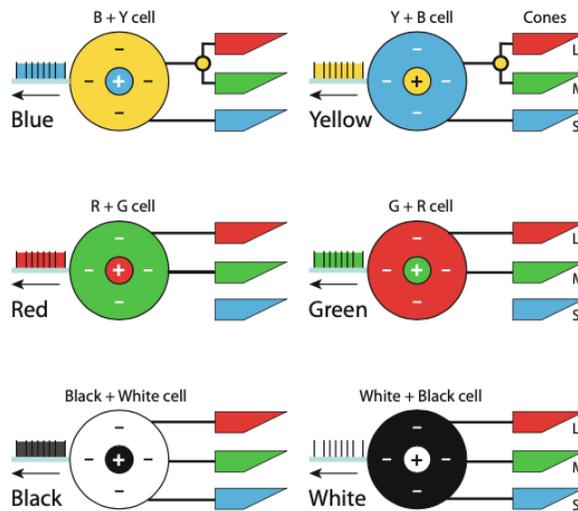


Figure 3-6 The opposing mechanism of the colour of LGN-Ganglion

The opposing mechanism for the transmission of the colour of LGN-ganglion cells receiving signals from cones. The right colour light in the receptive field centre (+) increases the tension firing rate of the ganglion cell, while the right colour light in the surroundings decreases the tension firing rate.

3.4 Physical vs perceived light intensity contrast: Weber- Fechner's law

Psychophysics is the science that aims at mathematically rigorous modelling of the magnitude of human perception in response to external stimuli. The physicist German E. Weber was one of the first scientists in history to develop psychophysical experiments to test the perception of light intensity contrast, with results obtained in the second half of the 19th century.

Weber operated in a very restricted setting: a dark-adapted human observer was positioned in a dim room in front of a white screen, at the centre of the visual field, reflecting a narrow beam of light. The light intensity L of the beam was very slowly increased, and the observer was asked to say whether he/she could sense a change in intensity. For Just Significant Difference, the least noticeable change in severity is called JND . Weber discovered that the JND increased proportionally to light intensity, i.e., somewhere:

$$W = \frac{\delta^*L}{L} = \frac{L_{object} - L_{background}}{L_{background}} \quad \text{Equation 3-1}$$

Where W is called the constant of Weber and the relationship between ubiquitous L_{object} and $L_{background}$ is called the law of Weber.

Weber's law states that if we raise the background light L , the difference L will rise proportionally in order to be able to understand the difference between L_{object} and $L_{background}$; this explains why dark areas of a visual scene are more sensitive to noise, and therefore, denoising in dark areas of digital images is more important than in bright ones. This last consideration is a use of a psychophysical phenomenon in practice.

German experimental psychologist G. Fechner the father of psychophysics. Fechner gave the following interpretation of Weber's law: he introduced the dimension quantity $s(I)$ called light sensation and stated that the difference between sensation and sensation (I) was proportional to a slightly modified Weber ratio.

$$\delta(\text{Perceived Signal}) = k \frac{\delta(\text{Sensed Signal})}{\text{Signal}} \quad \text{Equation 3-2}$$

Where $k > 0$ is a constant, Fechner assumed that this finite-difference equation might be valid for arbitrarily minor differences, which is not possible, of course, because of *JND's* very definition. However, the finite-difference Equation 3-2 following Fechner's theorem, is a differential equation:

$$\text{Perceived Signal} = k \cdot \ln(\text{Sensed Signal}) + C \quad \text{Equation 3-3}$$

Where $C > 0$ is a quantity, often interpreted by the visual mechanism as internal noise. Wherever $C = -k \cdot \ln(n + L_o) + C$. where $\delta(L_o) = 0$. This last theorem is called the law of Weber-Fechner, and it states that the sensation of light differences grows as the logarithm of the light intensity in the very restricted sense of Weber's experiment.

3.5 Retina Circuitry Computational Model

There is a vertical pathway, “Photoreceptors, Bipolar, and Ganglion cells” for a central data process, and lateral networks, “horizontal and amacrine” for the surrounding data process. There are two layers of Bipolar and Ganglion of centre/surround processes and four colour opponents (R, G, B and Y). There are two sub-types for each bipolar (same as ganglion): ON-bipolar and OFF-bipolar. All combinations of ON/OFF and colour opponents are +R/-G +G/-R, +B/-Y, -R/+G and -G/+R (but no -B/+Y). LGN and cortical regions perform the successive feature extraction and pattern recognition. Figure 3-7 illustrates how the human retinal network is presented.

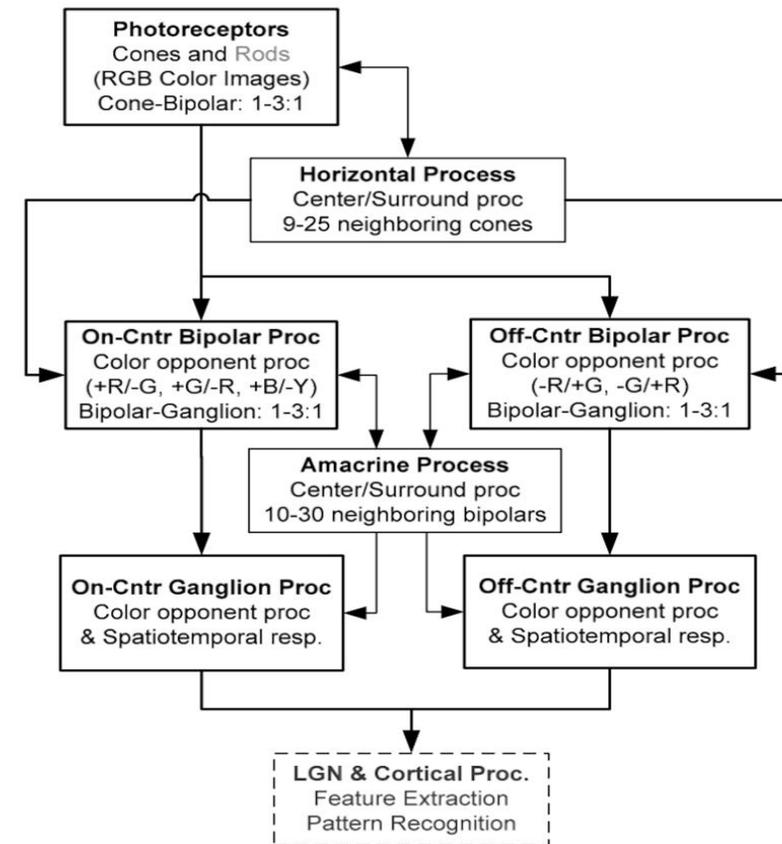
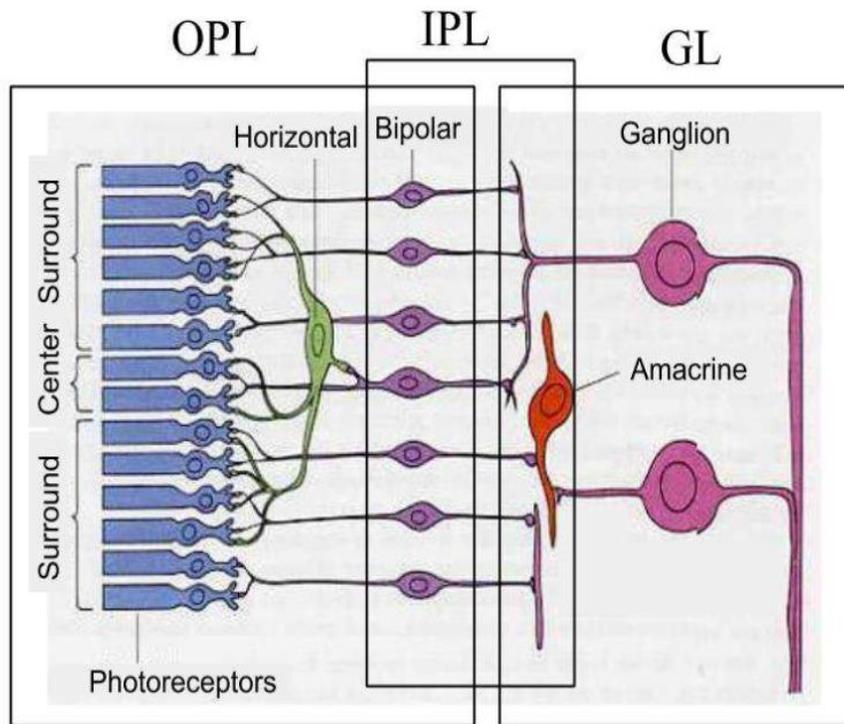


Figure 3-7 On the left: The anatomy of the retinal layers This figure shows the connectivity and hierarchical structure of the retinal cells. On the right: Diagram of the bio-inspired colour enhancement model. Two key processes are bipolar and ganglion, centre/surround, colour-opponent processes and ganglion spatiotemporal responses. The double arrows show feedback and control between horizontals and cones, and amacrine and bipolar. The LGN and cortical processes are not investigated in this thesis (Yufeng Zheng, 2009)

3.6 Summary

This Chapter reviewed the main retina layers, the IPL, which consists of photoreceptors (rods and cones), Bipolar cells and Horizontal cells. The cones are for colour day vision, and rods are sensitive to light and responsible for night vision; photoreceptors act more like camera sensors (R, G, B), Bipolar act like edge and motion detection, Horizontal cells function like information compressors. Then the OPL, which consists of Bipolar and Amacrin cells, where Amacrin function like a background motion detector, next is the GL, which consists of the Amacrin and the Ganglion cells.

Then the chapter reviews the Receptive Field concept as the area of space through which a stimulus influences neuron firing. As an example, the central cell takes the surrounding information from one or more Horizontal cells; and the Bipolar has sub-types ON and OFF; for example, a +R/-G Bipolar (red-centre- ON/green-surround-OFF) will be excited if only the centre is illuminated or inhibited if only the surroundings Bipolar are opponents with ON-centre/OFF-surround, +G/-R and +B/-Y, follow the same rules. Finally, Weber's law, where a contrast enhancement method will be proposed in chapter Six based on Weber's Law

Chapter 4

Image Formation Models

The scenes depend on four fundamental elements: an illuminate, a medium, a material and a vision system. The illuminate represents the source of visible electromagnetic energy and is characterized by its PSD. The medium is the medium in which electromagnetic waves travel. The surface of the material modulates the incident electromagnetic energy and is represented by the surface spectral reflectance, the fraction of incident radiation reflected by this surface. The vision system is identified by the spectral sensitivities of its photosensitive sensors, representing the response of such element to the received reflected light.

Another critical parameter is geometrical features, representing the scene structure, the illuminate orientation, the surface roughness and the viewing geometry. These features combine non-linearly to form a digital image (Sedky, 2014).

Recovering these features from images is an essential problem in image processing; however, this recovery is generally hard with limited information provided by standard commercial imaging devices.

4.1 Vision System

The light emitted by sources of illumination and modulated by surfaces in the scene arrives at the capturing sensors of the colour vision system that is observing the scene. The vision system senses the captured electromagnetic signal and then transforms the information carried by light into a colour image of the physical world.

Colour Charge Coupled Device (CCD) camera, an example of a typical vision system, contains a set of sensors that convert electromagnetic energy into electric signals, which are then sampled and quantized. Conventional CCD cameras insert colour filters, with different spectral sensitivity to the various wavelengths, over each sensory element, typically red, green, and blue filters, to obtain colour information. Figure 4-2 shows the Sony ICX098BQ CCD sensors (Sony ICX098BQ), excluding lens characteristics and

light source characteristics, as an example of a typical surveillance camera. Apart from the spectral sensitivity of the colour filters, the formation of digital image colour values includes other factors, such as lens characteristics and the electronics of the camera.

As Figure 4-1 shows, once a material surface is hit by a bundle of light emitted by an illuminate, the electromagnetic waves may be transmitted, absorbed, or reflected into the air. The quantities of transmitted, absorbed and reflected energy sum to the incident energy at each wavelength. Those quantities are typically measured in relative terms as a fraction of the incident energy. Some materials may emit light, or a fluorescence effect may occur, where the material absorbs light at specific wavelengths and then reflect light at different wavelengths. The surface's absorption, transmittance and reflectance are obtained. In this thesis, materials are assumed to be opaque, so light transmission through the material is not considered. The materials are assumed not to be fluorescent, so the emission does not focus on reflection.

The surface spectral reflectance of a material refers to the ability of the material to reflect different spectral distributions when some light shines on it. A reflectance model is a function that describes the relationship between incident illumination PSD and reflected light at a given point on a surface and each wavelength.

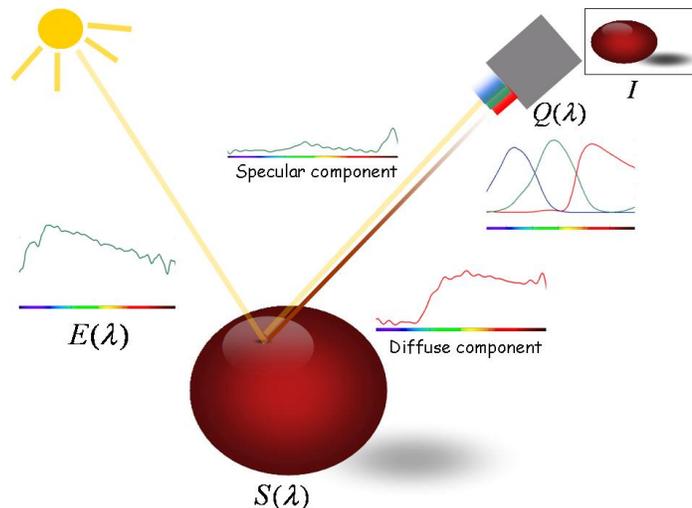


Figure 4-1 Schematic diagram of image formation

$S(\lambda)$ is defined as the ratio between the reflected PSD to the incident PSD. Different materials have different mechanisms of reflection, optically; most materials can be divided into two categories:

1. Homogeneous materials and,
2. Inhomogeneous materials

Homogeneous materials have a uniform refractive index throughout their surface and bodies. On the other hand, inhomogeneous materials have varying refractive index throughout their surfaces and bodies. If light hits its surface, part of the light reflects (specular reflection), while the other part enters the object and then reflects the air causing diffuse reflection. Such material's surfaces are known as diffuse (matte) surfaces or Lambertian surfaces.

4.2 The Dichromatic Model

The dichromatic model represents the light reflected by an inhomogeneous dielectric material as a linear combination of diffuse and specular reflections. Each of these parts is further divided into two elements, one accounting for the geometry and another purely spectral. There are many reflection models, most of them developed in the field of computer graphics. Among these methods, the dichromatic reflection model in (Klinker 2013) is a usual choice for those algorithms employing a physical model to represent colour images, as shown in Figure 4-1, and the equation below:

$$I_c = w_d \int_{\Omega} E(\lambda) \cdot S(\lambda) \cdot Q_c(\lambda) d\lambda + w_s \int_{\Omega} E(\lambda) \cdot Q_c(\lambda) d\lambda \quad \text{Equation 4-1}$$

Where Ω is the visible range from 400nm to 700nm, I_c is the measured colour intensity of the reflected light, w_d and w_s are geometrical parameters for diffuse and specular reflection respectively, $E(\lambda)$ is the spectral power distribution function of the illumination, $S(\lambda)$ is the SSR of the object, $Q_c(\lambda)$ The camera sensor spectral sensitivities characteristic and c represent the colour channels (*Red*, *Green* and *Blue*).

Assuming Lambertian surfaces, theoretically, an image taken by a digital colour camera (for diffuse only reflection) can be described as:

$$I_c = w_d \int_{\Omega} E(\lambda) \cdot S(\lambda) \cdot Q_c(\lambda) d\lambda \quad \text{Equation 4-2}$$

4.3 Surface Spectral Reflectance Estimation

4.3.1 Linear Models

Several researchers (Bajcsy *et al.*, 1990), (Maloney, 1986) and (Marimont *et al.*, 1992) show that both illumination and surface spectral reflectance is relatively smooth functions of the wavelength of light in the visual spectrum and that they can be expressed using finite-dimensional linear models.

The surface reflectance of a great variety of materials has been studied. Parkkinen *et al.* (Parkkinen *et al.*, 1989), Maloney (Maloney, 1986) studied the reflectance properties of the Munsell chips, which is a database of experimentally measured surface spectral reflectance characteristics. Parkkinen concludes that eight basis functions can cover almost all existing data in the Munsell chips database (Westland *et al.*, 2004). However, it has been shown in the literature (Klinker *et al.*, 1990) that the spectral reflectance calculated using the first three basis functions has an average error of 0.0055 and 0.01. Figure 2 shows the first 3 Parkkinen basis functions. As shown from the figure, a significant characteristic of Parkkinen basis functions is the flat distribution of his first basis function.

$$S(\lambda) = \sum_{i=1}^n w_i \cdot \varphi_i(\lambda) \quad \text{Equation 4-3}$$

Where $\varphi_i(\lambda)$ is the i^{th} reflectance linear basis function, and w_i is its corresponding weight, n is the number of basis functions used.

Using finite-dimensional linear models to represent the surface spectral reflectance provides a compact description of data; with few basis functions, we can represent surface spectral reflectance for general materials. The linear models have been extensively used

4.4 Illuminant estimation

in some colour constancy algorithms where the main aim was to recover either the illumination or the reflectance functions, or both of them.

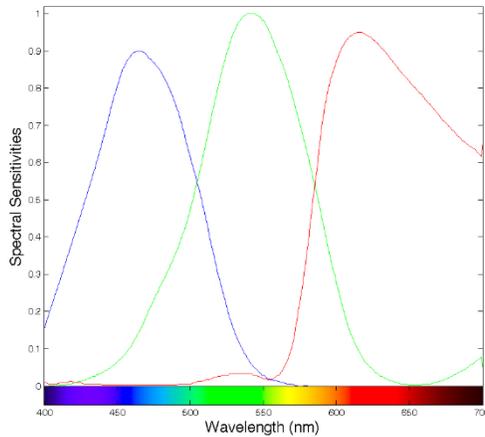


Figure 4-2 Spectral sensitivity characteristics of the Sony ICX098BQ camera (Sony ICX098BQ)

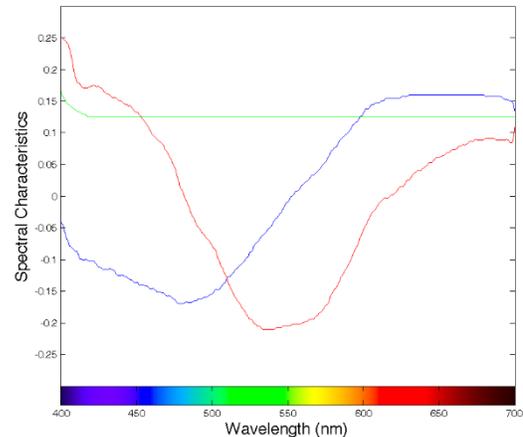


Figure 4-3 Parkkinen's first three basis-functions

The assumptions made by the dichromatic reflection model are:

1. There is a single light source, that can be a point source or an area source,
2. The illumination has a constant PSD across the scene,
3. The amount of illumination can vary across the scene.

For what concerns the surface properties, the model assumes that:

1. The surface is opaque,
2. The surface is not optically active (no fluorescence),
3. The colourant is uniformly distributed

The assumption of illumination is due to only one source of illumination, is limiting the application of such models to scenes where there is a dominant illuminant. The assumptions about the surface are typical for reflection models and not too unrealistic.

4.4 Illuminant estimation

The goal of the illumination estimation is to mimic the capacity of humans to distinguish PSD light from surface reflection; this includes methods relying on linear models (Ho, et al., 1990), neural networks (Cardei, et al., 2002), reliance on highlights and mutual

reflection (Klinker, et al., 1990), and Bayesian and probabilistic approaches (Sapiro, 1999).

The problem of estimating the illuminant is under-determined from a statistical point of view, using a camera that only obtains three spectral samples. In theory, the luminant can be accurately determined if there are enough surfaces on the scene, and then surface reflectivity, in turn, can be measured. For example, the scene is assumed to be metrically impartial colour, with a uniform set everywhere on the average. It is an adaptive reflection that does not overpower a specific colour. The specific spatial colour then matches the light colour.

The Grey-world assumption is one of the earliest developed and is based on the assumption that the surface reflectance spatial average in a scene is achromatic. As the light reflected from an achromatic surface changes equally at all wavelengths, it follows that the colour of the unknown illuminant is the average of the light leaving the scene.

The indicated diagonal transformation is essentially the proportion of the average grey image underneath the canonical to the unknown. The Retinex model makes use of similar assumptions to the Grey-world algorithm. This approach assumes the existence of a white patch from which the chromaticity of the illuminant is perfectly maintained. The maximum channel responses are assumed to arise from a white surface reflectance and are subsequently used in diagonal transform calculation. The method is susceptible to specularities, as they can quickly produce maximum reflection larger than that of a pure white surface. However, it is beneficial to know the presence of specularities first as the illuminating chromaticity is maintained.

Gamut mapping is a method of calculating an unknown illuminant that recovers the best transformation in which the measured gamut is projected into the canonical one. The diagonal transformations are calculated by determining the convex hull of the canonical and unknown light reflections, calculating the set of maps from the unknown Hull onto

those of the canonical Hull, and choosing the best transforming function from the intersection of all the transformative sets. By choosing the diagonal transformation from the feasible set, each variant of the gamut mapping technique can be differentiated. The feasible range is further reduced by restricting the quantity of illumination variation.

A basic correlated colour temperature estimation equation has been derived by McCamy *et al.* (1992) from the chromaticity coordinates x and y of CIE-1931, which helps design sources to simulate CIE-Illuminant.

$$n = \frac{x - x_e}{y_e - y} \quad \text{Equation 4-4}$$

where $x_e = 0.3320$ and $y_e = 0.1858$ Then

$$CCT = 449n^3 + 3525n^2 + 6823.3n + 5520.33 \quad \text{Equation 4-5}$$

This equation was derived from the assumption that CCT can be defined by a polynomial third-order function of the line's reciprocal slope, from specular pixels to the light chromaticity; this approach is based on the fact that the key interest iso-temperature lines for CCTs almost converge to a point on the diagram of chromaticity.

4.5 Colour space

Colour spaces enable different intensity and colour information representations in colour images. The colour spaces RGB, rgb, HSV, Ycbcr, and CIE-Lab will be summarised in the following section to relate to the standard colour space RGB.

4.5.1 RGB

Red, green and blue are the primary colours, and it is now known that their various practical combinations create almost all colour shades. This colour space was the basis for designing CRTs, TV, and computer screens. In this space of colour, most cameras and scanners save their pictures. However, the high correlation between the channels and the

mixing of information about chrominance and luminance renders RGB space a sub-optimal option for colour-based detection systems.

4.5.2 Normalised *rgb*

The normalised *rgb* is a colour space for the normalisation of intensity information, which leads to reduced dependence on the luminance information. Nevertheless, the standardisation property establishes redundancies between the three elements.

$$r = \frac{R}{R + G + B}, \quad g = \frac{G}{R + G + B}, \quad b = \frac{B}{R + G + B} \quad \text{Equation 4-6}$$

For $b = 1 - r - g$, for example, there is no additional detail. For this case, r and g components are called pure colours because of the lack of brightness dependence.

The normalised *rgb* is a colour space for the normalisation of intensity information, which leads to reduced dependence on the luminance information. Nevertheless, the standardisation property establishes redundancies between the three elements. In $b = 1 - r - g$, for example, there is no additional detail. For this case, r , and g components are called pure colours because of the lack of brightness dependence.

