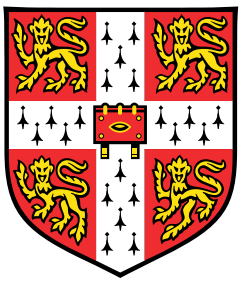
Perception-Aware Optimisation Methodologies for Quantum Dot Based Displays and Lighting



Chatura Theekshana Bandara Samarakoon

Supervisors: Prof. G. A. J. Amaratunga
Prof. J. M. Kim

Advisor: Prof. P. Stanley-Marbell

Department of Engineering
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Girton College

March 2021

Declaration

I hereby declare that, except where specific reference is made to the work of others, the content of this report are original and have not been previously submitted in whole or in part for the consideration of any other qualification or degree at the University of Cambridge or any other institution. This report is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements. This report contains fewer than 65,000 words.

Chatura Theekshana Bandara Samarakoon

March 2021

Perception-Aware Optimisation Methodologies for Quantum Dot Based Displays and Lighting

Chatura Theekshana Bandara Samarakoon

Human colour vision acuity is limited. This presents opportunities to leverage these perceptual limits to achieve engineering optimisations for devices and systems that interact with the human vision system. This dissertation presents the results of few investigations we carried out into quantifying these limits