4.5.3 HSV

In three quantities, this colour space defines any colour: Hue, Saturation, and Value. It is implemented to fulfil the necessity to define colour properties numerically. Hue determines the dominant pixel colour, Saturation measures the brightness of a pixel, and value correlates with luminance by colour. HSV is connected with RGB non-linearly through the following equations.

$$H = \begin{cases} \theta & \text{if } B \leq G, \\ 2\pi - \theta & \text{if } B > G \end{cases} \quad \text{Equation 4-7}$$

$$\text{Where } \theta = \cos^{-1} \frac{1/2[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}}$$

$$S = 1 - \frac{3}{R + G + B} [\min(R, G, B)] \quad \text{Equation 4-8}$$

$$V = 1/3(R + G + B) \quad \text{Equation 4-9}$$

4.5.4 Ycbcr

For most image compression formats, such as JPEG, MPEG, and TV studios, this colour space is used. Pixel intensity is expressed by Y luminance, measured as a weighted sum of RGB values; the weight matrix converting the RGB pixel value to Ycbcr is indicated in Equation. The pixel information chrominance portion is found in both the c_b and c_r channels. The separation of the luminance and the chrominance components characterises the colour space. It is identical to the colour spaces YIQ and YUV and is linearly related RGB as follows:

$$\begin{bmatrix} Y \\ c_b \\ c_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.711 & -0.587 & -0.114 \\ -0.299 & -0.587 & 0.886 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{Equation 4-10}$$

The elements in the matrix are fixed; this colour space initially suggested by G. Wyszecki has been standardised by the Commission Internationale de L'Eclairage (CIE) to approximate perceptually uniform spatial information. By "perceptually uniform", one means that it was designed to approximate human vision. The L-channel holds pixel intensity/brightness information, during a and b store colour information. The CIE-Lab colour space is non-linearly related to CIE-XYZ. The RGB to Lab conversion is done through a transformation.

$$M: \begin{bmatrix} L \\ a \\ b \end{bmatrix} = M \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Equation 4-11

Standard methods occur when the RGB device coordinates and reference white are defined to determine the transformation M. One such M was used for sRGB, device-dependent colour space. CIE-LUV and CIE-LCH are related spaces of colour.

4.6 Summary

This research has identified the inability of image enhancement techniques to deal with challenging imaging conditions as the critical bottleneck. Image enhancement techniques based solely on statistical methods fail to tackle the practical operational requirements of natural scenes. This research concludes that understanding the underlying physics, which govern the image formation process, is crucial to deal with the variation in imaging parameters. The physics-based approach enables self-adapting to various lighting conditions and better identification of colour. These advantages and benefits mean that image formation models provide an excellent base to get the best chance of achieving image enhancement.

This research used physics-based image formation models to extract meaningful physical features suitable to represent the scene and the objects contained within it. The proposed image enhancement technique exploited physics-based features, namely estimating the illumination characteristics of the scene and surface spectral reflectance based properties of objects. It can work automatically when illumination and scene's characteristics are unknown, which are the main challenges for the applications at hand.

Chapter 5

Proposed Physics-Based Retina-Inspired Models

5.1 Introduction

This chapter proposes two enhanced models that have been modified by the integration of the image formation model; the original models are the Multiscale Adaptation Model (MAM) and the Naturalness preservation model (Joshi & Prakash, 2020) .

Inspired by the retinal network and motivated by the centre-surround operation, a novel image enhancement model is proposed, Ganglion Receptive Field (GRF) model. A detailed explanation for the implementation techniques is provided.

5.2 Multiscale Adaptation Model MAM

The Multiscale Adaptation Model (MAM) is discussed in detail in Section 2.6.6

5.2.1 MAM computational implementation

Figure 5-1 provides a flow chart of each significant step in the computational model. The model processes an input image to encode the perceived contrasts for the chromatic and achromatic channels in their band-pass mechanisms. Then, the model correlates brightness, lightness, colourfulness, chroma, hue and saturation derived from the encoded visual representation.

The Multi-scale adaptation model starts by multiplying each pixel in the image by the spectral responses of the Long, Middle and Short wavelengths sensitive cones (LMS) and the Rods. Where the Cone signals can be calculated as a linear transform of the CIE 1931 XYZ tristimulus values as shown in Equation 5-1,

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \text{Equation 5-1}$$

While Equation 5-2 represents Rod's response for a pixel.

$$R = -0.702X + 1.039 * Y + 0.433Z \quad \text{Equation 5-2}$$

This process results in four images representing the calibrated photoreceptor responses (L , M , S , R), representing *Long*, *Medium*, *Short* and *Rod* signals, respectively.

Spatial decomposition is the second step, carried out on these four images, by applying the Laplacian pyramid (difference-of-Gaussian pyramid) approach proposed by Burt and Adelson (1983). The Gaussian pyramid consists of seven levels. Each level of the Gaussian pyramid represents a low-pass image limited to spatial frequencies, half of those of the next higher level. Each level is then up sampled such that each image is returned to the size of the initial image. Taking the image at each level and subtracting it from the next lower level will give us the differences of Gaussian images; this results in six levels of band-pass images. These images can be thought of as representations of the signals in six band-pass mechanisms in the human visual system. The lowest-level low pass image is retained since it must reconstruct the image for reproduction applications.

Luminance gain control is the third step, where the difference-of-Gaussian images are converted to adapted contrast signals. The gains are set using TVI-like functions representing the increment thresholds of the rod and cone systems; this step represents the growth in response required to allow perceived contrast to increase with luminance level (sub-Weber's law behaviour). The gain functions are given for the cones and the rods in Equation 5-3 and Equation 5-4.

$$G_{cone}(I) = \frac{1}{0.555(I + 1.0)^{0.85}} \quad \text{Equation 5-3}$$

$$G_{rod}(I) = \left[\frac{10}{I^2 + 10} \right] \frac{1}{0.908(I + 0.001)^{0.85}} \quad \text{Equation 5-4}$$

$G_{\text{cone}}(I)$ Represents the Rod or Cone signals used to set the level of adaptation, and $G(I)$ is the gain-control factor. Both Equation 5-3 and Equation 5-4 were derived from matching available psychophysical TVI and brightness matching data.

After applying a multiscale Gaussian Filter, each pixel in a given difference-of-Gaussian image is multiplied by the gain derived from the corresponding pixel in the lower-level low-pass image used in its derivation; this is illustrated in Figure 5-1.

$$ACI_n = G(LP_{n+1})[LP_n - LP_{n+1}] \quad \text{Equation 5-5}$$

Where ACI_n is the adapted contrast image at level n and LP represents the various low-pass images. In this model, the magnitude of the images is a function of luminance level specified by the gain control functions, which allow the prediction of luminance dependant effects. Also, performing the gain control at this point in the model allows proper prediction of chromatic adaptation effects. The luminance gain controls are applied in the same manner to each of the difference-of-Gaussian images for each photoreceptor. Equation 5-3 is used to calculate the gains for each cone, and Equation 5-4 is used for the rods.

Opponent Colour Processing is the following step of the model, where the adapted contrast images for the cones use the transform of (Hunt, 1982) to be transformed into the opponent, as given in Equation 5-6:

$$\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 2.0 & 1.0 & 0.05 \\ 1.0 & -1.09 & 0.09 \\ 0.11 & 0.11 & -0.22 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad \text{Equation 5-6}$$

5.2 Multiscale Adaptation Model MAM

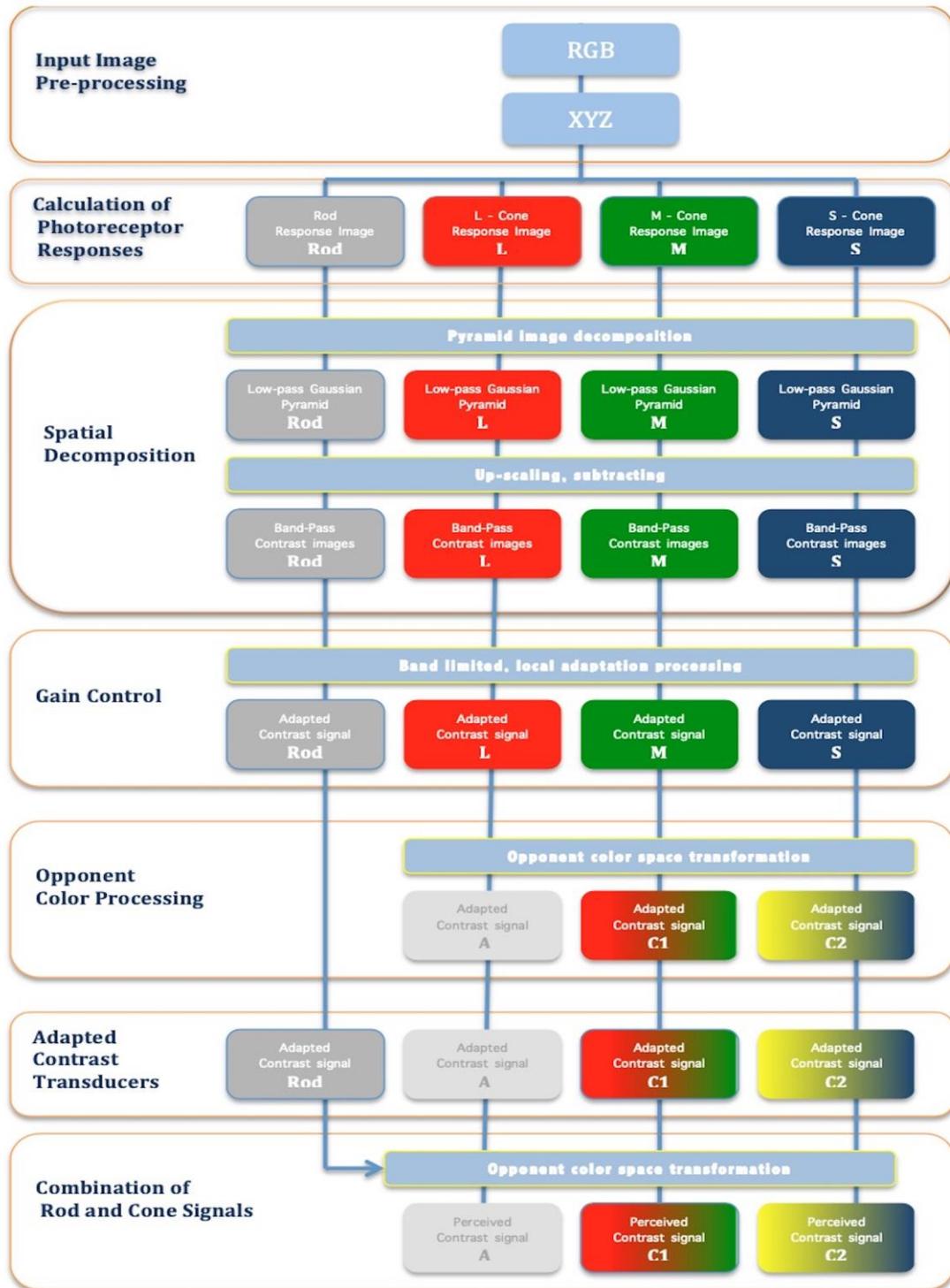


Figure 5-1 Multi-scale adaptation model flow-chart

$$T_{cone,Achromatic}(c) = \begin{cases} 22.4\left(\frac{c}{0.536}\right)^{1/2} & \text{if } c \geq 0.536 \\ 22.4\left(\frac{c}{0.536}\right)^p & \text{otherwise} \end{cases} \quad \text{Equation 5-7}$$

$$T_{cone,Chromatic}(c) = \begin{cases} 22.4\left(\frac{c}{0.176}\right)^{1/2} & \text{if } c \geq 0.0.1766 \\ 22.4\left(\frac{c}{0.176}\right)^p & \text{otherwise} \end{cases} \quad \text{Equation 5-8}$$

$$T_{cone,Achromatic}(c) = \begin{cases} 22.4\left(\frac{c}{0.0335}\right)^{1/2} & \text{if } c \geq 0.0335 \\ 22.4\left(\frac{c}{0.0335}\right)^p & \text{otherwise} \end{cases} \quad \text{Equation 5-9}$$

L, M, and S represent the cone signals and A, C1 and C2 represent the luminance, red-green, and yellow-blue opponent signals. This transformation is necessary to model differences in the spatial processing of luminance and chromatic signals.

The adapted contrast signals are then passed through the next stage, which is the contrast transducer functions; these functions are similar to those described by (Watson & Solomon, 1997) but differ in their ability to be analytically invertible.

Different transducer functions are applied to each spatial frequency mechanism to model the human spatial contrast sensitivity functions. The transducers represent the chromatic channels lower sensitivities and low pass, rather than band-pass nature. The Rod system has a distinct set of transducers to represent its unique spatial characteristics. The transducer functions converge to a common square root form at high contrast levels to properly represent perceived contrast constancy. The contrast transducers used in this model are Equation 5-7, Equation 5-8 for the Cones and Equation 5-9 for the Rods.

Where c represents the adapted contrast signals (ACI 's) and $T(c)$ represents the output of the transducers. The exponent, p , Equation 5-7, Equation 5-8 and Equation 5-9, differs from each spatial frequency mechanism as given in the following table.

Table 5-1 p -values for different spatial frequencies

Peak(cpd)	0.5	1.0	2.0	4.8	8.0	16.0
p for A	1.93	1.35	1.15	1.04	1.15	1.40
p for C ₁ & C ₂	1.93	1.93	2.35	2.85	-	-
p for Rod	3.39	3.39	4.50	7.64	-	-

Transducers' essential function is to set the threshold level where the imperceptible image content for a given set of viewing conditions can be removed; for this to be accomplished, the output of the transducer functions is threshold such that all absolute values less than 1.0 are set to 0.0. This will better replicate the appearance of visual noise at low contrast levels. In addition to establishing thresholds, the transducer functions are used to model saturation of the visual neurons that signal contrast. The transducer functions are limited to maximum values of 50 to simulate the typical physiological dynamic range

The combination of Rod and Cone Signals is the following stage, whereas, in previous stages, it was essential to keep the rod and cone signals separated to appropriately integrate their unique adaptation and spatial vision properties. After the contrast transducers, the rod and cone signals can be combined to produce signals representing the three-dimensional colour appearances of the input image. As it is assumed from the retina neurology that the rods contribute mainly to the luminance signal and thus combine the A signal from the cones with the rod signal, denoted A_{Rod} , to produce a total achromatic signal, Total, using

$$A_{total} = A_{cone} + \frac{A_{Rod}}{7} \quad \text{Equation 5-10}$$

To obtain a total achromatic output that is monotonic with luminance, the model needs to be calibrated by adding a differential weighting of the rod and cone signals to establish the rod and cone gain controls and transducer functions.

At this stage, the model has accomplished a significant level of dynamic-range compression. Both the gain control functions and the nonlinear transducers accomplish

this compression. The model has three channels representing achromatic, red-green, and yellow-blue apparent contrast for six band-pass mechanisms.

Low Pass Image treatment comes before the reconstruction stage to produce maximum dynamic-range compression by multiplying each pixel in the low-pass image by a gain factor derived from the pixel itself. The gain factors are derived using Equation 5-3 and Equation 5-4. The lowest level, low-pass image from the up-sampled Gaussian pyramid must be retained to reconstruct an image from the adapted contrast images that have been passed through the model.

The Low Pass Image treatment tends to mimic the human vision assuming that the eye fixated on each image location and judged them independently. Low-pass transducers that are simple power functions are applied to preserve equal magnitude of signals for the low-pass and band-pass model output for a sinusoidal grating. Those transducers are given in Equation 5-11, Equation 5-12 and Equation 5-13 for the achromatic and chromatic cone signals and rod signals, respectively.

$$T_{LP_cone,Achromatic}(LP) = 30.5 (LP)^{\frac{1}{2}} \quad \text{Equation 5-11}$$

$$T_{LP_cone,Chromatic}(LP) = 53.3 (LP)^{\frac{1}{2}} \quad \text{Equation 5-12}$$

$$T_{LP_rod}(LP) = 122 (LP)^{\frac{1}{2}} \quad \text{Equation 5-13}$$

T is the output of the low-pass transducers, and LP is the pixel values in the adapted low-pass image; before going to the last stage, the model's final outputs are signals in an achromatic two chromatic channels six spatial band-pass mechanisms and a low-pass image.

Image Reconstruction is the last stage model where the model must be inverted; the inversion procedure does not “undo” the processes of the model since the threshold and saturation procedures are accomplished, and the gain control parameters differ for the original scene.

The first step to inverse the transducer functions is Equation 5-7, Equation 5-8, Equation 5-9, Equation 5-10 and Equation 5-11. The AC1C2 signals are then transformed to adapted LMS cone signals using the inverse of the matrix transformation given in

Equation 5-1. At this point, we have adapted contrast signals that have been subjected to the appropriate visibility threshold and saturation by the contrast transducer functions.

The second step is to reverse the gain control process for the viewing conditions of the output display by determining the gain control factors for the mean luminance of the used display device using Equation 5-1. The adapted low-pass images are then divided by the display-mean gain control factors to produce images representing the appropriate LMS cone responses for the display low-pass image. This low-pass display image is then used to reconstruct a full-resolution image from the six adapted contrast signal images.

Gain control factors are calculated for each pixel of the display low-pass image using Equation 5-1. The image is then added to the low-pass image to produce a new low-pass image and then used to calculate gain control factors that are applied to the next level. Again, the resulting image is added to the new low-pass image to generate yet another low-pass image that incorporates the information from the two previous levels. The process is repeated at each level until all spatial frequency mechanisms have been scaled to the display and added to the output image.

The MAM model mimics the retinal photoreceptors by converting the *RGB* pixel values to *L*, *M*, *S* and *Rod values*. In the second step, the Laplacian pyramid is recruited for the spatial decomposition, which mimics the concept of receptive field in the retina, where it emphasizes the relationship between the pixel and its neighbours, In the third step it mimics the contrast response of the photoreceptors concerning the light intensity the equations used are derived from matching available psychophysical TVI and brightness matching data.

The next stage converts the *L*, *M*, *S* and rods to *A*, *C₁*, and *C₂*, where it corresponds to *K-pathway*, *M-pathway*, and *P-pathway* in the retina “the three main parallel paths that carry image signals from the retina to the brain), this transformation is necessary to model differences in the spatial processing of luminance and chromatic signals.

The adapted contrast signals are then passed through the next stage, which is the contrast transducer functions. Those functions model the human spatial contrast sensitivity functions; the transducer functions are used to model the saturation of the visual neurons

that signal contrast; after the contrast transducers, the rod and cone signals can be combined to produce signals representing the three-dimensional colour appearances of the input image. In retina neurology, the rods contribute mainly to the luminance signal and combine the signal from the cones with the rod signal, denoted, to produce a total achromatic signal; the model is calibrated by adding a differential weighting of the rod and cone signals. In the last stage, where the model must be inverted, the inversion procedure does not “undo” the processes of the model since the threshold and saturation procedures are accomplished.

5.2.2 Physics-Based-Multiscale Adaptation Model (MAM_p): Enhanced model

The new enhanced proposed computational model integrates the physics-based image formation model with the retina-inspired multi-scale adaptation model. Where the Surface Spectral Reflectance, SSR, is estimated based on the dichromatic colour reflectance model. This approach uses image formation models to computationally estimate, from the camera output, a consistent physics-based colour descriptor of the spectral reflectance of surfaces visible in the image, and then to convert the full-spectrum reflectance to Long (L), Medium (M), Short (S) Cone responses and Rod (R) response.

This method represents a new approach to image enhancement, using explicit hypotheses about the physics that create images. This step makes the proposed different from the original model, the Multi-scale Adaption Model (Pattanaik *et al.*, 1998), in that it relies on models, which can represent broad classes of surface materials. It uses the pre-trained linear SSR models to represent the SSR of the objects in the scene. For what concerns the surface properties, this model assumes that: the surface is opaque; not optically active (no fluorescence), and the colourant is uniformly distributed.

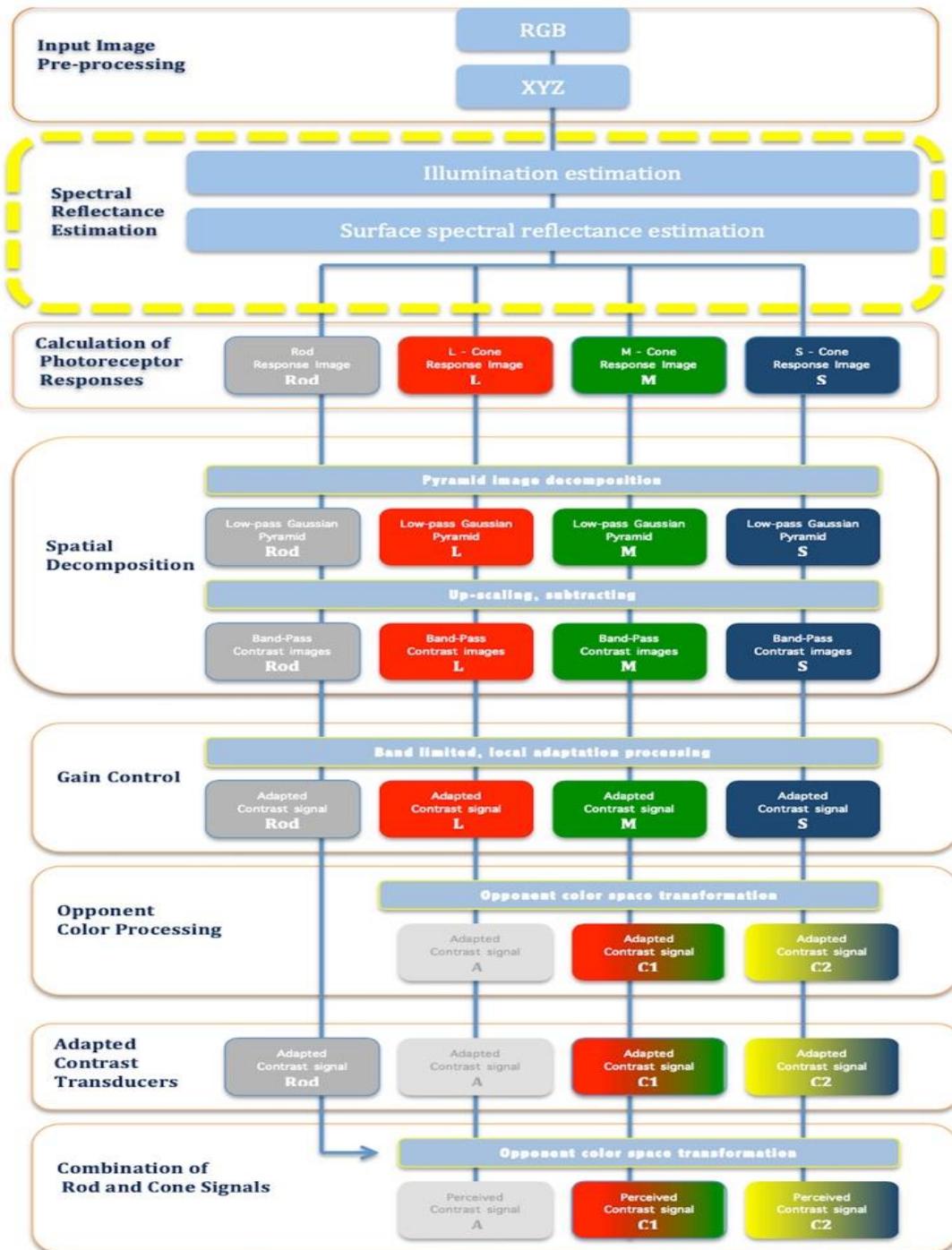


Figure 5-2 In the new enhanced model, the cones and rod spectral characteristics are multiplied by the SSR, and the result is integrated to obtain the cone and rod responses, four images representing the calibrated photoreceptor responses (L, M, S, R)

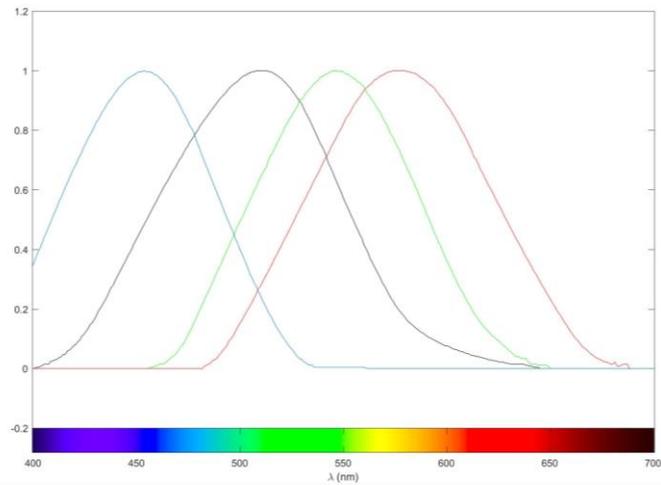


Figure 5-3 Long(red), Medium (Green) and Short(Blue) Cones Sensivities and Rod sensitivities(Black)

The Cones and Rod spectral characteristics are multiplied by the SSR, and the result is integrated to obtain the cone and rod responses. The result is represented by four images representing the calibrated photoreceptor responses (L , M , S , R), where they represent Long, Medium, Short and Rod signals, respectively, shown in Figure 5-3. Those four signals will be subjected to spatial processing.

The new enhanced model is identical to the original multi-scale adaption model from stage two to the last stage. Which consists of image processing, spatial decomposition, gain control, opponent colour processing, orientation filtering, contrast transducers, thresholding, a combination of rod and cone signals, and treatment of the low pass image, as the flow chart Figure 5-2.

Method 1: Physics-Based Multi-Scale Adaptation

Input: RGB test image

Output: RGB restored image

Step1: {Illumination Estimation} Apply Mccamy formula to The Correlated Colour Temperature (CCT)

Step2: {Spectral Reflectance Estimation} Calculate the spectral reflectance weights using the CCT and the input image

Step3: {Spectral Components to LMS-Rod Conversion} Transform the Parkkinnen weights to cones (Long, Medium and Short cone responses) and Rod

Step4: {Laplacian Pyramid (difference-of-Gaussian pyramid)} Apply the Gaussian pyramid (seven levels).

Step5: {Opponent Colour} Calculate The adapted contrast images for the cones use the transform of Hunt (1995) to be transformed into opponent, where L , M and S represent the cone signals and A , C_1 and C_2 represent the luminance, red-green, and yellow-blue opponent signals respectively. The contrast transducer functions are then applied to the adapted contrast signals

Step6: {Rod and Cone Signals Re-combination} The rod and cone signals are combined by adding a differential weighting of the rod and cone signals

Step7: {Low Pass Image treatment} Multiply each pixel in the low-pass image by a gain factor derived from the pixel itself.

Step8: {Image Reconstruction} Inverse the transducer functions, then reverse the gain control process, then divide the adapted low pass image by the reversed gain control factors, This new low-pass image is then used to begin the process of reconstruction of a full-resolution image from the six adapted contrast signal images by adding this new image to layer 1

5.3 Naturalness preservation Model

The naturalness preservation model is discussed in detail in section 2.6.7

5.3.1 Computational Model

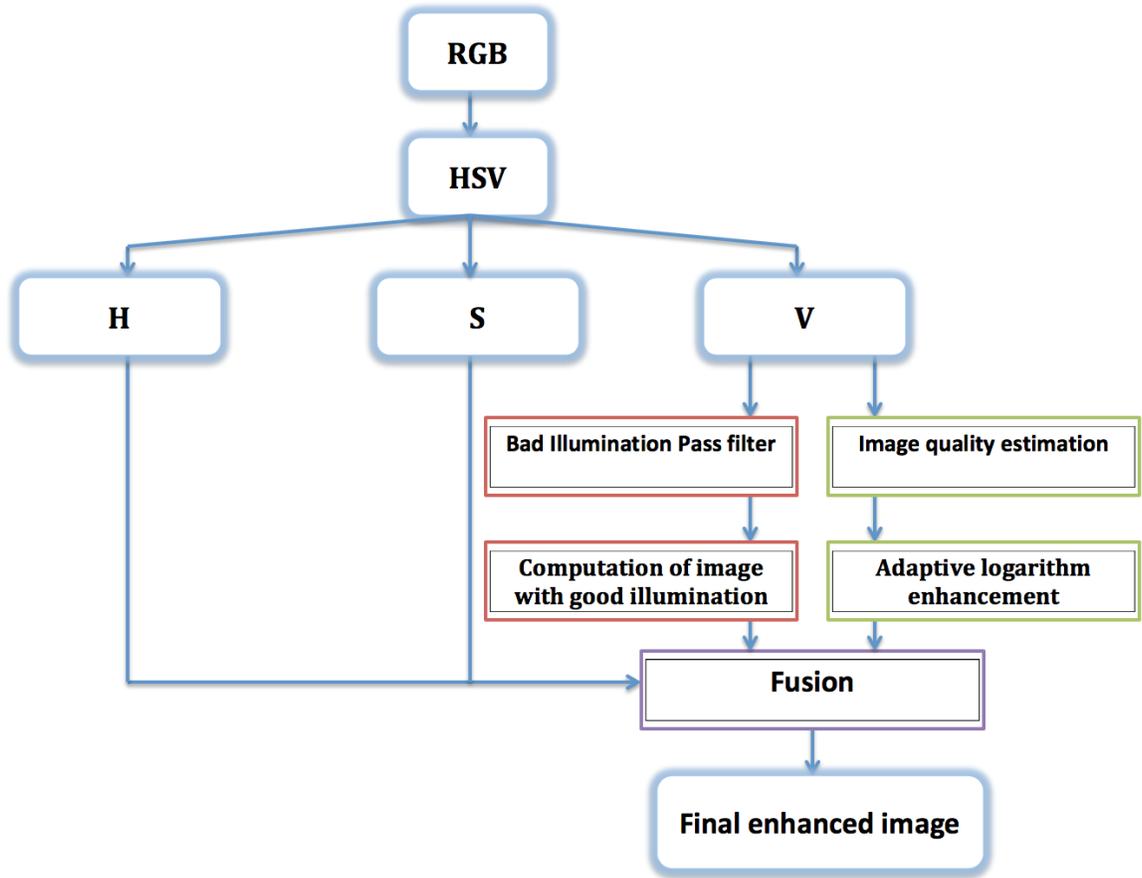


Figure 5-4 Framework of the naturalness preservation model technique

Step One: Extraction of details (reflectance) and good illumination

Bad illumination pass filter (BIPF)

The basic idea in BIPF designing is that due to the effect of its adjacent neighbouring pixels, a pixel gets poorly illuminated. The proposed filter considers an $n \times n$ window around each pixel and calculates the presence of bad or good lighting on it. BIPF is defined as the weighted average of adjacent pixels with the weight associated with the frequency

of bad pixels of illumination and is given as:

$$I_f(I(x_c, y_c)) = \sum_{(i,j) \in win} [f(i,j) \cdot I(i,j)] \cdot F(I(i,j)) \quad \text{Equation 5-14}$$

where win is the window centered at pixel location (x_c, y_c) and, $F(I(i,j))$ and $f(i,j)$ are two functions. Function F is used to decide if a pixel (x,y) of the image I is poorly illuminated or not and is given as:

$$F(I(x,y)) = \begin{cases} 1, & \text{if } I(x,y) < T_1 \text{ or } I(x,y) > T_2 \\ 0, & \text{otherwise} \end{cases} \quad \text{Equation 5-15}$$

T_1 and T_2 are lower and upper thresholds to determine, respectively, whether or not a pixel is poorly illuminated. A pixel is considered very dark if its intensity is below the T_1 threshold and if it is higher than the T_2 the threshold is considered very bright. Both dark and bright pixels (black and bright spots) are the source of poor lighting. Additionally, Function f is used to count the number of poorly illuminated pixels around pixels (I_j) and is defined as:

$$f(i,j) = \sum_{(k,l) \in win} F(k,l) \quad \text{Equation 5-16}$$

In Equation 5-4, I_f is the filtered image in which bad illumination is located. This image is again filtered using Gaussian function G with standard deviation σ (Retinex theory) in order to concentrate on illumination part as follows fully:

$$I_{fG} = I_f * G \quad \text{Equation 5-17}$$

Computation of image with good illumination (I_{GI})

$$I_{GI} = \log(I) - \log(I_{fG}) \quad \text{Equation 5-18}$$

where I is an input image and I_{fG} is the BIPF filtered image is containing only bad illumination.

Step Two: Insertion of brightness

The method incorporates the correct amount of light in dark images to enhance them and achieve pleasing perceptual consistency. This approach demands the deteriorated image characteristics. The proposed technique first estimates the illumination-based quality of the image and uses this quality information to inject the required amount of brightness into the non-uniform illuminate images.

Bad illumination-based quality estimation

The proposed technique computes quality based on illumination by coding each pixel based on its neighbours. Code at pixel location $c = (x_c, y_c)$ (provided I_c, T_1) is defined as:

$$\text{code}(x_c, y_c) = \sum_{(x_p, y_p) \in \text{win}} \frac{E(I_p - I_c)}{(N - 1)} \quad \text{Equation 5-19}$$

win is the window centered at pixel (x_c, y_c) and N is the number of pixels falling under the window. I_p and I_c define intensity values at pth neighboring pixel and central pixel (for which code is being computed), respectively. T_1 is a threshold, used to detect dark regions in the image. Function E (D) is defined as follows:

$$E(D) = \begin{cases} 1, & \text{if } \text{abs}(D) > \varepsilon \\ 0, & \text{otherwise} \end{cases} \quad \text{Equation 5-20}$$

Where ε is a difference parameter. The value of E (D) for pth neighbour is assumed to be 1 if difference in intensity I_c of the centre pixel c and intensity I_p of its pth neighbour is greater than threshold, else it is assumed to be 0. In the dark regions of the image, difference in intensities of neighbouring pixels is found to be very less.

Equation 2 gives code value (code (x_c, y_c)) between 0 to 1 for a pixel with intensity value below the threshold T_1 (dark intensity pixel), whereas pixels with intensity value above T_1 are coded as 1. Overall illumination-based quality is obtained by computing the sum of all codes,

$$Q_I = \frac{\sum_{i=1}^M \sum_{j=1}^N \text{code}(x_i, y_i)}{W * H} \quad \text{Equation 5-21}$$

W and H are the width and the height of the input image, Q_I helps in the insertion of the required amount of brightness in the images. This quality estimation is based on one threshold T_1 (lower threshold). However, in order to achieve proper enhancement, both dark and bright regions should not be present in the final enhanced images. Hence, quality estimation of final enhanced images must include both parameters T_1 and T_2 , for estimating quality based on dark and bright regions. For an effective quality assessment of the final enhanced images, we also code bright pixels along with dark pixels in the image and for this, an upper threshold T_2 is included. With new constraint, a pixel is coded using Equation 5-21 provided $I_c(T_1)$ and $I_c(T_2)$. This means an image which has either very bright or very dark regions is considered having bad illumination and obtains a low-quality score. The quality thus estimated using thresholds T_1 and T_2 is the quality for final enhanced image.

Adaptive logarithm enhancement

The adaptive logarithm enhances the image based on input image quality score. This adaptive logarithm inserts the required amount of brightness in images with non-uniform illumination. The logarithm is defined as,

$$I_{ALE} = \log(I + e) \quad \text{Equation 5-22}$$

I is the input image, and e is an adaptive parameter of brightness varying from 0 to I . It determines how much light to introduce into an image in order to obtain pleasing naturalness. The low e value results in high dynamic range compression (more addition of brightness) and vice versa.

For example, if the image input is dark, e is set near 0 to insert higher brightness; it is observed that we must carefully choose the value of the parameter brightness e . The selection of e plays a significant role, and a proper enhancement can only be accomplished

by choosing the correct value of e . The value e is set accordingly.

Q_I is the estimated quality score. We classify the images into 4 categories, normal images ($Q_I \geq 0.85$), shadow images ($Q_I \geq 0.60$ & $Q_I < 0.85$), dark images ($Q_I \geq 0.40$ & $Q_I < 0.60$) and very dark images ($Q_I < 0.40$) and assign a value e to each class.

Computation of final enhanced image

The proposed technique generates two enhanced images, one with good illumination (I_{GI}) and image (I_{ALE}) obtained after the use of adaptive logarithm enhancement. Image (I_{GI}) contains details (reflectance) and good illumination, whereas the image obtained after adaptive logarithm enhancement contains proper amount of brightness. Final enhanced image (I_E) is obtained by combining images I_{GI} and I_{ALE} as follows.

$$I_E = I_{GI} + I_{ALE} \quad \text{Equation 5-23}$$

The naturalness technique primarily performs two parallel tasks. It generates an improved intermediate image in the first task with the aid of BIPF, which enhances by identifying image areas with information and good illumination. In the second phase, using image quality and applying the ALE logarithm calculates another enhanced image from the input image. Finally, to get the final enhanced image, output images of these two tasks are fused into one.

5.3.2 Physics-Based – Naturalness preservation model: Enhanced model

The new enhanced model integrates the physics-based image formation model with the naturalness preservation model, where the spectral surface reflectance is estimated based on the dichromatic colour reflectance model. This approach uses image formation models to estimate the camera output computationally, a consistent physics-based colour descriptor of the spectral reflectance surface visible in the image. The weights will represent the colour space in the Parkinson basis function in the enhanced model. Instead of presenting the image in the HSV (*Hue, Saturation, Value*) colour space, it is represented in W_3, W_2, W_1 , respectively.

Method 2: Physics-based Bad Illumination Filter

Input: RGB test image

Output: RGB restored image

Step1: {Illumination Estimation} Apply Mccamy formula to The Correlated Colour Temperature (CCT)

Step2: {Spectral Reflectance Estimation} Calculate the spectral reflectance weights using the CCT and the input image

Step3: {Allocation of Bad Illumination} Apply Bad Illumination Pass Filter (BIPF) on $W1$ **Step4: {Computation of image with good illumination}** Eliminating areas with bad illumination.

Step5: {Estimate the quality of the image based on illumination} Compute quality by coding each pixel based on its neighbours.

Step6: {Utilize the Adaptive Logarithm Enhancement (ALE)} Final enhanced image is obtained by combining the image of good illumination and the image obtained after using ALE.

5.4 Ganglion Receptive Field Model (GRF): Novel model

The proposed retina-inspired image enhancement computational model GRF is an attempt to represent the retinal network. The model considers the four layers in the retina, the Bipolar cells, the Horizontal cells, the Amacrine cells and the Ganglion cells and the interaction between these different layers to mimic the retinal neural process as shown in Figure 5-6. This computational model will be evaluated in Chapter 6 by introducing various real-world images to prove that a better representation of the retinal network could yield results comparable to or even better than the state-of-the-art methods.

The first stage in the model is the image presentation. To have a physical representation of colour, the Dichromatic model was applied. It uses the physics-based features in the image (e.g. Spectral Power Distribution of the dominant illuminate in the scene and the estimation of the Surface Spectral Reflectance of the objects contained in the image), so

the image will be presented in the full spectrum instead of three colour samples used by other colour spaces.

The second stage is the image filtering or the spatial decomposition, instead of considering only the interactions between each pixel and its surroundings within a single colour layer. The proposed framework introduces the interaction between the colour layers (red, green, blue, yellow [add the red and green], black, white) to mimic different retinal layers, targeting the interconnectivity between the Bipolar receptive field and the Ganglion receptive field. Then the Ganglion cells operation is presented using a tone mapping function given with a new log-power-exponential function.

This approach is novel and differs from all the previous work, as it makes use of the pre-trained linear Surface Spectral Reflectance (SSR) models, i.e. Parkkinen basis functions, to represent the SSR of the objects in the scene. The method proposes an image enhancement method that restores an image with better luminance and contrast.

The rationale behind this approach is that the Ganglion receptive field concept can better present a basis for image enhancement and that the physics that govern the image plays a vital role in the way the image features are introduced to an algorithm. A potential computational representation for the Ganglion cell functionality is the log-power-exponential function. This function can map over-saturated and under-saturated pixels back to a normal range. The computational model is discussed in the next section.

5.4 Ganglion Receptive Field Model (GRF): Novel model

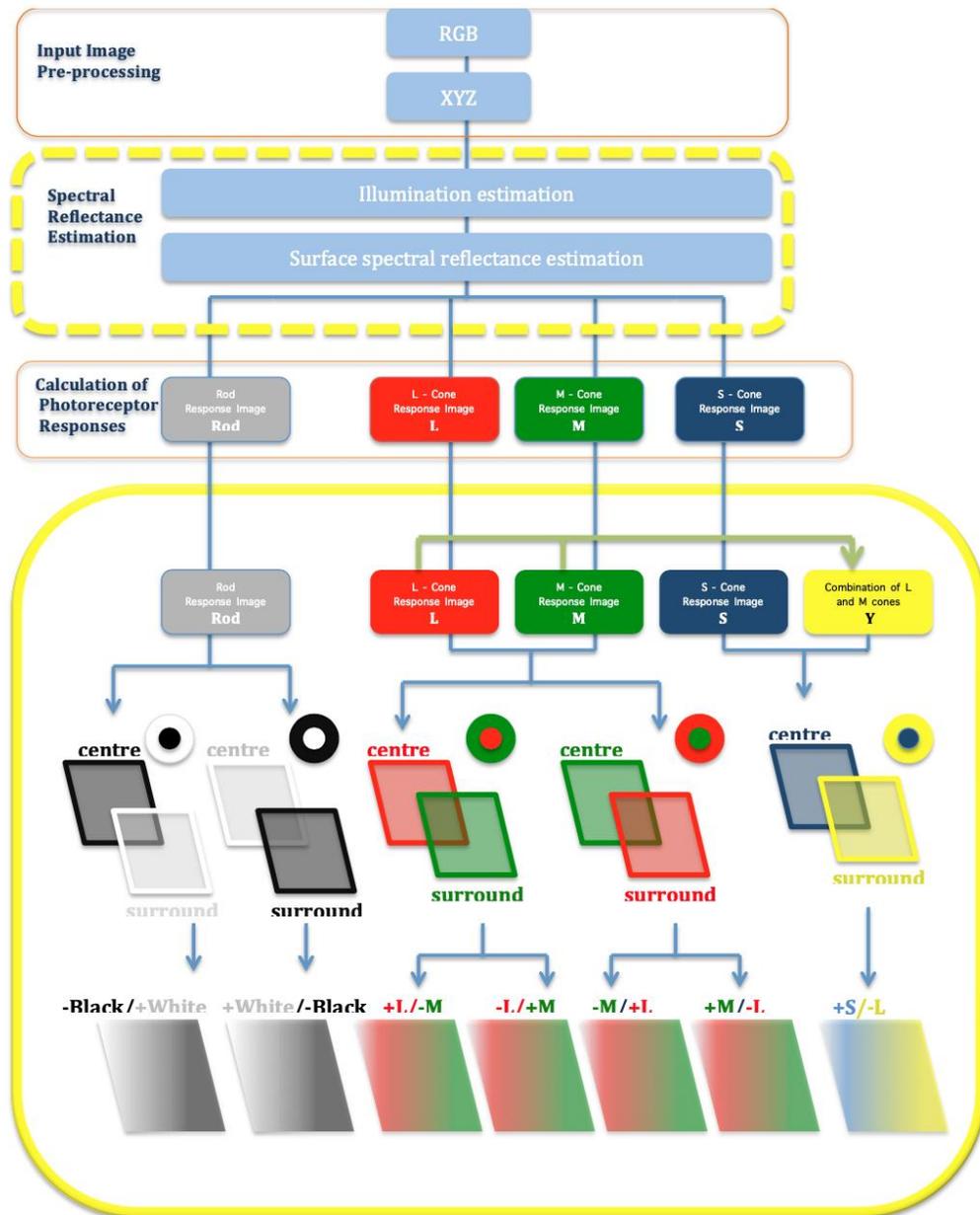


Figure 5-6 Proposed method block diagram

5.4.1 The computational model

Surface Spectral Reflectance (SSR) calculations

Starting with the dichromatic model, presented in Section 4.3, the Surface Spectral Reflectance is calculated using the Parkinen basis functions presented in Figure 4-3. The characteristic of the camera sensors shown in Figure 4-2 is used to represent the camera sensitivities. The SSR recovery module aim is to calculate the weights of Parkkinen basis functions. To obtain the SSR of the object, the model could be rewritten now as:

$$\begin{aligned} R &= w_d \int E(\lambda).S(\lambda).Q_R(\lambda) d\lambda + \tilde{w}_s \int E(\lambda).Q_R(\lambda) d\lambda \\ G &= w_d \int E(\lambda).S(\lambda).Q_G(\lambda) d\lambda + \tilde{w}_s \int E(\lambda).Q_G(\lambda) d\lambda \\ B &= w_d \int E(\lambda).S(\lambda).Q_B(\lambda) d\lambda + \tilde{w}_s \int E(\lambda).Q_B(\lambda) d\lambda \end{aligned} \quad \text{Equation 5-24}$$

Where λ is the wavelength (the visible range is approximately from 400nm to 700nm), $E(\lambda)$ is the illuminant spectral intensity of radiation, $S(\lambda)$ is the spectral reflectivity of the object, $Q_R(\lambda)$ is the red camera sensor spectral response characteristics, $Q_G(\lambda)$ is the green camera sensor spectral response characteristics, $Q_B(\lambda)$ is the camera sensor spectral response characteristics, w_d The geometrical parameter for the diffuse reflection component, The geometrical \tilde{w}_s parameter for the specular reflection component

The SSR of an object is a measure of the amount of light reflected by an object, or radiance, relative to the amount of incident light shone on the object, or irradiance, and is indicative of the reflectance or intrinsic brightness of an object.

The SSR of an object can represent a signature of an object and can be used to physically represent the objects in the scene. The SSR of an object is decomposed into diffuse (Lambertian, body) and specular (surface) components. This work accounts for diffuse only materials. In order to extract the SSR from a colour camera output, an illumination estimation algorithm is required.

The illumination estimation stage aims to calculate the correlated colour temperature (CCT) as an illumination parameter of the illuminant. McCamy's method is used in this invention, which is able to estimate the CCT from CIE 1931 chromaticity coordinates x and y . McCamy's method has a maximum absolute error of less than 2° K for colour temperatures ranging from 2,856 to 6,500° K (corresponding to CIE illuminants A through D65). This method proves useful for implementation in real-time applications. It was derived from the assumption that CCT may be represented by a third-order polynomial function of the reciprocal of the slope of the line from specular pixels to the chromaticity of the light and the fact that the isothermperature lines for CCTs of principal interest nearly converge towards a point on the chromaticity diagram.

The illumination estimation algorithm starts by segmenting areas in the image which represent high specularities (highlights); a value for the illuminant CCT is initially set depending on the available knowledge of the environment. The spectral illumination characteristic is then calculated using Plank's formula:

$$E(\lambda) = c_1 \cdot \lambda^{-5} \cdot (e^{\frac{c_2}{\lambda \cdot T}} - 1)^{-1} \cong c_1 \cdot \lambda^{-5} \cdot e^{-\frac{c_2}{\lambda \cdot T}} \quad \text{Equation 5-25}$$

Where $E(\lambda)$ is the illuminant spectral power distribution in Watts / m² / wavelength interval; λ is the wavelength in m; c_1 is a constant in Watt.m²; c_2 is a constant in m.°K; T is the colour temperature of the illuminant, in °K. This work assumes the colour temperature is $T = 4400$ °K for Fluorescent F2.

In order to build a computational physical model, the exact model is used as in Equation 5-14, we have then used 3 basis functions of Parkkinen ($n = 3$), as shown in Figure 4-3, so the reflectance function is represented with

$$S(\lambda) \approx \sum_{i=1}^n w_i \phi_i(\lambda) \quad \text{Equation 5-26}$$

From the manufacturer data sheet, the actual characteristics of the camera sensors are

obtained, as shown in Figure 4-2.

The aim now is to calculate the weights of Parkkinen basis functions, to obtain the spectral reflectivity of the object. Equation 5-27 to Equation 5-35 present a derivation for the linear basis functions weighs. The model could be written now as:

$$\begin{aligned}
 R &= w_d w_1 \int E(\lambda) \cdot \varphi_1(\lambda) \cdot Q_R(\lambda) d\lambda + w_d w_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_R(\lambda) d\lambda + w_d w_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_R(\lambda) d\lambda \\
 G &= w_d w_1 \int E(\lambda) \cdot \varphi_1(\lambda) \cdot Q_G(\lambda) d\lambda + w_d w_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_G(\lambda) d\lambda + w_d w_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_G(\lambda) d\lambda \\
 B &= w_d w_1 \int E(\lambda) \cdot \varphi_1(\lambda) \cdot Q_B(\lambda) d\lambda + w_d w_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_B(\lambda) d\lambda + w_d w_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_B(\lambda) d\lambda
 \end{aligned}$$

Equation 5-27

Or

$$\begin{aligned}
 R &= \tilde{w}_1 \int E(\lambda) \cdot \varphi_1(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_R(\lambda) d\lambda \\
 G &= \tilde{w}_1 \int E(\lambda) \cdot \varphi_1(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_G(\lambda) d\lambda \\
 B &= \tilde{w}_1 \int E(\lambda) \cdot \varphi_1(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_B(\lambda) d\lambda
 \end{aligned}$$

Equation 5-28

Where

$$\tilde{w}_1 = w_d w_1, \quad \tilde{w}_2 = w_d w_2 \quad \text{and} \quad \tilde{w}_3 = w_d w_3$$

Equation 5-29

Knowing that the first basis function of Parkkinen is constant so $\varphi_1(\lambda) = k_\phi$, so

$$\begin{aligned}
 R &= \tilde{w}_1 \cdot k_\phi \int E(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_R(\lambda) d\lambda \\
 G &= \tilde{w}_1 \cdot k_\phi \int E(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_G(\lambda) d\lambda \\
 B &= \tilde{w}_1 \cdot k_\phi \int E(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_s \int E(\lambda) Q_B(\lambda) d\lambda
 \end{aligned}$$

Equation 5-30

which then could be merged with the specular component to give:

$$\begin{aligned}
 R &= (\tilde{w}_1 \cdot k_\phi + \tilde{w}_s) \int E(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_R(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_R(\lambda) d\lambda \\
 G &= (\tilde{w}_1 \cdot k_\phi + \tilde{w}_s) \int E(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_G(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_G(\lambda) d\lambda \\
 B &= (\tilde{w}_1 \cdot k_\phi + \tilde{w}_s) \int E(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_2 \int E(\lambda) \cdot \varphi_2(\lambda) \cdot Q_B(\lambda) d\lambda + \tilde{w}_3 \int E(\lambda) \cdot \varphi_3(\lambda) \cdot Q_B(\lambda) d\lambda
 \end{aligned}$$

Equation 5-31

By taking

$$\begin{aligned}
 X_i &= \int E(\lambda) \cdot \varphi_i(\lambda) \cdot Q_R(\lambda) d\lambda \\
 Y_i &= \int E(\lambda) \cdot \varphi_i(\lambda) \cdot Q_G(\lambda) d\lambda \\
 Z_i &= \int E(\lambda) \cdot \varphi_i(\lambda) \cdot Q_B(\lambda) d\lambda
 \end{aligned}$$

Equation 5-32

These integrations are calculated to obtain the transformation matrix

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{bmatrix} \begin{bmatrix} \tilde{w}_1 \cdot k_\phi + \tilde{w}_s \\ \tilde{w}_2 \\ \tilde{w}_3 \end{bmatrix} \quad \text{Equation 5-33}$$

Now the weights of the basis functions can be obtained from RGB values by:

$$\begin{bmatrix} \tilde{w}_1 \cdot k_\phi + \tilde{w}_s \\ \tilde{w}_2 \\ \tilde{w}_3 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{bmatrix}^{-1} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{Equation 5-34}$$

Or

$$\begin{bmatrix} w_d \cdot w_1 \cdot k_\phi + w_s \cdot k_s \\ w_d \cdot w_2 \\ w_d \cdot w_3 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{bmatrix}^{-1} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{Equation 5-35}$$

As a particular case for diffuse-only reflection $w_s = 0$ then:

$$\begin{bmatrix} \ddot{w}_1 \\ \ddot{w}_2 \\ \ddot{w}_3 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{bmatrix}^{-1} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{Equation 5-36}$$

where $\ddot{w}_1 = w_d \cdot w_1 \cdot k_\phi$, $\ddot{w}_2 = w_d \cdot w_2$ and $\ddot{w}_3 = w_d \cdot w_3$

Using the transformation discussed in the previous section, an RGB image could be now represented by basis function weights used in the expression of its spectral reflectivity.

Figure 5-7 shows an example of a reconstructed SSR.

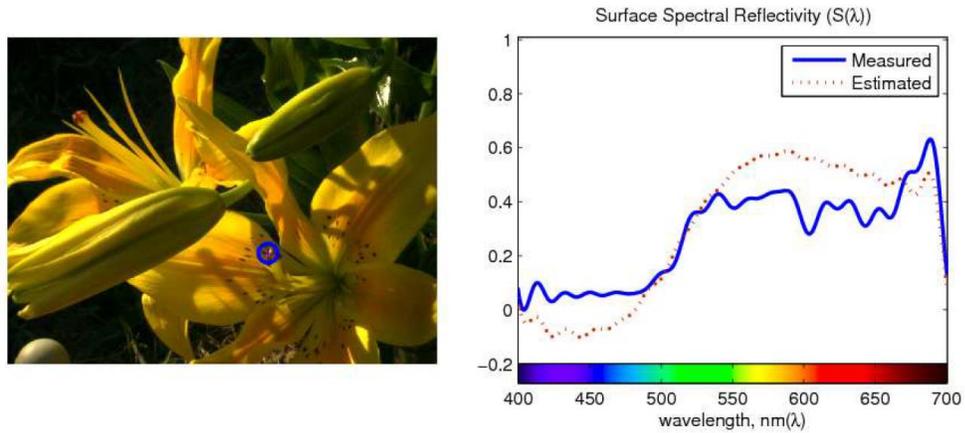


Figure 5-7 Reconstructed SSR

The SSR weights are then converted to LMS by multiplying the LMS spectral sensitivity with the SSR and integrating as shown in the following equations:

$$\begin{aligned}
 L &= \int_{400nm}^{700nm} SSR(\lambda)L(\lambda)d\lambda \\
 M &= \int_{400nm}^{700nm} SSR(\lambda)M(\lambda)d\lambda \\
 S &= \int_{400nm}^{700nm} SSR(\lambda)S(\lambda)d\lambda
 \end{aligned}
 \tag{Equation 5-37}$$

Spatial decomposition

Horizontal Cells

The L , M , S pixels are averaged over a window of 2×2 , while all the Rod pixels are averaged over a window of 30×30 .

Bipolar Cells

The output of the Horizontal cells is subtracted from the relevant center channels as follows:

$$L_{centre}M_{surround}On - Off = L_{centre} - M_{surround}
 \tag{Equation 5-38}$$

$$Rod_{centre}Rod_{surround}On - Off = 1 - Rod_{centre} - Rod_{surround}$$

Where Y is the Yellow component = L+M,

In the ON-centre Red / OFF-surround Green (+L/-M) receptive field, there will be an operation between each pixel from the L layer and the two surrounding pixels from the M layer; the M layer mask represents Horizontal Cells' function in the retina. the output of this stage is seven layers, this stage represents the Bipolar cell function in the retina.

Amacrine Cells

The output of the bipolar cells is averaged over a window of 4x4,

Ganglion Cells

To better mimic the centre-surround retinal receptive field concept, different photoreceptors' outputs will be combined. From the output of the Bipolar and American cells. seven layers of outputs will be the input for the Ganglion cells, the receptive field concept will be applied again but instead of being a subtraction between the L and M , and between the Yellow and blue as in the Bipolar cells , it will be between the seven new layers, it will be between L/M on/off layer as a centre and the average of the M/L off/on layer as the surround, and so on with rest of the layers. This thesis presents the Ganglion cells as a log-power-exponential map as follows:

$$f(x, y) = (2 * I - \log(I^2 \cdot 9 + 1))^{(1 - \exp(K)/2)} \quad \text{Equation 5-39}$$

Method 3: Ganglion Receptive Field (GRF)

Input: RGB test image

Output: RGB restored image

Step1: {Illumination Estimation} Apply Mccamy formula to calculateThe Correlated Colour Temperature (CCT)

Step2: {Spectral Reflectance Estimation} Calculate the spectral reflectance weights using the CCT and the input image

Step3: {Spectral Components to LMS-Rod Conversion} Transform the Parkkinnen weights to cones (Long, Medium and Short cone responses) and Rod

Step5: {Horizontal Cells} Average the cells in the L, M, S and Rod (surround pixels), for L, M and S the surround window = 2, for the Rods the surround window = 30,

Step6: {Bipolar Cells} Subtract each centre pixel from the average surround to calculate $L_{\text{centre}}-M_{\text{Surround}}$ On-Off, $L_{\text{centre}}-M_{\text{Surround}}$ Off-On, $M_{\text{centre}}-L_{\text{Surround}}$ On-Off, $S_{\text{centre}}-Y_{\text{Surround}}$ On-Off, $Rod_{\text{centre}}-Rod_{\text{Surround}}$ On-Off, and $Rod_{\text{centre}}-Rod_{\text{Surround}}$ Off-On

Step7: {Amacrine Cells} Average the cells in the all the components generated from the bipolar cells, using a surround window =4 and average the centre cells using centre window = 2,

Step8: {Ganglion Cells} Enhance the input image by applying the log-power-exp map to the L, M and S cone signals using a gain factor derived from the output of Amacrine cells.



Figure 5-8 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods

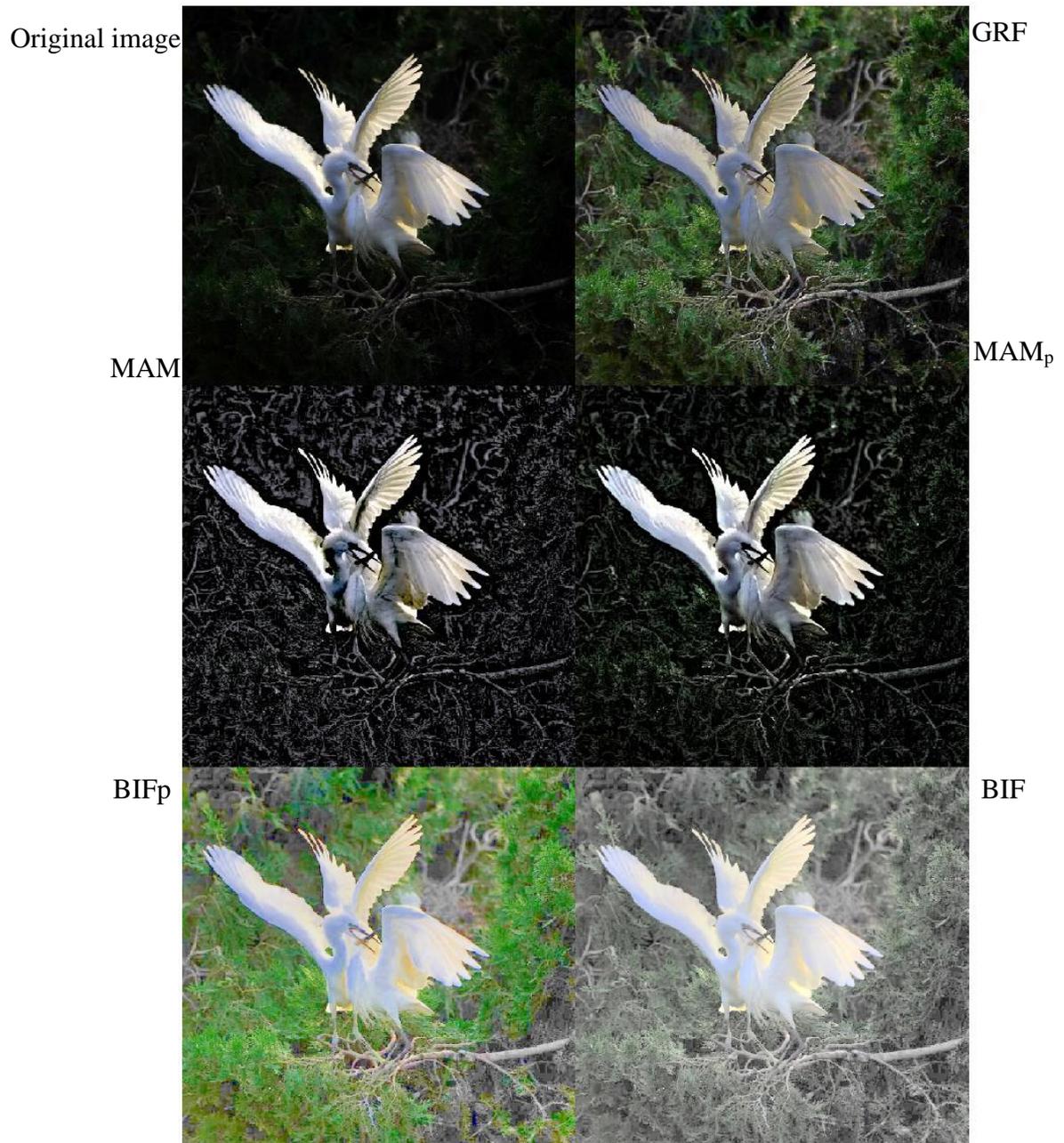


Figure 5-9 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods

5.5 Summary

The challenge of the image enhancement approach based on surface spectral reflectance arises from the new idea of processing the full spectrum of the surface spectral reflectance instead of the three samples used by other colour spaces. The spectral representation uses a linear model consisting of several basis functions (pre-trained from a set of materials) and weights (calculated for the object under investigation). Numerical estimation of the physics-based model for image formation and the real-time transformation from the camera output to the physical parameters is carried out. The assumptions are that only diffuse reflections are considered, and the existence of a dominant illuminant is applicable. This Chapter introduced two newly modified state-of-the-art image enhancement methods through their integration with image formation models. The new algorithms have been developed based on a physical meaning of image formation and which use advances in colour constancy techniques. In particular, physics-based features (e.g. Spectral Power Distribution of the dominant illuminate in the scene and the Surface Spectral Reflectance of the objects contained in the image) are estimated and used as inputs to the multiscale model for adaptation.

This Chapter has successfully proposed a novel Ganglion receptive field computational model for image enhancement. Instead of considering only the interactions between each pixel and its surroundings within a single colour layer, the proposed framework introduces the interaction between different colour layers to mimic the retinal neural process; to better mimic the centre-surround retinal receptive field concept, different photoreceptors' outputs are combined. The Ganglion cells operation is presented using a tone mapping function presented with a new log-power-exponential function.

Chapter 6

Performance Evaluation

6.1 Introduction

Image enhancement algorithms are crucial for a variety of computer vision applications. However, the targeted performance of such techniques differ depending on the application at hand and the characteristics of the captured scene. Contrast enhancement techniques aim to increase the contrast without losing visual information, while tone mapping application converts real-world luminance into displayable luminance to preserve naturalness while enhancing details (Hautière, et al., 2008).

Due to the diversity of computer vision applications and the increasing development of image enhancement techniques, an objective assessment procedure is desired to compare the results provided by different methods. This procedure performance evaluation compares the output of each method with the test image and measures different qualities according to objective metrics. The key challenge here is the presence of a decision process parameter(s) (threshold(s)) that influence the outcome of the methods. So, such evaluation procedures should assist in selecting optimal parameters for each algorithm and ranking different algorithms, depending on specific application requirements.

This chapter aims to evaluate the performance of the proposed Retina-inspired image enhancement methods proposed in Chapter 5 and several state-of-the-art methods that devise the same functionality. Discrepancy empirical evaluation methods for image enhancement approaches, discussed in Chapter 3, are adopted in this chapter. An objective criterion that indicates the quality of the converted image will be used. Public-domain datasets will be used to test and evaluate the proposed methods.

In physics-based algorithms, the synthetic data will not represent the physical parameters of the real monitored scene. Three evaluation methods are adopted, illustrative visual comparison, objective, and computational complexity evaluation. Section 6.2 discusses

the performance metrics used in the threshold selection criterion and objective evaluation throughout this Chapter. Section 6.3 proposes a new contrast enhancement metric. The datasets used in this chapter are introduced in Section 6.3. The illustrative visual comparison is presented in Section 6.4. In section 6.5, the performance metrics are applied to evaluate the proposed algorithms objectively. Finally, the results are analysed.

6.2 Performance metrics

Objective assessment consists of computational methods called “metrics” that produce values that score the image quality. In this thesis, objective metrics are adopted to evaluate the performance of the proposed image enhancement methods by assessing the quality of the contrast-restored image compared to the quality of an input image. Low contrast, high saturation, and excessive brightness change are considered types of defects, which the objective metrics are attempting to spot. The performance evaluation in this Chapter will assess the quality of the contrast-restored image from five different aspects, naturalness preservation, illumination quality, structural consistency, saturation maintenance and contrast enhancement.

6.2.1 Naturalness preservation assessment

Lightness-Order-Error (LOE) is a quality estimation metric proposed in (Wang, et al., 2013), aiming to assess naturalness preservation objectively, where smaller values of LOE scores show better preservation of lightness order and pleasing natural appearance. LOE scores will be used for the objective evaluation of the proposed enhancement technique. LOE scores calculated from the images generated from state-of-the-art methods chosen in this investigation will be used to compare their naturalness preservation alongside LOE scores produced by the proposed Retina-inspired methods.

Wang et al. (2013) assume that the naturalness of an enhanced image is related to the relative order of lightness in different local areas. Since the relative order of lightness represents the light source directions and the lightness variation, they define the quantitative LOE measure based on the lightness order error between I_o and its enhanced

version I_e . They considered the lightness $L(x, y)$ of an image as the maximum of its three components as follows:

$$L(x, y) = \max I_c(x, y) \quad \text{Equation 6-1}$$

I is the image, $c \in \{r, g, b\}$ and (x, y) the pixel's position in the image domain. For each pixel (x, y) , the Lightness Relative Order Difference (LROD) between the original image I and its enhanced version I_e is defined as follows:

$$\begin{aligned} LROD(x, y) \\ = \sum_{i=1}^m \sum_{j=1}^n U[L(x, y), L(i, j)] \oplus U[L_e(x, y), L_e(i, j)] \end{aligned} \quad \text{Equation 6-2}$$

Where m and n are the image's height and width, $U(x, y)$ is the unit step function, \oplus is the eXclusive-OR operator (XOR).

$$U(x, y) = \begin{cases} 1, & \text{for } x \geq y \\ 0, & \text{otherwise} \end{cases} \quad \text{Equation 6-3}$$

The LOE measure is defined as:

$$LOE = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n LROD(i, j) \quad \text{Equation 6-4}$$

The smaller the LOE value is, the better the lightness order is preserved. In their implementation, Wang et al. (2013) have taken the down-sampled versions LROD and LROD_e of size $dm \times dn$ instead of L and L_e . The ratio r between the size of the down-sampled image and that of the original images is set as $r = \frac{50}{\min(m, n)}$. The new size for $dm \times dn$ of the down-sampled image is now $[m \cdot r] \times [n \cdot r]$. Figure 6-1 shows three images generated from using three different approaches, the first image to the left did not preserve naturalness and hence scored $LOE = 825$, the middle image preserved better naturalness and scored $LOE = 539$, while the image to the right of the Figure preserved

naturalness better than both of the previous ones and the achieved LOS score was 393.

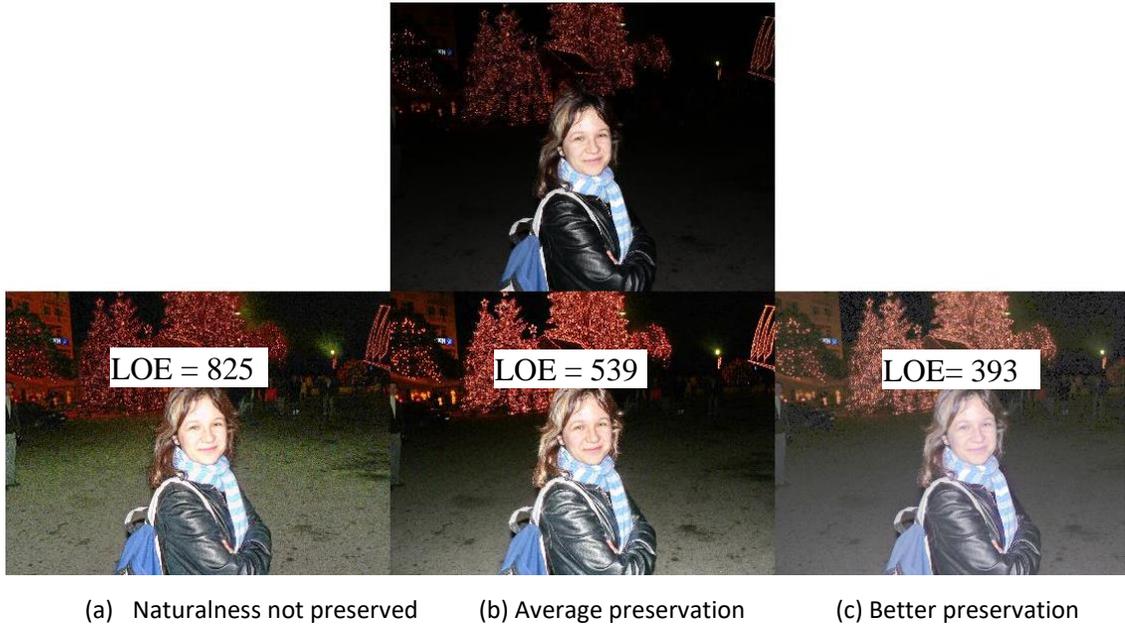


Figure 6-1 Naturalness-preservation examples

6.2.2 Illumination-based quality estimation

An illumination assessment technique was proposed in (Subramanyam et al., 2016), aiming to assess the quality of an image based on illumination. The authors have assumed that there would be very dark regions present if there is bad illumination on the image. Their technique computes such dark regions in the image by coding each pixel based on its neighbours to generate an illumination-based quality score. The technique starts by segmenting bad-illuminated areas from the image. All pixels with $I_c \leq \tau_{low}$, where $\tau_{low} = 43$ is the empirically chosen threshold, the Code at pixel location x, y is defined as:

$$Code(x, y) = \frac{1}{8} \sum_{p=0}^7 E(I_{pixel} - I_{centre}) \quad \text{Equation 6-5}$$

Where p (ranges from 0 to 7) defines index for 8-neighbouring pixels. I_p and I_c define intensity values at p^{th} neighbouring pixels and central pixel, for which code is computed,

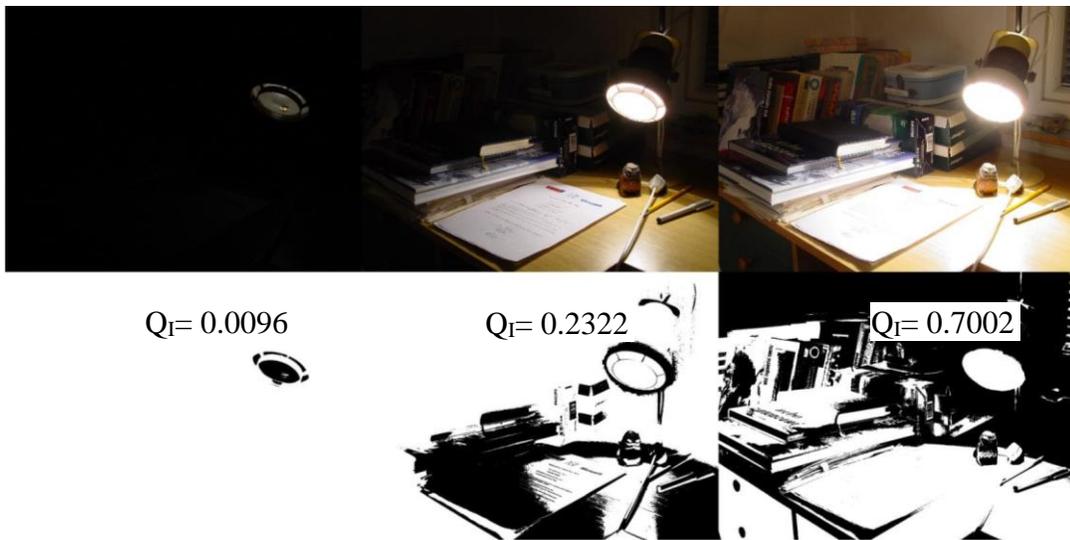
respectively. τ_{low} is the threshold used to detect dark regions in an image. Function $E(.)$ is defined as follows:

$$E(I) = \begin{cases} 1, & \text{for } abs(I) > \epsilon \\ 0, & \text{otherwise} \end{cases} \quad \text{Equation 6-6}$$

If the difference between intensity (I_{centre}) of the centre pixel 'centre' and intensity (I_{pixel}) of its p^{th} neighbour is greater than a threshold ϵ , where $\epsilon = 212$ is the empirically chosen threshold, then the value of $E(I)$ for p^{th} neighbour is assumed to be 1 else it is assumed to be 0. This method gives a value between 0 to 1 for a pixel whose intensity value is below τ_{low} threshold. Values above τ_{low} threshold is coded 1. Overall illumination-based quality is obtained by summation of all codes as follows:

$$Q_I = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n Code(x_i, y_i) \quad \text{Equation 6-7}$$

The images shown in Figure 7 2 represent the same scene under three different illumination conditions, (a) low illumination, (b) dull illumination and (c) where more illumination was presented.



(a) Dark image, Bad illumination

(b) Dull image

(c) Better illumination

Figure 6-2 Illumination-based quality examples

The Codes calculated from Equation 6-7 and generated from each image are directly proportional to illumination change. The darkness in the left image induced a Code of

0.0096, while the middle image produced 0.2322, and the right image showed the highest Q_1 value of 0.3025.

6.2.3 Structure consistency and contrast restoration

One of the main challenges in image enhancement is to keep the image structure intact. Keeping the consistency between the structure of the source and produced image is crucial for several computer vision systems, and mainly for image and video analytic solutions, which would operate on the image/video frame generated from the enhancement technique, expecting to access the same spatial arrangements of the source image.

A structure consistency metric is proposed in (Hautière, et al., 2008), where they considered the set of edges that have a local contrast above 5%. The detected edges are then counted and presented as two cardinal numbers n_o and n_r to denote visible edges in the original image I_o and the contrast-restored image I_r , respectively. They propose to compute a metric to denote the rate of new visible edges, e , in the restored image, as follows:

$$e = \frac{n_r - n_o}{n_o} \quad \text{Equation 6-8}$$

This metric, e , evaluates the ability of the image enhancement method to restore edges that were not visible in I_o but are in the I_r .

Also, they proposed the computation of the geometric mean of the ratios of the VL s calculated as follows:

$$r = \frac{VL_r}{VL_o} \quad \text{Equation 6-9}$$

Where VL_r is the visibility level of the object in the contrast-restored image, calculated using the following formula:

$$VL = \frac{(\delta L/L_b)_{actual}}{(\delta L/L_b)_{threshold}} = \frac{L_{centre} - L_{surround}}{L_{surround}} \quad \text{Equation 6-10}$$

The geometric mean, \bar{r} , emphasises the quality of the image enhancement method to restore the image contrast. Unlike the image restoration metric, e , the contrast restoration metric, \bar{r} , takes into consideration both invisible and visible edges in I_o .

$$\bar{r} = \exp \left[\frac{1}{n_r} \sum_{p_i \in \mathcal{P}_r} \log(r_i) \right] \quad \text{Equation 6-11}$$

Where p_i belongs to all edge pixels in the contrast-restored image, \mathcal{P}_r .

The difference, σ , between the number of saturated black and white pixels after and before restoration process,

$$\sigma = \frac{n_{so} - n_{sr}}{m \times n} \quad \text{Equation 6-12}$$

Where n_{so} denotes the number of pixels saturated in the original image, and n_{sr} denotes the number of pixels saturated in the original image, m and n are the width and the height of the image respectively.

The Patch-based Contrast Quality Index (PCQI), proposed in (Wang, et al., 2015), uses an adaptive representation of local patch structure, which allows the decomposition of any image patch, represented as an N -dimensional vector, into its mean intensity, signal strength and signal structure components and then evaluate their perceptual distortions. They assume that x and y are a pair of co-located patches in the original image and contrast-restored image, respectively, each patch, y , is represented as follows:

$$y = c_1^y \cdot v_1 + c_2^y \cdot v_2 + r \quad \text{Equation 6-13}$$

where,

$$c_1^y = y^T v_1 + \sqrt{N} \mu_y \quad \text{Equation 6-14}$$

and

$$c_2^y = y^T v_2 \quad \text{Equation 6-15}$$

where v_1 is a unit vector, $v_1 = \frac{1}{\sqrt{N}}$, v_2 is another unit vector pointing to a specific direction in the signal space and r is the residual signal perpendicular to both v_1 and v_2 . The contrast change, q_c is formed by scaling the signal strength along the direction of v_2 .

$$q_c(x, y) = \frac{4}{\pi} \cdot \arctan \left(\left| \frac{c_2^y}{c_1^y} \right| \right) \quad \text{Equation 6-16}$$

The structural distortion, q_s is determined by the relative strength of r , as it leads the test signal structure to go away from v_2 , which is defined as follows:

$$q_c(x, y) = \frac{c_2^y + r^T v_2}{\|c_2^y \cdot v_2 + r\|} \quad \text{Equation 6-17}$$

An exponential function based on the absolute difference between c_1^x and c_1^y has been applied to compare the mean intensity, as follows:

$$q_i(x, y) = e^{-|c_1^x - c_1^y|/\sqrt{NL}} \quad \text{Equation 6-18}$$

Where L is the dynamic range of the pixel values. The patch-based contrast image quality index is calculated for M patches as follows:

$$PCQI(x, y) = \frac{1}{M} \sum_{i=1}^M q_c(x, y) \times q_i(x, y) \times q_s(x, y) \quad \text{Equation 6-19}$$

Figure 6-3 present three examples of contrast-restored enhanced images. The enhanced image in Figure 6-3(b) is perceptually more appealing than the original image. Some parts of Figure 6-3(a) are over enhanced, which results in the loss of structural details.

It can be seen that the PCQI values in Figure 6-3 captures the quality degradations, the left image exhibits enhancement in fine details compared to the test image and the middle image, while the right image presents further fine details enhancement. On the other hand, some artificial artefacts near the edges of the objects are introduced. The e , \bar{r} and σ values for the left image are -0.1213 , 7.2867 and 0.0176 , while for the middle image these metrics are -0.0733 , 4.9228 and 0.0092 , while for the right image -0.138 , 4.2023 and 0.015 respectively. The right image has improved the signal intensity, this is reflected in lower e value. While the increase in \bar{r} gives credit for the contrast enhancement. On the other hand, the decrease of σ reflects the reduction in white and black areas.

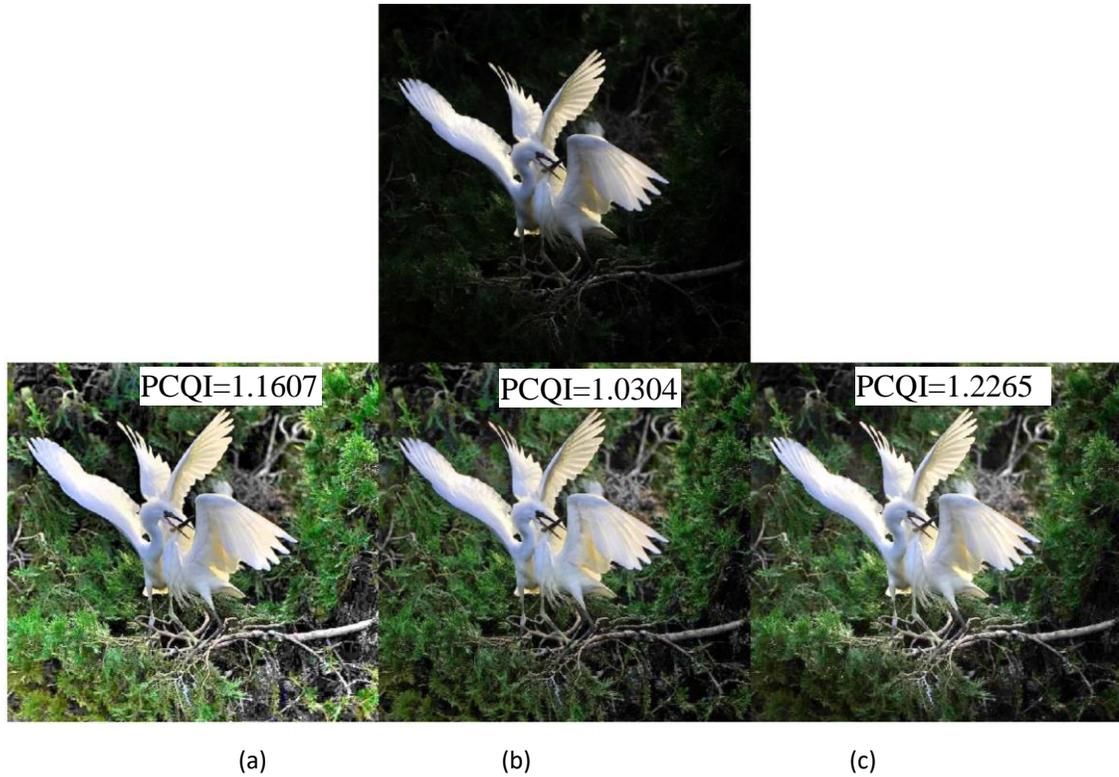


Figure 6-3 Structure consistency quality examples

6.3 Proposed contrast enhancement assessment criterion

This thesis proposes a novel assessment criterion for contrast enhancement, based on the famous law in perceptual (both visual and aural) psychology and psychophysics known as Weber's Law. Weber luminous contrast C , is defined as:

$$C = \frac{\delta L}{L_b} = \frac{L_{centre} - L_{surround}}{L_{surround}} \quad \text{Equation -6-20}$$

Where δL is the luminance difference between *centre* and *surround*, L_{centre} is the luminance of the *centre* and $L_{surround}$ is the luminance of the *surround*. Weber discussed the Just Noticeable Difference (JND) threshold, δL_{jnd} , a value at which a centre becomes perceptible with high probability. This threshold depends on the size of the centre, ambient light's intensity, it decreases with increase of light intensity (Hautière, et al., 2008). The Visibility Level (VL) of a centre can be presented as in Equation 6-10.

In this method, the image contrast defined in Equation -6-20 calculates the number of pixels with low contrast. Let $L_{surround}$ denotes the luminance of the mean field surround and Let δL be the difference or contrast between the mean field surround pixels and the centre pixel embedded within them. For each centre pixel, one thousand surround pixels are averaged then subtracted from the centre pixel, and the result is divided by the average surround pixels as follows:

$$C = \frac{\delta L}{\text{mean}(L_{surround})} = \frac{L_{centre} - \text{mean}(L_{surround})}{\text{mean}(L_{surround})} \quad \text{Equation 6-21}$$

Weber's law proves that an average human observer can identify the difference between a centre pixel and its counterpart surround pixels if the luminance of the centre pixel and the luminance of the surround pixels is more than the just noticeable difference threshold, $\hat{\delta}L$.

Similar to (Shen, 2003), let $T(\cdot)$ be the yes-no binary decision for the low-high contrast pixels, the researcher proves that the psychological evidence shows that $T(L_{centre}|L_{surround})$ behaves like a Heaviside function $H(\delta L - \hat{\delta}L)$ for some threshold $\hat{\delta}L$ depending on $L_{surround}$.

$$T(L_{centre}|L_{surround}) = H(\delta L - \hat{\delta}L) = \begin{cases} 1, & \text{for } \delta L > \hat{\delta}L \\ 0, & \text{otherwise} \end{cases} \quad \text{Equation 6-22}$$

The sum of the T values for all pixels is then divided by the total number of pixels to calculate the Contrast Enhancement (CE) in the image.

$$CE = \frac{\sum_{i=1}^m \sum_{j=1}^n T(x_i, y_i)}{m \times n} \quad \text{Equation 6-23}$$

This thesis proposes the contract enhancement metric to be used as an additional evaluation metric. This new metric is important for evaluating image enhancement algorithms because it demonstrates the effectiveness of the algorithm and underpins the robustness of the algorithm to scene variations.

Such a metric can give a better indication of the enhancement achieved by a specific algorithm to preserve the image's contrast. When it comes to the comparison between different algorithms, instead of comparing each image with the original image, a single metric, the CE metric, could be used to compare between several algorithms and at the

same time to present the enhancement achieved compared to the original image. Each algorithm will be featured with its CE value, and this will be used to conduct better performance analysis. There is no image enhancement evaluation criterion that identifies the contrast enhancement based on the physiological basis to the author's knowledge. The proposed algorithm can produce a local contrast quality map, which could be used for helping other image enhancement techniques improve their results. Figure 6-4 shows a test image (top) and three contrast-restored images, with their corresponding CE maps and binary maps, indicate the spatial locations where the contrast is degraded. The middle image has higher contrast than the left image, while the right image has the highest contrast. The maps generated from the proposed method is shown below each image, where the black areas are the areas that require further contrast enhancement.

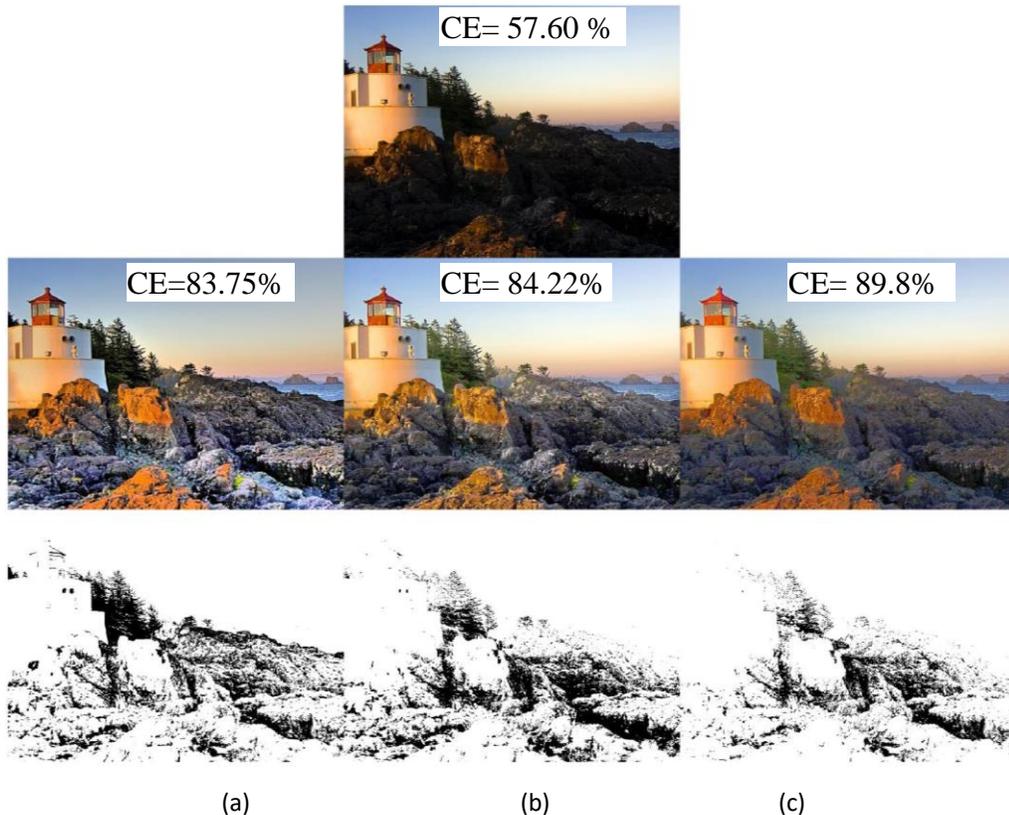


Figure 6-4 Contrast enhancement quality example and their associated maps

6.4 Datasets

Five datasets have been used to evaluate the performance of the proposed techniques. These datasets are DICM (Lee, *et al.*, 2012), and NPE-data-1 (Wang, 2013), NPE-data-2 (Ma, *et al.*, 2015), NPE-data-2 (Wang, *et al.*, 2015) and LIME-Data (Guo & Li, 2016). DICM dataset, collected by (Wang, 2013), includes 69 outdoor, indoor, underwater well space images acquired using commercial digital cameras. The images present dusk, low-light and poorly illuminated scenes.

NPE-data-1 dataset, used in (Wang, 2013), consists of 77 outdoor nature scenes and includes some aerial images. In general, the images presented are well-illuminated; some images present night-time scenes. The dataset includes daytime, daybreak, nightfall, night-time as well as cloudy scenes.

NPE-data-2 dataset, used in (Ma, *et al.*, 2015), contains 23 high-quality poor-exposed outdoor images. Each scene in this dataset includes a part of the correctly exposed image and another part that is severely under/over-exposed.

NPE-data-3 dataset, used in (Wang, 2013), consists of 131 low quality outdoor natural scenes images

LIME-data, used in (Guo & Li, 2016), consists of 10 low-light images. NPE. This dataset contains 85 low-light images downloaded from the Internet.

6.5 Illustrative visual comparison

The test images provided by the datasets, discussed in Section 6.4, are used as input to the image enhancement methods under investigation in this Section. The results are evaluated by visual comparison between the images restored using the proposed image enhancement methods and the images generated from the same state-of-the-art methods used in techniques used in the objective evaluation, section 6.6, namely NPEA, MSRCR, MF, LIME and LECARM. The contrast-restored result is compared visually to the test image and the output of the state-of-the-art methods.

Due to space constraints, only specific images from the processed sequences are presented.

The image in the top left section of Figure 6-5 shows a view taken from a sea cave looking across a beach. In this image, there is a shadow in the foreground at the entrance to the cave. The cave wall appears to be saturated, almost black, and the surface details are not defined. The middle image in the top section is the output obtained by applying the multi-scale model of adaptation. The result shows that the model concentrated predominantly on the darker areas, highlighting the shadowed areas. This resulted in an almost CGI rendition of the original image and appeared unrealistic. When applying the proposed model (right-hand image top section of Figure 6-5), the definition appears to be more precise and transparent. On the other hand, the visual model used in the middle image highlights the details of the sand patterns within the cave shadow.

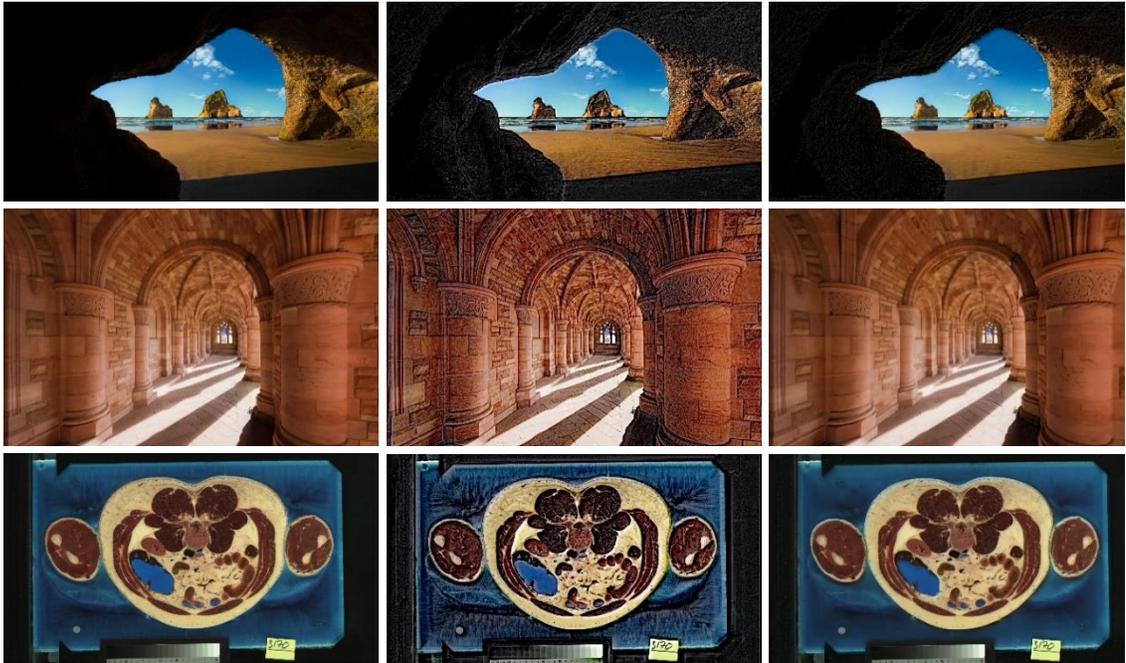


Figure 6-5 Illustrative visual comparison of the results obtained, the original image (left), the output of the multiscale model of adaptation (middle) and physics-based retina-inspired model (right).

The first image of the second row (from the left-hand side) is of a cloister in a monastic building highlighting the shadows bouncing off sandstone columns. The image also defines the details in the carvings above the pillars. The second image of the same row is less natural due to the oversaturation of detail. However, the intricate carvings are

highlighted but, to some extent, exaggerated. The third image in the same row on this occasion is visually very similar to the first image.

The bottom left image in Figure 6-5 shows an example of a medical image. The output of the multi-scale adaptation model is showing good improvement by extracting details such as blood vessels that could improve medical diagnosis. However, the results obtained by using the proposed model (right bottom off show fewer details) could improve particular objects' segmentation quality, depending on the required application.

The image on the top shows an image taken at night. The gable ends, roofs and building facades are blurred. The results obtained by applying a multi-scale adaptation model (the top middle image of Figure 6-5) show that too many artefacts were identified, making it challenging to identify essential features such as rooftops or eaves. In comparison, the top left image shows the output of our proposed model. This brings out the finer details with a reasonable level of unwanted artefacts. Clear lines of the gable end, roofs and building facades can also be observed.



Figure 6-6 Illustrative visual comparison of the results obtained, the original image (left), the output of the multi-scale model of adaptation (middle) and physics-based retina-inspired model (right). The left middle image of Figure 6-6 represents a bright image with glare from objects. The multi-scale adaptation model (the second image in the middle row), although it finds more features, also detects unwanted artefacts on a desk draw, which can give misleading results. It can also be noticed that some book titles are distorted due to unwanted artefacts.

With the proposed model (middle right of Figure 6-6), the output of our algorithm detects less unwanted data yet still provides good sharpness of objects.

Finally, the bottom left image in Figure 6-6 shows a similar scenario as the middle image, i.e. a desk with a lamp and books on one side of the table. However, this darker image where the lamp was the only source of light, and, therefore, the book titles are not visible. Both models improved the image, and straightforward book edges are visible, and book titles can be legible. However, the multi-scale adaptation model (the second image in the last row of Figure 6-6) generates an unwanted mist effect around the table lamp.

Figure 6-7, Figure 6-8 and Figure 6-9 show several comparisons between the proposed methods, LGRF and SGRF, with state-of-the-art methods. The operations of NPEA, MSRCR, MF, LIME, and LECARM are executed on the input images' value channel by first converting the RGB colour space to the HSV. The generated output is then converted back to RGB colour space. It can be observed that NPEA, MSRCR, MF, LECARMp can recall the information in the dark regions by increasing the luminance. However, they produce visual artefacts in the areas with shadow, as Figure 6-8 indicates.

The MSCRC highlights visual artefacts around the lamp in Figure 6-7. The use of a physics-based model has made the proposed GRF method prone to such artefacts. The illumination estimation module accommodates the ambient illumination, and the Rod signal integration in the centre-surround scheme allowed better adaptability to illumination changes across the scene. The proposed algorithm, GRF, maintained a better balance between luminance and contrast.

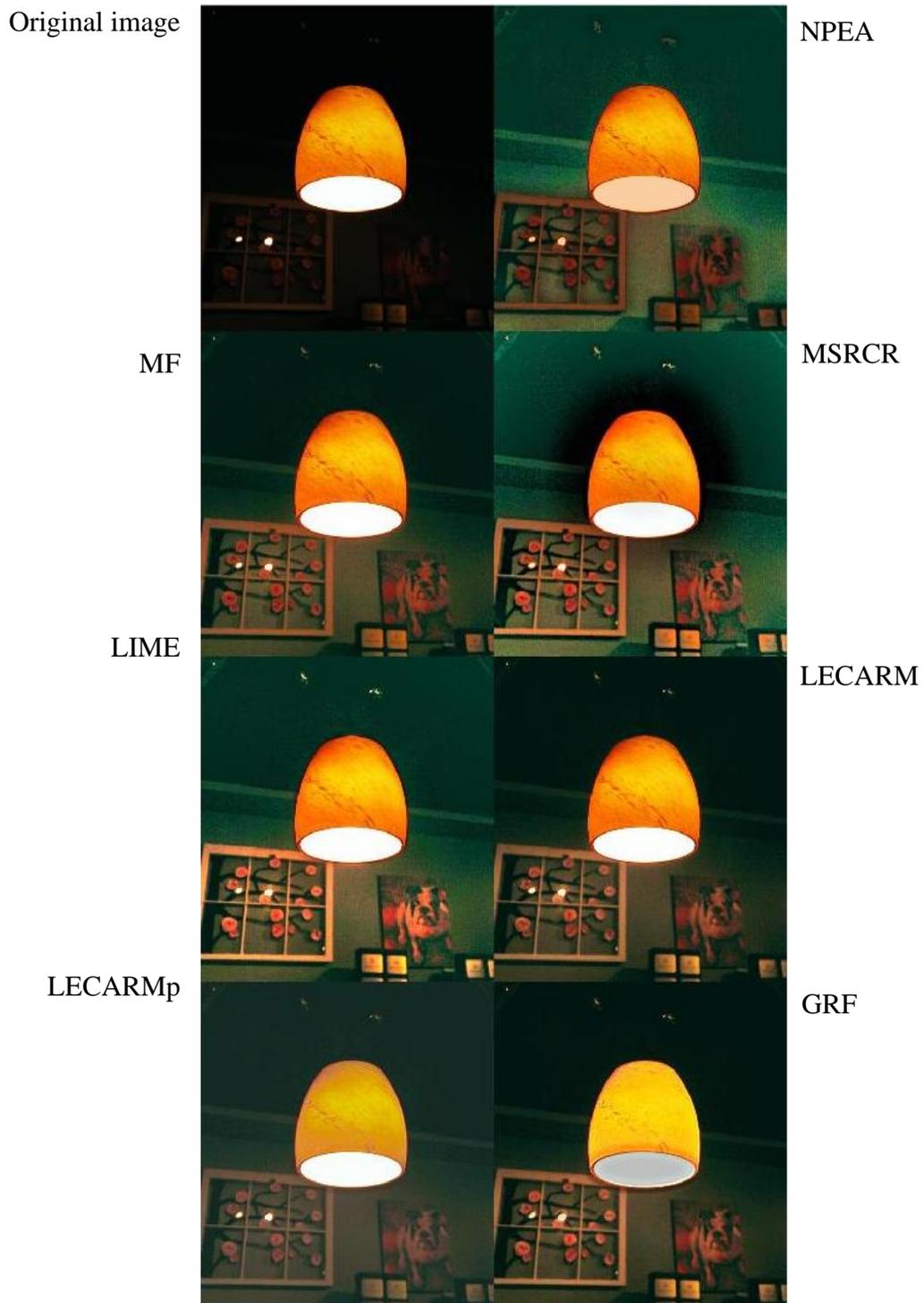


Figure 6-7 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods

Original image



NPEA

MF



MSRCR

LIME



LECARM

LECARMp



GRF

Figure 6-8 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods

6.5 Illustrative visual comparison

Original image



NPEA

MF



MSRCR

LIME



LECARM

LECARMp



GRF

Figure 6-9 Illustrative visual comparison of the results obtained, the original image (top left), the output of the proposed methods and state-of-the-art methods

6.6 Objective evaluation

In this section, the objective evaluation metrics defined in Section 6.2 are used to compare and rank the proposed image enhancement algorithms with several state-of-the-art algorithms, including Naturalness Preserved Enhancement Algorithm (NPEA) (Wang & L., 2017), Multi-Scale Retinex Colour Restoration (MSRCR) (Petro, et al., 2014), Multi-deviation Fusion (MF) (Fu, et al., 2016), Low-Light Image Enhancement via Illumination Map Estimation (LIME) (Guo & Li, 2016) and Low-light Enhancement Camera Response Model (LECARM) (Ren, et al., 2018). All the algorithms are implemented and tested on Matlab, which ensures fair comparison. These results complement the illustrative visual comparison results in Section 6.5 and those presented in this Chapter. The evaluation of the enhancement is carried out using the LOE, QI, e , \bar{r} , PCQI and CE. The quantitative performance results are presented in Table 6-1, Table 6-2, Table 6-3, Table 6-4, Table 6-5, Table 6-6 and Table 6-7 respectively.

Table 6-1 Quantitative measurement results of LOE

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
DICM	311	390	1136	819	532	640	256
NPEA-Data-1	325	384	1250	863	563	705	256
NPEA-Data-2	241	176	1155	470	238	276	160
NPEA-Data-3	197	350	1276	824	513	651	151
LIME-Data	362	250	937	551	298	349	164
<i>Average</i>	287.2	310	1150.8	705.4	428.8	524.2	197.4

The LOE values for all methods under investigation, presented in Table 6-1, show that the proposed GRF method achieved the lowest LOE values over all five datasets, followed by the NPEA, MF and the LECARM methods, respectively. The proposed GRF method outperforms all state-of-the-art methods; this matches the qualitative assessment presented by the visual illustration examples shown in the previous section, where these algorithms present natural luminance restoration compared to the other methods. This indicates that the proposed method maintains a better luminance signal compared to the state-of-the-art methods.

Table 6-2 Quantitative measurement results of CE

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
DICM	92.02%	88.84%	89.80%	89.28%	88.11%	94.92%	95.86%
NPEA-Data-1	93.64%	92.75%	88.29%	92.51%	93.64%	94.53%	97.11%
NPEA-Data-2	80.51%	72.96%	79.10%	76.30%	72.46%	75.57%	82.28%
NPEA-Data-3	91.47%	90.54%	86.23%	90.43%	90.82%	92.46%	97.50%
LIME-Data	95.06%	93.88%	90.05%	92.98%	95.07%	95.07%	97.89%
<i>Average</i>	90.54%	87.79%	86.69%	88.30%	88.02%	90.51%	94.13%

Table 6-2 illustrates the contrast enhancement measured by the proposed CE metric. The results indicate that the proposed GRF method outperforms all state-of-the-art methods, followed by the proposed physics-based LECARM method, original LECARM and the NPEA methods. This confirms the visual illustration examples shown in the previous section, where these algorithms present better contrast enhancement compared to the other methods. This shows the effectiveness of the proposed contrast enhancement metric as it is directly proportional to the human perception, which confirms the suitability of applying Weber's law to assess contrast enhancement quantitatively.

Table 6-3 Quantitative measurement results of PCQI

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
DICM	0.9303	0.9434	0.9773	0.9289	0.8979	0.8693	0.8128
NPEA-Data-1	0.9547	<u>0.9825</u>	1.0201	0.9652	0.9289	0.9174	0.7906
NPEA-Data-2	0.9921	<u>0.9981</u>	0.9795	0.9985	0.9105	0.9111	0.8402
NPEA-Data-3	1.0281	0.9682	0.9889	0.9578	<u>1.0113</u>	1.0026	0.8962
LIME-Data	1.1150	1.0199	1.0339	1.0403	<u>1.0704</u>	1.0829	0.9648
<i>Average</i>	1.00404	0.98242	0.99994	0.97814	0.9638	0.95666	0.86092

Table 6-3 presents the PCQI values performed on all five datasets and the restored images for all methods under investigation. The results show that the NPEA method achieved the highest PCQI values, followed by MSRCR, MF and LIME. By contrasting the results obtained from Table 6-1, Table 6-2 and Table 6-3 and knowing that the PCQI metric combines three different metrics, structure, contrast and luminance, it can be concluded that the proposed GRF algorithm has not scored High in the structure preservation assessment carried by the PCIQ criterion. This is due to the centre-surround framework

used in this method, which affects the image structure. This observation confirms that the human visual contrast manipulates the image structure to achieve an appealing image presentation with the psychological findings.

Table 6-4 Quantitative measurement results of QI

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
DICM	93.03%	94.34%	97.73%	92.89%	89.79%	86.93%	81.28%
NPEA-Data-1	93.43%	92.82%	90.46%	85.73%	90.29%	90.96%	95.58%
NPEA-Data-2	86.32%	74.69%	88.75%	79.04%	70.45%	74.79%	86.05%
NPEA-Data-3	92.01%	91.51%	90.42%	84.06%	87.04%	87.68%	95.09%
LIME-Data	92.21%	85.58%	87.80%	85.87%	79.33%	81.06%	92.16%
<i>Average</i>	91.40%	87.79%	91.03%	85.52%	83.38%	84.28%	90.03%

The QI values are presented in Table 6-4. This metric measures the percentage of restored black and white areas. The NPEA algorithm achieved better overall performance, the performance of the GRF proposed method outperforms all other methods in two datasets (NPEA-Data-1 and NPEA-Data-3) and matches the performance of NPEA in the other two datasets (NPEA-Data-2 and LIME-Data), but it achieved much lower QI value in the DICM dataset., which has resulted in lower overall QI by 1.4%. Such drop relates to the nature of the DICM dataset, as it includes images taken in space and underwater. The results show that such scenes, which are not fully correlated with the image formation model, have dropped the performance of the GRF proposed method. Apart from space and underwater scenes the proposed GRF method has shown very good performance in recovering black and white patches in the scene. This confirms the suitability of the log-power-exp model used in this thesis to simulate the functionality of the Ganglion cells. Such model allows balanced weighing for the enhancement applied to different image patches depending on their luminance value.

Table 6-5 Quantitative measurement results of e

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
DICM	0.0228	0.0344	0.1566	0.0270	-0.01902	-0.1765	-0.09314
NPEA-Data-1	0.1324	0.0070	0.4750	0.0001	-0.0649	-0.0649	-0.1628
NPEA-Data-2	0.2947	0.3153	0.2728	0.2963	0.00627	0.0937	-0.0245
NPEA-Data-3	-0.0194	-0.0114	-0.0374	-0.0284	-0.0413	-0.04135	-0.1554
LIME-Data	0.1799	0.0578	0.1915	0.1050	0.0538	-0.2659	-0.0802
<i>Average</i>	1.22e-1	8.06e-2	2.12e-1	8.00e-2	-1.30e-2	-9.10e-2	-1.03e-1

Table 6-5 illustrates the e metric values measured by the formula in Equation 6-8. The results indicate that the modified physics-based LECARM method outperformed other methods, followed by the original LECARM and the proposed GRF method. This confirms the results achieved in Table 6-1 Quantitative measurement results of LOE

Table 6-6 Quantitative measurement results of \bar{r}

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
DICM	3.2227	2.8613	5.4478	3.5783	2.6080	2.4171	2.6041
NPEA-Data-1	2.3861	2.3714	3.9285	<u>3.0327</u>	2.5021	2.502	1.6881
NPEA-Data-2	3.9748	2.9704	6.7157	<u>4.6146</u>	2.9168	2.8509	2.8686
NPEA-Data-3	2.3826	2.4873	4.0294	<u>3.2701</u>	2.6197	2.6163	1.7501
LIME-Data	<u>4.1688</u>	3.360	5.2846	<u>3.5299</u>	3.254	3.2054	2.639
<i>Average</i>	3.227	2.81008	5.0812	3.60512	2.78012	2.71834	2.30998

The \bar{r} values are presented in Table 6-4. This metric measures the structure preservation by comparing the edges in the test image and the ones recovered from the restored image. Although the MSRCR outperforms all the other methods, but, the visual assessment show that such recovered edges are not realistic and does reduce the naturalness of the restored image. Again, this confirms the results in Table 6-3, where the proposed GRF method changes the scene structure in the restored image to give more naturalness preservation, which is consistent with the psychological findings. The GRF adopts centre-surround mechanism, which makes changes to the structure.

Table 6-7 Quantitative measurement results of σ

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
DICM	8.86e-3	9.79e-3	1.09e-2	8.83e-3	9.96e-3	9.4e-3	8.57e-3
NPEA-Data-1	6.90e-3	7.81e-3	1.25e-2	7.2e-3	9e-5	5.167e-5	6.8e-3
NPEA-Data-2	2.15e-3	3.50e-3	2.113e-3	2.119e-3	1.45e-6	2.07e-7	2.05e-3
NPEA-Data-3	4.61e-3	2.89e-3	1.06e-2	2.50e-3	4.59e-3	0	4.59e-3
LIME-Data	8.95e-3	3.33e-3	9.20e-3	3.16e-3	0	0	8.91e-3
<i>Average</i>	6.29E-03	5.46E-03	9.06E-03	4.76E-03	2.93E-03	1.89E-03	6.18E-03

The σ values are presented in Table 6-7. This metric measures the number of saturated pixels in both the test image and the restored image. The modified physics-based LECARM method outperforms the other methods.

Table 6-8 Overall Quantitative measurement results

	NPEA	MF	MSRCR	LIME	LECARM	LECARMp	GRF
LOE	<u>287.2</u>	310	1150.8	705.4	428.8	524.2	197.4
CE	90.54%	87.79%	86.69%	88.30%	88.02%	90.51%	94.13%
PCQI	1.00404	0.98242	<u>0.99994</u>	0.97814	0.9638	0.95666	0.86092
QI	91.40%	87.79%	<u>91.03%</u>	85.52%	83.38%	84.28%	90.03%
e	1.22e-1	8.06e-2	2.12e-1	8.00e-2	<u>-1.30e-2</u>	-9.10e-2	-1.03e-1
\bar{r}	<u>3.227</u>	2.81008	5.0812	3.60512	2.78012	2.71834	2.30998
σ	6.29E-03	5.46E-03	9.06E-03	4.76E-03	<u>526900</u>	1.89E-03	6.18E-03

The average performance over five databases is summarized in Table 6-8 Overall Quantitative measurement results at the rightmost of Table 6-8, which demonstrates the superior performance of the proposed GRF method in luminance restoration and contrast enhancement, however it does change the appearance of the scene's structure in the restored image. The MSRCR performs the highest in preserving the structure, however this reduces the contrast and the overall naturalness of the output image. The NPEA gives a good balance between the luminance restoration and structure preservation, however, it performs less in contrast enhancement. The modified physics-based LECARM method shows superior performance in illumination restoration as well as naturalness preservation.

6.6.1 Quantitative testing – video signal

The Changedetection.net (CD.net) challenge (Goyette *et al.*, 2012), presents a comprehensive academic benchmarking effort for testing and ranking existing and new algorithms for change and moving object segmentation. This section compares the performance of Spectral-360 change detection method (Sedky *et al.*, 2014) and its integration with (a) the multiscale model of adaptation method (Pattanaik *et al.*, 1998), a Retina-inspired model and (b) with our proposed tone mapping method, a physics-based and Retina-inspired model.

6.6.2 Evaluation

The evaluation of the accuracy of detection is carried out using Precision, Recall and the overall metric F-measure for the Thermal category that includes five videos, corridor, dinningRoom, lakeside, library and park.

Sample frames of the processed videos, as well as their corresponding image enhancement outputs and output masks are shown in Figure 6-11.

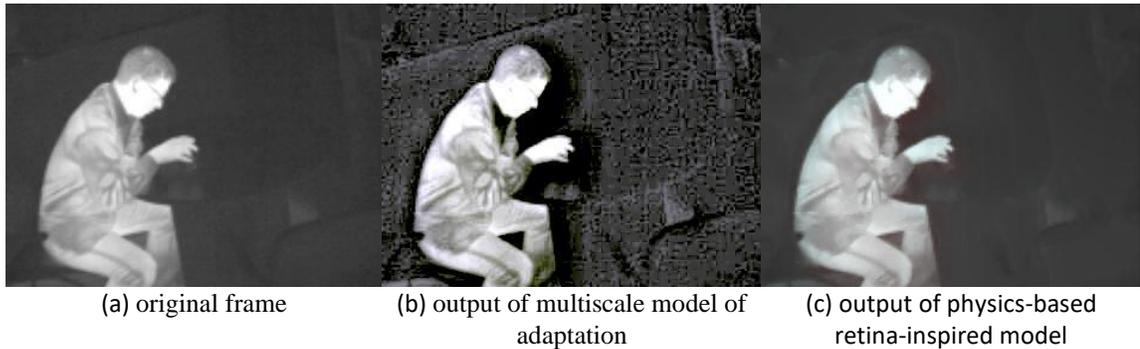


Figure 6-10 Frame number 1500, library video, thermal dataset

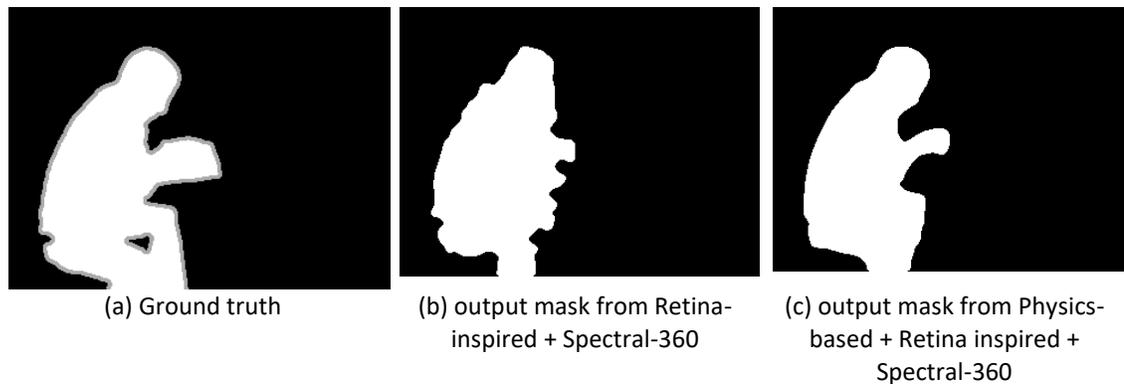


Figure 6-11 Illustrative visual comparison of the results obtained, the original image (left), the output of the multi-scale model of adaptation (middle) and physics-based retina-inspired model (right) By analysing the F-measure, the three methods could be ranked as shown in Figure 6-12. The best performance on average is shown by our proposed method with 79.4%, followed by Spectral-360 method with 77.64%, then Spectral-360 modified by the Retina-inspired method with 57.91%.

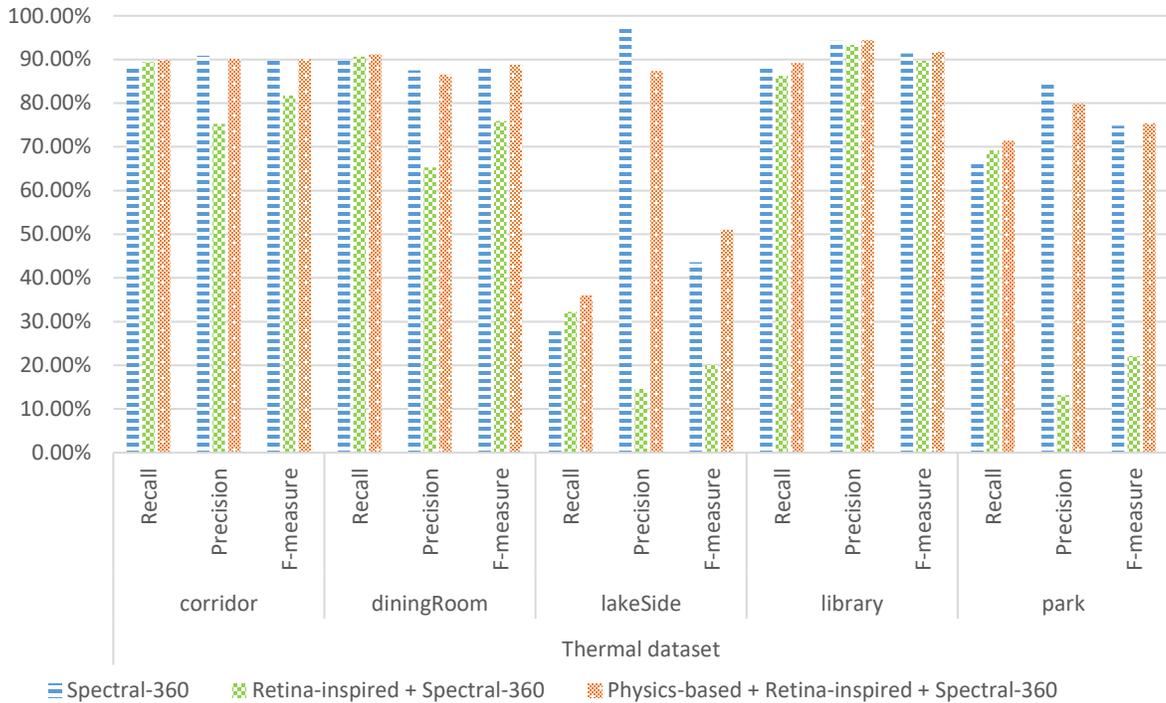


Figure 6-12 Average Precision, Recall and F-measure for all videos of the Thermal category. The method with the lowest over-segmentation (the highest Recall) is our proposed method with 75.54%, it is followed by Spectral-360 modified by the Retina-inspired method with 73.57%, then Spectral-360 with 72.38%. On the other hand, the algorithm with the lowest under segmentation (the highest Precision) is Spectral-360 with 91.14%, it is followed by our proposed method with 87.68%, then Spectral-360 modified by the Retina-inspired method with 52.34%. The proposed image enhancement method integrated with the Spectral-360 change detection method performs better than the original Spectral360 method in all the videos of the Thermal category.

6.7 Summary

In this Chapter, the proposed physics-based modified algorithms, Multi-scale Adaptation Model (MAM) and LECARM, and the novel GRF image/video enhancement method, were tested.

The quantitative performance analysis shows that embedding both contrast and colour constancy by extracting physical features from the camera output, applying a centre-surround framework alongside the log-power-exp model has achieved better contrast enhancement, luminance restoration and naturalness preservation unlike other image enhancement methods that use the camera output directly without considering its physical meaning or the application of simplified Retinal connection models.

The estimation of the spectral characteristics of the dominant illuminant allows the proposed technique to adapt itself to different illumination conditions; this means that it would be applicable for a variety of scenes and not to be limited to certain environments.

The results have shown that the estimation and use of physics-based spectral image representations, deduced from the dichromatic model, represent a more realistic input to the retina-inspired models and would mimic the signal received by the human eye.

A new contrast enhancement metric, based on Weber's law, has been proposed and used to evaluate the performance of state-of-the-art methods. Experimental results show that the proposed CE metric is well correlated with image quality's subjective assessment, signifying that the proposed metric is promising at assessing contrast changed images. The CE map produced by the proposed assessment method also indicates local contrast variations, which could be beneficial in developing new contrast enhancement methods.

Chapter 7

Conclusions and Further Work

7.1 Overview

This research investigated the proposal of a retinal network computational model for image enhancement that mimics different retinal layers, namely targeting the interconnectivity between the Bipolar receptive field and the Ganglion receptive field. The research started with a literature review of image enhancement techniques and their architect, which has two main stages: image representation and image filtering. The image representation had two main methods to represent an image, based on being physics-based and non-physics based. In image filtering, the two required methods found were the frequency-domain and the spatial-domain methods; the principle of frequency-domain methods is to compute the Fourier transform of an image, then multiply the resulting filter. In contrast, the spatial domain deals directly with the pixel value. There are a wide variety of techniques to achieve visually accepted images. Conventional image enhancement techniques; have been widely used with their different advantages and drawbacks; there is a recent interest in Bio-inspired image enhancement techniques, i.e., Retinex, Convolutional Neural Networks (CNN), Multiscale Adaptation Model MAM (Pattanaik, et al., 1998), and Naturalness (Joshi & Prakash, 2020). They attempt to mimic the human visual system. So, for example, the main advantages of Retinex are image sharpening and colour constancy (Land, 1986), while the critical benefit of CNNs is their accuracy in object recognition. The chosen models for more in-depth studying were the MAM and Naturalness preservation models, where the details of implementation and modification were further discussed in chapter five.

In chapter three, a detailed study of the Human Visual System (HSV) and the neural structure of the retina was carried out to model the visual pathway. Chapter three started by identifying the retinal four primary layers, beginning with the Rods and Cones,

analogous to camera sensors, rods are scotopic. They detect mainly luminance, cones that are photopic detect the colours, cones are primarily three types: long (L), medium (M) and short (S) cones, just like the camera sensors R, G, B.

The second layer of the retina consists of the Horizontal cells and the Gap cells, which operate horizontally, this type of cells acts in data compression. For high signal speed, the gap cells connect the rods to the bipolar cone cells.

The third layer of the retina consists of bipolar cells and amacrine cells; bipolar cells have a receptive field comprised of a centre and surround; this type of cell functions are edge detection, intensity information, and motion detection of objects. At the same time, the Amacrine cells function as detectors of background motion.

The fourth layer is the ganglion cell layer, which leads to the three major parallel pathways connecting the optic nerve to the visual brain cortex; the P-pathway carries signals from the L and M cones, the K-pathway carries signals from the yellow (L+M) and S cones, as for the M-pathway carries the luminance signal.

Then the chapter will review in-depth the Receptive Field concept and how it is applied in colour vision, finally Weber's law, where a contrast enhancement method will be proposed in chapter six based on Weber's Law.

In chapter four, the physics-based image formation models were reviewed to extract meaningful physical features suitable to represent the scene and the objects contained within it. This research has identified the inability of image enhancement techniques to deal with challenging imaging conditions based solely on statistical methods; it was concluded that understanding the underlying physics, which govern the image formation process, is crucial to deal with the variation in imaging parameters. The physics-based approach enables self-adapting to various lighting conditions and better identification of colour. These advantages and benefits mean that image formation models provide an excellent base to get the best chance of achieving image enhancement.

The proposed image enhancement technique exploited physics-based features, namely estimating the illumination characteristics of the scene and the surface spectral

reflectance-based properties of objects and working automatically when illumination and scene's elements are unknown. The challenge of the image enhancement approach based on surface spectral reflectance arises from the new idea of processing the full spectrum of the surface spectral reflectance instead of the three samples used by other colour spaces. The spectral representation uses a linear model consisting of several basis functions (pre-trained from a set of materials) and weights (calculated for the object under investigation). Numerical estimation of the physics-based model for image formation and the real-time transformation from the camera output to the physical parameters is carried out. The assumptions are that only diffuse reflections are considered, and the existence of a dominant illuminant is applicable.

Chapter five introduced two newly modified state-of-the-art image enhancement methods through their integration with image formation models and proposed a novel Ganglion receptive field computational model for image enhancement (GRF model). Instead of considering only the interactions between each pixel and its surroundings within a single colour layer, the proposed framework introduces the interaction between different colour layers to mimic different retinal layers, namely targeting the interconnectivity between the Bipolar receptive field and the Ganglion receptive field. The Ganglion cells operation is presented using a tone mapping function given with a new log-power-exponential function.

In Chapter Six, the proposed physics-based modified algorithms, MAM, naturalness preservation model, and the novel GRF image enhancement model were tested.

The quantitative performance analysis shows that embedding both contrast and colour constancy by extracting physical features from the camera output; applying a centre-surround framework alongside the use of the log-power-exp model has achieved better contrast enhancement, luminance restoration and naturalness preservation, unlike other image enhancement methods that use the camera output directly without considering its physical meaning or the application of simplified Retinal connection models.

The results have shown that the estimation and use of physics-based spectral image representations, deduced from the dichromatic model, represent a more realistic input to the retina-inspired models and mimic the signal received the human eye.

Based on Weber's law, a new contrast enhancement metric has been proposed and used to evaluate the performance of state-of-the-art methods. Experimental results show that the proposed CE metric correlates with image quality's subjective assessment, signifying that the proposed metric is promising at assessing contrast changed images.

The CE map produced by the proposed assessment method also indicates local contrast variations, which could be beneficial in developing new contrast enhancement methods.

7.2 Contribution to Knowledge

The challenges that face the use of physics-based Retina-inspired models as a basis for image enhancement algorithms to target various images are challenges imposed by the operational requirements of image enhancement application and challenges related to the mathematical complexity of the image formation models.

The challenge to develop robust image enhancement techniques that can deal effectively with different scene structures is tackled by considering the physics that govern image formation. Robustness means the algorithm's ability to adapt automatically with changes in the image details to cope with changes in lighting, scene geometry and scene activity. The image is represented in the proposed image enhancement algorithms using physics-based cues. Information about the illumination and intrinsic features of the material in the scene is extracted and used to solve colour constancy.

There are different reflection models proposed in the literature, as discussed in Section 4.2. However, the choice of a suitable reflection model for such an image enhancement application is not easy. There is a need to consider the assumptions of different models and adopt the model with assumptions that do not contradict real-world operational conditions.

The other consideration is the variability of surface spectral reflectance for different materials. Ideally, the algorithm should be able to model all types of materials contained

in a workplace environment. However, the available models are trained for only a limited number of materials.

One of the main problems is the illumination variation and how to represent a scene's reference under varying illumination conditions. This leads to the challenge of estimating the illumination and choosing the suitable colour constancy technique, which can ensure a convenient estimation of the illumination and extraction of the surface spectral reflectance.

The challenge of the proposed image enhancement approaches based on surface spectral reflectance arises from the new idea of processing the full spectrum of the surface spectral reflectance instead of the three samples used by other colour spaces. The spectral representation uses a linear model consisting of several basis functions (pre-trained from a set of materials) and weights (calculated for the object under investigation). Numerical estimation of the physics-based model for image formation and the real-time transformation from the camera output to the physical parameters is carried out.

The challenge related to the mathematical complexity of the image formation models is tackled. Firstly, by choosing an appropriate reflection model, the dichromatic reflection model proves its effectiveness in computer vision applications.

Secondly, setting a feasible set of assumptions, diffuse-only reflection and the existence of a dominant illuminant. No assumption is made about surface geometries, surface texture, or types and shapes of shadows, objects, and background, which shows the model's applicability for practical operational requirements.

Thirdly, to tackle the variability of surface spectral reflectance for different materials, the linear model was adopted, consisting of several basis functions, pre-trained from a set of materials; and weights, calculated for the scene investigation.

Fourthly, the challenge of choosing suitable colour constancy techniques that can ensure a rapid and convenient estimation of the illumination and extraction of the surface spectral reflectance is solved by choosing an illumination estimation method (McCamy's method), which has been proven successful and valuable for real-time processing.

The main contributions presented in this work can be divided into primary and secondary contributions.

1. The primary contributions are as follows :
 - a. A novel physics-based computational model of the Retina, which mimics biological Retina functioning layers, integrated with the dichromatic image formation model. The algorithm computationally estimates a consistent physics-based colour descriptor model of surface spectral reflectance from the camera output and then uses the full-spectrum reflectance of the scene to mimic the response of the Rods and Cones. Such responses are manipulated to computationally model the centre-surround receptive field behaviour of the Bipolar and Ganglion cells and their interactions, alongside the proposal of a new log-power-exp tone mapping function to mimic the response of the ganglion cells.
 - b. Two modified image enhancement models integrate a physics-based image formation model, where the surface spectral reflectance has been extracted and adopted as a reference of the scene under investigation, instead of using the three colour samples used by other colour spaces. The results show that these modified techniques can adapt to scene variations, such as a change in illumination and scene structure, and give superior performance over the original model.
2. The secondary contribution is :
 - a. A new contrast enhancement assessment criterion based on Weber's law image. The proposed CE metric correlates with image quality's subjective assessment, signifying that the proposed metric is promising at assessing contrast changed images.

Performance assessment results show that the novel algorithm (contribution 1.a) matches the requirements of image enhancement application better in terms of flexibility and robustness to cope with changes in illumination and scene structure adaptively. These

match better with the requirements, which boosts the algorithm's applicability in real-world scenarios, in line with the proposed contrast enhancement assessment criterion (contribution 2.a), where technical capability and applicability are essential considerations.

7.3 Work Limitation

The Bidirectional Reflectance Distribution Function (BRDF) is the general model of reflectance; however, the measurement of such a model is complex and computationally expensive; in this research, simplified models have been chosen to model the reflectance for the proposed GRF model. The dichromatic model was chosen to represent the image formation. A feasible set of assumptions have been set for the Dichromatic model, which best matches the reduction of model complexity.

Two assumptions have been considered, diffuse-only reflection and the existence of a dominant illuminant. Due to the computational complexity of the Dichromatic model, it has been simplified further, where the specular reflectance was neglected. The main limitation of this simplification is that it can be applied only to inhomogeneous dielectrics, Lambertian surfaces. Many common materials can be described this way, including paints, varnishes, paper, ceramics, and plastics. Metals, glass, and crystals are excluded as they are optically homogeneous. Only opaque surfaces are considered in this model. Such assumption limits the performance of the proposed technique when it comes to highlights in the image.

To tackle the challenge of the variability of surface spectral reflectance for different materials, the linear model, proposed by Parkkinen, was adopted in this work, which consists of several basis functions, pre-trained from a set of materials; and weights, calculated for the scene under investigation. Using finite-dimensional linear models to represent the SSR provides a compact description of data, with few basis functions SSR for general materials could be represented. Although, Parkkinen calculated Nine basis functions to represent the SSR for the Munsell dataset, in this research, only Three

Parkkinen basis-functions were used, to recover the scene's SSR from a camera input with three colour channels, which limits the accuracy of the SSR estimation.

The chosen illumination estimation method, McCamy's method, assumes a single dominant illuminant and does not take into consideration light reflected from other materials.

The spatial decomposition of the proposed Retina-inspired computational model, the GRF model, represents a feedforward neural circuitry. Neuroscience findings show evidence of feedback connections between the Horizontal cells/Gap cells and the photoreceptors/Bipolar cells, besides feedback links between the Amacrin cells and the Bipolar and Ganglion cells. However, due to the limited studies in such feedback mechanisms, it was challenging to build a computational model representing the behaviour of such feedback connections.

7.4 Recommendations for Further Work

The modular structure of the proposed approach to image enhancement makes it particularly suitable for extending its capabilities and performances. Each level of analysis can be independently extended and improved. Some directions for further work are proposed below.

- A future improvement to the spatial decomposition stage in the GRF model, by adding another layer. This layer represents the lateral inhibition cells in the retina, which is not represented in the proposed model as the Gap cells are responsible for high-speed signals. The Gap cells will impact the interconnectivity between the Bipolar receptive field and the photoreceptors in the new proposed retinal computational model.
- Furthermore, a feedback loop between the Horizontal cells and the Bipolar cells could be added. Also, a feedback loop between the Amacrin cells and the Ganglion cells could be introduced, as those neural circuits are presented in the retina.
- In the SSR estimation, the specular reflectance component was negated for simplicity, assuming diffuse reflectance only if SSR can be estimated by taking

both diffuse and reflectance reflections to give a more realistic colour re-orientation.

- The CE map produced by the proposed assessment method indicates local contrast variations, which could be beneficial in developing new contrast enhancement methods based on the proposed CE map.
- Apart from the results presented in this thesis, the proposed approaches have been tested on various images and videos. Different scenes have been used. The algorithms have been applied to many images and videos outside the scope of this work for testing purposes, showing a range of different indoor and outdoor scenes with different levels of complexity. Tests were conducted in image sequences containing targets of interest in various environments, e.g., offices, public buildings, subway stations, campuses, parking lots, airports, and sidewalks. All of the obtained results are consistent with those presented in the evaluation chapter. A normal extension is the application of such algorithms for other application areas.

Appendix A

The Retinal Anatomy

A.1. Introduction

The Human Visual System (HVS) is a dynamic retinal organization-dominated network; parallel processing, feed forward, feedback, and lateral connexons (Tuscany et al., 2014). The image processing in the HVS does not start in the brain but rather in the Retina which begins immediately. In addition, the retina and the optic nerve originate as an outgrowth of vertebrate embryonic development. Figure A-1 portrays a human eye and the cross-section of a retina. The retina is formed by many layers of neural cells, starting with about 130 million photoreceptors (rods and cones) and finishing with about 1 million ganglion cells. The basic processing that takes place in each cell type is complex and still not fully understood.

What we know for certain is that retinal cells can respond to stimuli in a nonlinear manner and are linked through links called synapses, which allow basic mathematical operations such as addition, subtraction, multiplication, division, amplification and gain control. Such operations are taken as a whole and result in a clever and sophisticated behaviour.

The total retina region is roughly $1,100\text{mm}^2$. The retina thickens in the macular region in the fovea to about $400\ \mu\text{m}$ and thins in the fovea to about $150\ \mu\text{m}$.

Retinal nerve fibres leave the eye via the optic nerve, nasally situated on the same plane as the retina's anatomical center. There is no retinal tissue that overlies the head of the optic nerve. The retina center provides the eye's greatest resolving power. This region is known as the macula which is responsible for central vision. The macula core is called the fovea. The retina's inner surface is parallel to the light that opens the eye. Tightly connected to the choroid is the outermost layer of the retina, the retinal pigment epithelium where the light sensors (Rods and Cones) reside. The photo sensors of the retina are reversed, not in the direction of light. The Retinal cells may be divided into three basic

types of cells, photoreceptor cells, neuronal cells, and glial cells as defined in (Andrew A Dahl, 2015) and as shown in Figure A-1:

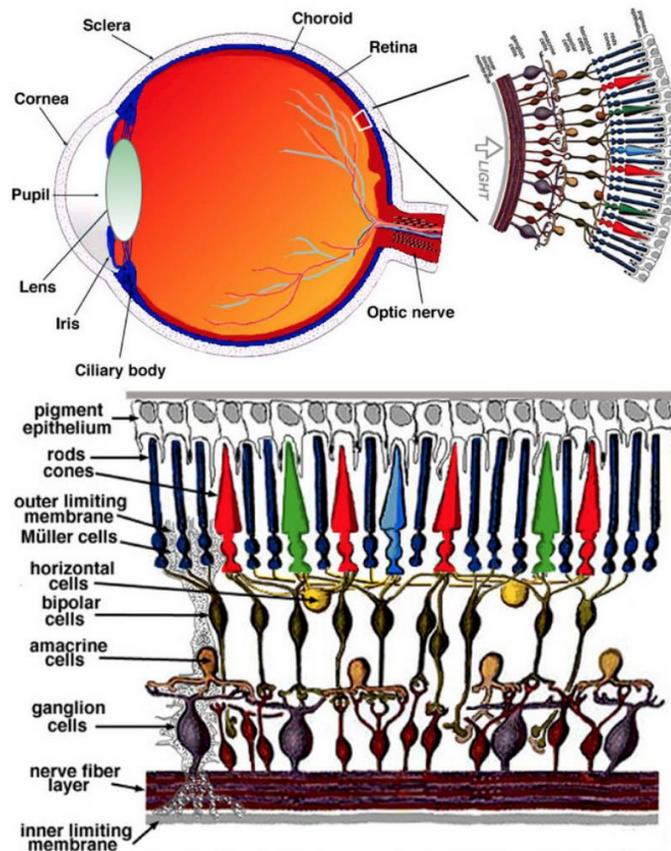


Figure A-1 Top: a human eye. Bottom: the cross-section of a human retina(Carreras, et al., 2017).

The photoreceptor cells primarily consist of cones and rods. Cones work best under bright conditions and offer vision of colour. Rods typically work in dim light and provide black and white vision. Growing human retina comprises roughly 120 million rods and 6 million photoreceptors for cones.

At the middle of the fovea lies the highest cone density. The middle of the fovea is without rods. The fovea, a pit where the cones are smallest and arranged in a mosaic to provide the maximum and most powerful optical density, is at the centre of the macula.

The retina includes a third class of photoreceptors (except for the rods and the cones). These are photosensitive retinal ganglion cells that are activated by light. The ganglion cells contain the pigment melanopsin which is photosensitive, those cells respond to light by regulating the pupil size and synchronize the circadian cycles.

Nerve cells include bipolar cells, ganglion cells, horizontal cells, and amacrine cells. Bipolar cells are completely enclosed within the retina, linking the photoreceptors to the ganglion cells. These are oriented vertically (perpendicular to the surface of the retinal tract). Nine forms of cells are bipolar. Bipolar cells are post-synaptic to rods and cones. Ganglion cells have dendrites that are synapsed with bipolar cells. The axons of the ganglion cells become a layer of nerve fibre within the retina and then become optic nerve fibres that end up within the brain. Horizontal cells connect bipolar cells to each other. They are the laterally interlinking neurons in the outer plexiform layer of the retina. Horizontal cells are responsible for allowing the eyes to adjust to seeing well under both light and dim light conditions. They are horizontally oriented (parallel to the retinal surface). Amacrine cells connect bipolar and ganglion cells with each other. Amacrine cells reside inside the inner plexiform layer, the second layer of retinal synaptic where synapses are produced by bipolar cells and retinal ganglion cells. There are approximately 40 different types of amacrine cells. Like horizontal cells, amacrine cells are horizontally oriented and work laterally, affecting the output of bipolar cells. Each amacrine cell type connects to a certain bipolar cell type and has a special neurotransmitter type.

A.2. Anatomic Layers of the Retina

Each of the retinal layers shown in Figure A-2 is given a name and contains different structures. Those layers are (starting from the innermost layer) as follows:

A.3. Neurological network function

The transformation of light into neuronal signals includes four basic processes: photoreception, transmission through synapses to bipolar cells, transmission to ganglion cells, and transmission along the optic nerve.

A.1.1. Photoreception

Focused (or unfocused) light will reach the photoreceptors (rods and cones) passing through the inner layers of the retina. Since the photoreceptive cells reside in the innermost layer of the retina, light first needs to pass through the ganglion cells and through the thickness of the retina until entering the rods and cones. The light does not penetrate the invisible pigment epithelium or choroid.

The outer segments of the rods and cones produce photo pigment, catching individual light photons and activating neural signalling. The inner segments of the photoreceptor include the axon terminal where neurotransmitters are released into the bipolar cells. The inside segments often work efficiently to channel light into the outer segments.

1. The inner limiting membrane
2. The nerve fibre layer,
3. The ganglion cells layer,
4. The inner plexiform layer
5. The inner nuclear layer,
6. The outer plexiform layer,
7. The outer nuclear layer,
8. The outer limiting membrane
9. The rod and cone layer
10. The Pigment eplithume

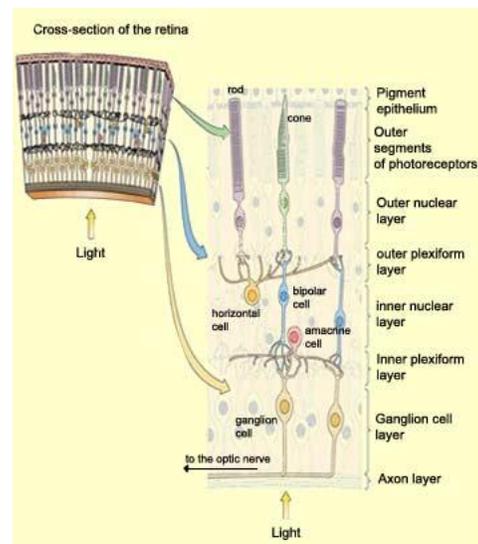


Figure A-2 Cross section of the retina (Dubuc, 2002)

Rods work mainly in dim light and give black-and-white vision, while cones support daytime vision and colour perception. The rods and cones are mainly positioned vertically inside the retina. Photoreceptors aren't evenly distributed across the retina. Most cones are in the fovea while peripheral vision is dominated by rods. The retina has 6–7 million cones. The whole retina is home to 125,000,000 rods. Figure A-3 shows the Rods-Cones distribution at the fovea.

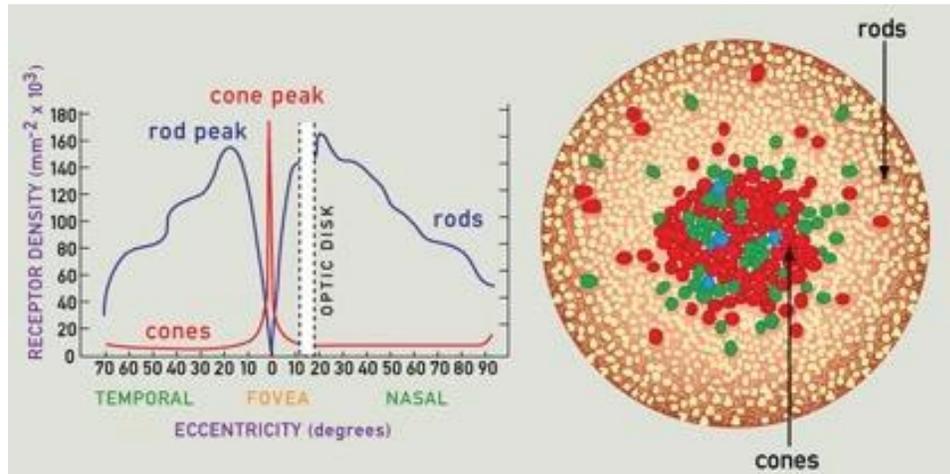


Figure A-3 Rod-Cones distribution in the fovea (A Dahl, 2015).

The macula is the Retina's central 3 mm. It has strong pigment supplied from the epithelium of the retinal pigment. The fovea is the middle part of the macula. It is situated 3-3.5 mm temporally up to the optic nerve head's temporal side. The fovea only contains cones inside a 500 μ m diameter rod-free region. The foveal cones have a diameter of 2,3 μ m and are very closely packed together.

At the middle of the fovea, the retina is thinnest, because there is no bipolar and ganglion cell here. Light directly strikes the receptors allowing for the best visual acuity. The bipolar cells which bind to these cones are displaced from the fovea concentratively.

A.1.2. Transmission to Bipolar Cells Through Synapses

The outer segments of the rods and cones transduce light and send the signal to their axons, through the cell bodies of the outer nuclear layer. Photoreceptor axons in the outer plexiform layer contact the dendrites of both the bipolar and horizontal cells.

Horizontal cells are horizontally oriented interneurons (parallel to the retinal surface) which aid in the processing of signals. The bipolar cells are oriented vertically (perpendicular to the retinal surface), as Figure A-4 shows.

Eight different types of bipolar cells synapse with cones. Five of these are called diffuse bipolar cells and up to 20 cones make synaptic contact. The other three forms only touch single cones, and they are called bipolar midget cells. Since there are 150 million photoreceptors and only 1 million optic nerve fibres, a convergence of signals and mixing of individual cone signals with others is needed. The horizontal action of the horizontal and amacrine cells may also enable one area of the retina to affect another.

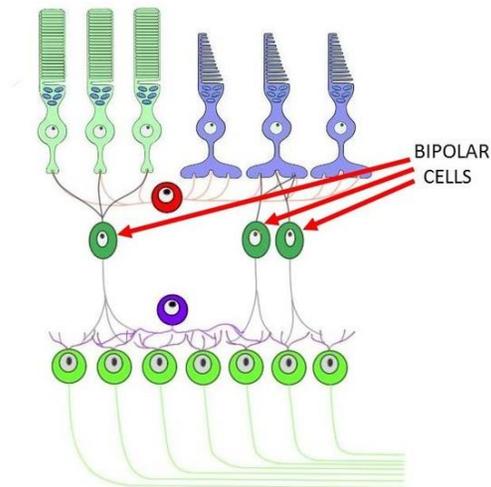


Figure A-4 Bipolar cells (A Dahl, 2015).

A.1.3. Transmission to Ganglion Cells

Information from photoreceptors and horizontal cells is given by the bipolar cells in the inner nuclear layer. The signal is sent to their axons. Bipolar axons contact dendrites of the ganglion cells and amacrine cells, another class of interneurons, through synapses in the inner plexiform layer. Ganglion cells are oriented vertically, while the amacrine cells are oriented horizontally.

A.1.4. Transmission along the Optic Nerve

The ganglion cells of the layer of ganglion cells send their axons through the layer of nerve fibres and converge to the middle of the retina at a point nasal, forming the optic nerve. The axons of the ganglion cells all set their eyes on the optic disk. Such axons fly all the way across the brain stem to the lateral geniculate nucleus.

No retinal photoreceptors, bipolar cells, ganglion cells, or accessory cells are present at the optic disk. Every single human optic nerve has 1,000,000 axons.

A.1.5. Receptive field

In the visual system, volumes in visual space are receptive fields, as shown in Figure A-5. The area of space (e.g., the visual field) through which a stimulus influences neuron firing (e.g., cone, bipolar cell, or ganglion cell) is called the photoreceptor's or neuron's receptive field. The receptive field is always determined relative to the given receptor or neuron. In other words, the stimulus is the activity we observe which affects the neuron's activity or state. If we test individual photoreceptor receptive fields, they'll all be identical in that they're determined by whether or not light hits the receptor. However, since many rods converge on their bipolar and ganglion cells while cones project in a 1:1 fashion, the receptive fields of rods will be larger than those of cones, consisting of the sum of the receptive fields of all the rods supplying input, as shown in Figure A-6.

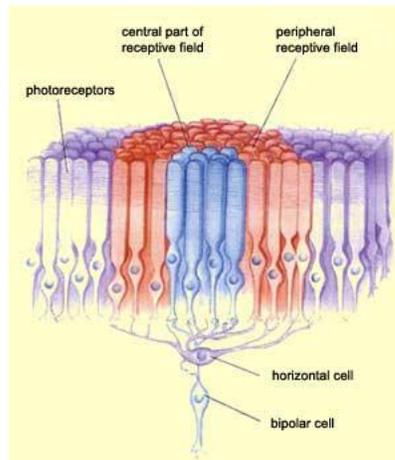


Figure A-5 Receptive field = centre + surround (A Dahl, 2015).

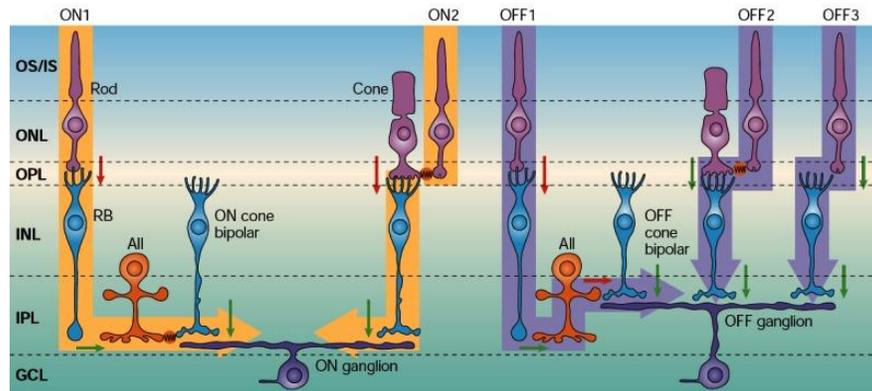


Figure A-6 Different Rods paths to the Ganglion cells

A retinal neuron's receptive field is the portion of the visual field that evokes a change in the behaviour or state of that specific neuron when stimulated. For example, in this hypothetical sequence, each photoreceptor's receptive field corresponds to the portion of the visual field with the same letter (e.g., photoreceptor A "sees" the head of the snake and photoreceptor J sees the rattles on its tail. Ganglion cell receptive fields are determined by the photoreceptor input pattern. For example, if we reported from ganglion cell 1, portions A and B of the visual field (head and upper body of the snake) will be included in its receptive region.

Ganglion cell 3 only gets input from photoreceptor E, so its receptive field is E (the last diamond-patterned segment of the snake's body). Ganglion cell 5 gets input from G, H, I and J, so it "sees" most of the black portion of the tail plus the rattles.

In particular, there are two bipolar / horizontal and ganglion / amacrine centre / surround layers; and four-color opponents, red (R), green (G), blue (B), and yellow (Y). The central cell (bipolar or ganglion) takes the surrounding information from one or more horizontal or amacrine cells; and both bipolar and ganglion have sub-types ON and OFF, as shown in Figure A-7. For example, a +R/-G bipolar (red-centre- ON/green-surround-OFF) will be excited if only the centre is illuminated or inhibited if only the surroundings (bipolar) are opponents with ON-centre/OFF-surround, +G/-R and +B/-Y, follow the same rules. The yellow (Y) channel can be obtained by averaging red and green channels.

On the other hand, Off-centre/ON-surround bipolar (i.e., -R/+G and -G/+R, but no - B/+Y) are inhibited when the centre is illuminated. An ON bipolar (or OFF-bipolar) only transfers signals to an ON ganglion (or OFF-ganglion), where amacrine provide surrounding information. (Lamb et al., 2007).

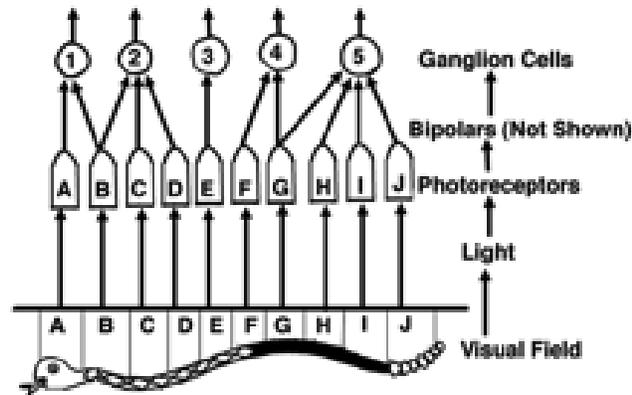


Figure A-7 Bipolar-cells-Ganglion cells interactions (Oguztoreli et al., 1985).

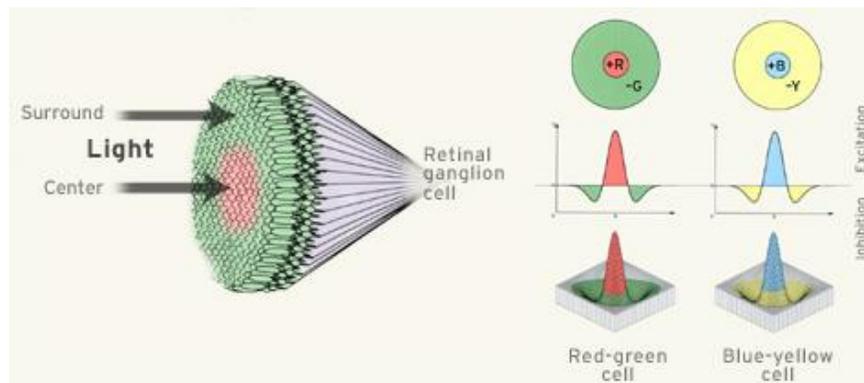


Figure A-8 Illuminated or stay neutral if both center and surroundings are illuminated. Likewise, other two-color

Ganglion cells do not actually relay the photoreceptor signals to the brain. The input signals from the photoreceptors undergo the initial processing in the ganglion cells. A ganglion cell's production is a sequence of voltage spikes (neural action potentials) that move down the optic nerve to the brain where further processing eventually contributes to the visual look. The output signal of the ganglion cells is determined not by the voltage value, but by the rate of voltage spikes. The means by which communication usually takes place in the brain is to fire up voltage spikes at different levels.

Many photoreceptor cells converge, particularly outside the fovea, into a single ganglion cell. It is because the collector's horizontal cell form interacts with multiple photoreceptors (Figure A-8). The means with which communication generally takes place in the brain is to fire up voltage spikes at different rates.

Many photoreceptor cells, particularly outside of the fovea, converge on a single ganglion cell. It is because the form of collector cells in horizontal cells connects with multiple photoreceptors (Figure A-8)The total area of the photoreceptor cells that converge on a given ganglion cell is called that cell's receptive field. Kuffler (1953) was the first to map receptive fields by radiating tiny light spots on the retina and by measuring electrical activity in ganglion cells using miniature electrodes.

The receptive fields grow larger farther down into the periphery of the retina, as more photoreceptors converge on a single ganglion cell. The bipolar cell connects the middle portion of the receptive field directly to the ganglion cell. The other photoreceptors of the same receptive field are connected more indirectly via a horizontal cell which interconnects the photoreceptors of the receptive field. This section of reception field is called the surrounding section.

Signals from the receptive field are not merely transmitted but are processed by the ganglion cell to facilitate scene interpretation. Light-dark patterns or contrasts are important for that interpretation, particularly at the edges of objects and surfaces.

Ganglion cells compare signals from an inner circular region of the receptive field with signals from the outer circular (surroundings) region of the same receptive field. Another type of ganglion cell increases its output ("excites" is the word usually used) when the centre circle of its receptive field is illuminated but its output decreases ("inhibits") when it illuminates the surroundings.

Figure A-10(b) shows the situation of the ON-centre ganglion cell when the light overlaps completely with both the centre and the receptive field surrounding.

The ganglion cell shows only spontaneous activity, just like in the case where there is no light Figure A-10(a)).

Figure A-10(c) demonstrates the condition of an ON-centre ganglion cell when it illuminates only the centre part of the receptive field. The ganglion cell fires a series of high-rate voltage spikes that pass into the brain (an ON response).

Opposite to this, when only the centre of the receptive field is illuminated by light (Figure A-10(d)), the firing of spikes stops completely (the OFF response). When both the centre and the surrounding area are partially illuminated (Figure A-10(e)), the spikes fire at a rate higher than the spontaneous rate, but at a rate lower than in the situation with all the light in the centre. The cell fires at a rate greater than the spontaneous rate if the middle and part of the surround are lighted (Fig. Figure A-10(f)).

There is another type that processes signals in the opposite direction, called the "off-centre ganglion cell." Light in the centre of the receiving field of an off-centre cell decreases or stops the spiking of spikes while light in the surrounding area increases the firing rate (Fig. Figure A-10(g)– I There are nearly equal percentages of ON-centre and Off-centre cells.

The ON-OFF retinal ganglion cells processing in the centre surround, described above, allows the detection of light-dark transitions and thus the edge detection of bright objects or light sources. This is shown in Figure A-10 as a light.

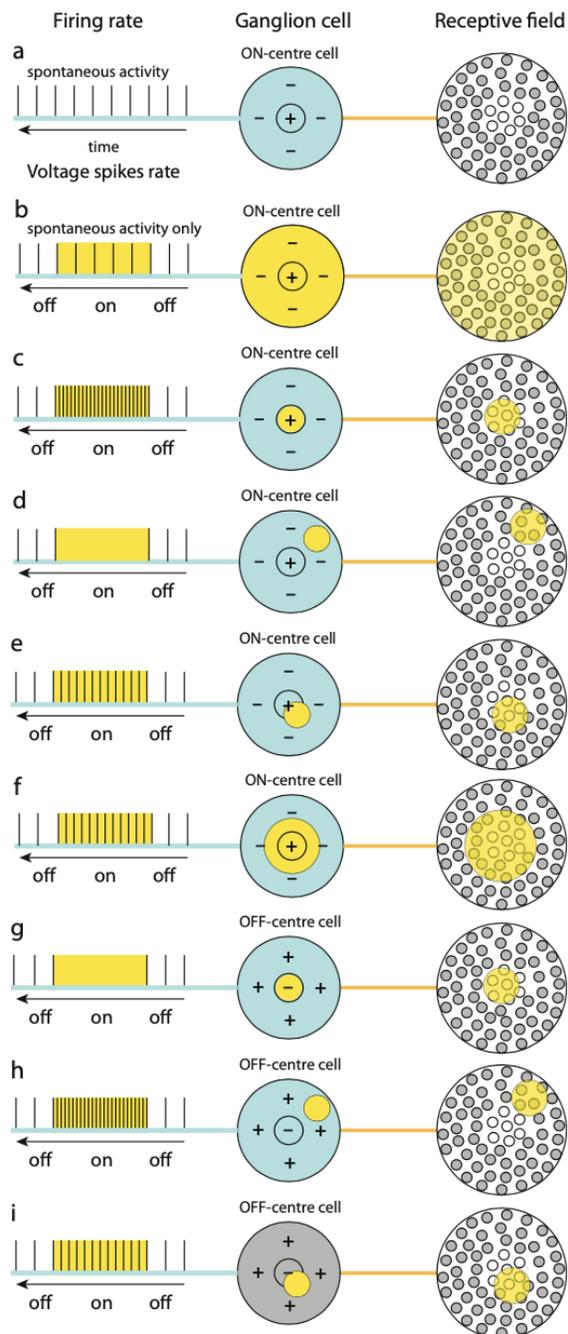


Figure A-9 Voltage firing rate of ganglion cell types ON-centre and Off-centre depending on a light spot (yellow circle) at different areas of the cell's receptive field. When the light spot is on, a yellow bar shows the voltage-spike indicator. (a) Rest (no light) situation with spontaneous activity; (b – f) ON-center type of ganglion cells; (g – i) off-center cell type

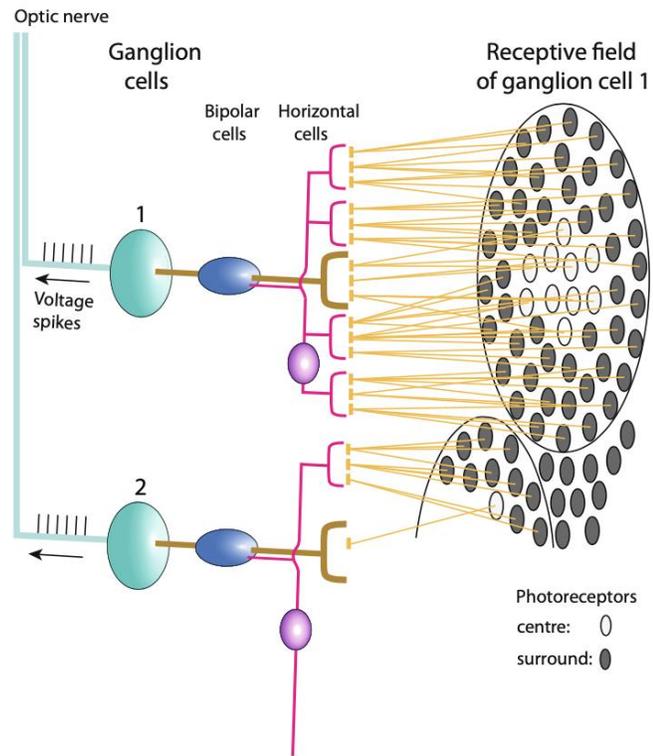


Figure A-10 Receptive fields of two ganglion cells

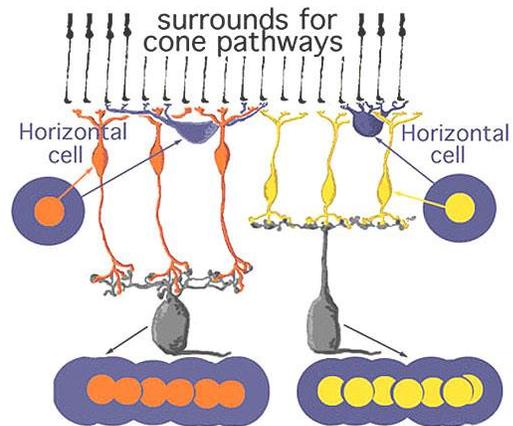


Figure A-11 Centre-surround organization using horizontal cells circuitry

Figure A-11 summarises the architecture for centre surround organisation in the mammalian retina's cone bipolar system by means of horizontal cell to cone bipolar cell circuits. The centre pathway is created by the cone to the bipolar to the ganglion cell through-channel, whereas the injection of horizontal cell information provides an antagonistic surrounding to the centre: an OFF-surround for the ON-centre channel (horizontal cell and orange bipolar, left hand pathway) and an ON-surround for the off-centre channel (horizontal cell and yellow bipolar, right hand pathway (Werblin and Dowling, 1979; Werblin, 1991)).

References

- Bajcsy, R., Lee, S. & Leonardis, A., 1990. Colour image segmentation with detection of highlights and local illumination induced by inter-reflections. *In: Proceedings of ICPR*, pp. 785-790.
- Barnard, K., Cardei, V. & Funt, B., 2002. A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data. *In: IEEE Transactions on Image Processing*, 11(9), pp. 972-984.
- Berns, R. S., Motta, R. J. & Gorzynski, M. E., 1993. CRT colorimetry. Part I: Theory and practice. *Color Research & Application*, 18(5), pp. 299-314.
- Berns, R. S., Motta, R. J. & Gorzynski, M. E., 1993. CRT colorimetry. Part I: Theory and practice. *In: Color Research & Application*, 18(5), pp. 299-314.
- Bio-inspired color image enhancement model. In Visual Information Processing XVIII (Vol. 7341, p. 7. I. S. f. O. a. P., 2009. Bio-inspired color image enhancement model. *International Society for Optics and Photonics.*, 7341(April), p. 73410D.
- Burt, P. J. & Adelson, E. H., 1983. The Laplacian Pyramid as a Compact Image Code. *In: IEEE Transaction on Communications*, 31(4), pp. 532-540.
- Cao, G., Zhao, Y., Ni, R. & Li, X., 2014. Contrast enhancement-based forensics in digital images. *IEEE transactions on information forensics and security*, 9(3), pp. 515-525.
- Cardei, V., Funt, B. & Barnard, K., 2002. Estimating the scene illumination chromaticity using a neural network. *In: Journal of the Optical Society of America A*, 19(12), pp. 2363-2373.
- Carreras, F., Delgado, Á., García-Serrano, J. & Medina-Quero, J., 2017. A virtual model of the retina based on histological data as a tool for evaluation of the visual fields.

Carreras, F., Delgado, Á., García-Serrano, J. & Medina-Quero, J., 2017. A virtual model of the retina based on histological data as a tool for evaluation of the visual fields.

Ciurea, F. & Funt, B., 2004. Tuning retinex parameters. *In: Journal of Electronic Imaging*, 13(1), pp. 58-64.

Dubuc, B., 2002. *The brain from top to bottom*. [Online] Available at: https://thebrain.mcgill.ca/flash/i/i_02/i_02_cl/i_02_cl_vis/i_02_cl_vis.html [Accessed 8 May 2021].

Finlayson, G. D., 1995. Color constancy in diagonal chromaticity space. *In: Proceedings of the Fifth International IEEE Conference on Computer Vision*, pp. 218-223.

Fu, X. et al., 2014. *A novel retinex based approach for image enhancement with illumination adjustment*. Florence, Italy, 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

Fu, X. et al., 2016. A fusion-based enhancing method for weakly illuminated images. *Signal Processing*, Volume 129, pp. 82-96.

Georgeson, M. A. & Sullivan, G. D., 1975. Contrast constancy: deblurring in human vision by spatial frequency channels. *In: The Journal of Physiology*, 252(3), pp. 627-656.

Guo, X. & Li, Y., 2016. LIME: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2), pp. 982-993.

Hautière, N., Tarel, J., Aubert, D. & Dumont, E., 2008. Blind contrast enhancement assessment by gradient ratioing at visible edges. *Image Analysis & Stereology*, 27(2), pp. 87-95.

Ho, J., Funt, V. & Drew, M. S., 1990. Separating a color signal into illumination and surface reflectance components: Theory and applications. *In: IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10), pp. 966-977.

- Hunt, R., 1982. A model of colour vision for predicting colour appearance. *Color research & application*, 7(2), pp. 95-112..
- Hunt, R., 2004. *The Reproduction of Color*. 6th ed. England: Kingstonupon-Thames, Fountain Press.
- Jan, Z. et al., 2018. A review on automated diagnosis of malaria parasite in microscopic blood smears images. *Multimedia Tools and Applications*, 77(8), pp. 9801-9826.
- Jan, Z. et al., 2018. A review on automated diagnosis of malaria parasite in microscopic blood smears images.. *Multimedia Tools and Applications*, 77(8), pp. 9801-9826.
- Joshi, P. & Prakash, S., 2020. Image enhancement with naturalness preservation. *The Visual Computer*, 36(1), pp. 71-83.
- Klinker, G. J., Shafer, S. A. & Kanade, T., 1990. A physical approach to color image understanding. *In: International Journal of Computer Vision*, 4(1), pp. 7-38.
- Land, E. H., 1986. Recent advances in Retinex theory. *In: Vision research*, 26(1), pp. 7-21.
- Land, E. H. & McCann, J., 1971. Lightness and retinex theory. *In: Journal of the Optical Society of America A*, 61(1), pp. 1-11.
- Laughlin, S., 1981. A simple coding procedure enhances a neuron's information capacity. *Zeitschrift für Naturforschung c*, 36(9-10), pp. 910-912.
- Lee, C., Lee, C. & Kim, C., 2012. Contrast enhancement based on layered difference representation. *19th IEEE International Conference on Image Processing* , pp. 965-968.
- Maini, R. & Aggarwal, H., 2010. *A comprehensive review of image enhancement techniques*.. s.l.:arXiv preprint arXiv:1003.4053..
- Majumder, A. & Irani, S., 2006. Contrast enhancement of images using human contrast sensitivity. *In Proceedings of the 3rd symposium on Applied perception in graphics and visualization, ACM* , pp. 69-76.

Ma, K., Zeng, K. & Wang, Z., 2015. Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, 24(11), p. 3345–3356.

Maloney, T. L., 1986. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *In: Journal of the Optical Society of America*, 3(1), pp. 1673-1683.

Marimont, D. H. & Wandell, B. A., 1992. Linear models of surface and illuminant spectra., vol. 3, pp. .. *In: Journal of the Optical Society of America A*, 3(1), pp. 1673-1683.

Marr, D. & Vision, A., 1982. *A computational investigation into the human representation and processing of visual information..* San Francisco: WH San Francisco: Freeman and Company.

McCamy & S., C., 1992. Correlated colour temperature as an explicit function of chromaticity coordinates. *In: Journal of Colour Research & Application*, 17(2), pp. 142-144.

McCann, J., 1983. *Method and apparatus for lightness imaging.* United States, Patent No. 4384336.

Morel, J. M., Petro, A. B. & Sbert, C., 2010. A PDE formalization of retinex theory. *In: IEEE Transactions on Image Processing*, 19(11), pp. 2825-2837.

Ortiz, S. H. C., Chiu, T. & Fox, M. D., 2012. Ultrasound image enhancement: A review. *Biomedical Signal Processing and Control*, 7(5), pp. 419-428.

Parkkinen, J. P., Hallikainen, J. & Jaaskelainen, T., 1989. Characteristic spectra of Munsell colors. *In: Journal of the Optical Society of America A*, 6(2), pp. 318-322.

Pattanaik, S. et al., 1998. A multiscale model of adaptation and spatial vision for realistic image display. *In Proceedings of the 25th annual conference on Computer graphics and interactive techniques.*

- Pattanaik, S. N., Ferwerda, J. A., Fairchild, M. D. & Greenberg, D. P., 1998. A multiscale model of adaptation and spatial vision for realistic image display. *In: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pp. 287-298.
- Petro, A., Sbert, C. & Morel, J., 2014. Multiscale retinex. *Image Processing On Line*, pp. 71-88.
- Rahman, M. A. et al., 2018. A review on brightness preserving contrast enhancement methods for digital image. *In Ninth International Conference on Graphic and Image Processing (ICGIP 2017)*, Volume 10615, p. 106152S.
- Ramesh, V., 2003. A class of photometric invariants: Separating material from shape and illumination.. *In Proceedings Ninth IEEE International Conference on Computer Vision*, pp. 1387-1394.
- Reddy, A. A. et al., 2013. Comparison of image enhancement techniques using retinex models. *In: International Journal of Advanced Computer Engineering and Communication Technology*, 2(3), pp. 7-12.
- Ren, Y., Ying, Z. & Li, 2018. LECARM: low-light image enhancement using the camera response model. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(4), pp. 968-981.
- Rundo, L. et al., 2019. MedGA: a novel evolutionary method for image enhancement in medical imaging systems.. *Expert Systems with Applications*, Volume 119, pp. 387-399.
- Rundo, L. T. A. N. M. S. M. C. B. D. M. G. & C. P. (. M., n.d.
- Saba, T. et al., 2018. Image enhancement and segmentation techniques for detection of knee joint diseases: A survey.. *Current Medical Imaging Reviews*, 14(5), pp. 704-715.
- Saba, T. et al., 2018. Image enhancement and segmentation techniques for detection of knee joint diseases: A survey. *Current Medical Imaging Reviews*, 14(5), pp. 704-715.

- Saichandana, B. & Ramesh, S., 2014. Image Fusion Technique for Remote Sensing Image Enhancement. *In: ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India*, Volume 2, pp. 235-242.
- Saichandana, B., Ramesh, S., Srinivas, K. & Kirankumar, R., 2014. Image fusion technique for remote sensing image enhancement. *In ICT and Critical Infrastructure Proceedings of the 48th Annual Convention of Computer Society of India*, Volume II , pp. 235-242.
- Sapiro, G., 1999. Colour and illuminant voting. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 21(11), pp. 1210-1215.
- Sarage, G. & Jambhorkar, S., 2011. Enhancement of Mammography Images for Breast Cancer Detection using Histogram Processing Techniques 1.
- Sedky, M., Moniri, M. & Chibelushi, C., 2014. Spectral-360: A Physics-Based Technique for Change Detection. *In: IEEE Conference In Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 405-408.
- Shen, J., 2003. On the foundations of vision modeling: I. Weber's law and Weberized TV restoration.. *Physica D: Nonlinear Phenomena*, 175(175), pp. 241-251.
- Stanford, 2021. *CS231n: Convolutional Neural Networks for Visual Recognition*. [Online] Available at: <http://cs231n.stanford.edu> [Accessed May 2021].
- Subramanyam, B., Joshi, P., Meena, M. & Prakash, S., 2016. Quality based classification of images for illumination invariant face recognition. *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA) (pp.)*, pp. 1-6.
- Valliammal, N. & Geethalakshmi, S. N., 2011. A hybrid method for enhancement of plant leaf recognition. *World of Computer Science and Information Technology Journal*, 1(9), pp. 370-375.

- Wang, S. & L., G., 2017. Naturalness preserved enhancement algorithm using a priori multi-layer lightness statistics. *IEEE Transactions on Image Processing*, 27(2), pp. 938-948.
- Wang, S. et al., 2015. A patch-structure representation method for quality assessment of contrast changed images. *IEEE Signal Processing Letters*, 22(12), pp. 2387-2390..
- Wang, S., Zheng, J., Hu, H. & Li, B., 2013. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, 22(9), pp. 3538-354.
- Wang, S. Z. J. H. H. a. L. B., 2013. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, 22(9), pp. 3538-3548.
- Watson, A. B. & Solomon, J. A., 1997. Model of Visual Contrast Gain Control and Pattern Masking. In: *Journal of the Optical Society of America A*, 14(9), pp. 2379-2391.
- Whittle, P., 1986. Increments and decrements: luminance discrimination. *Vision research*, 26(10), pp. 1677-1691.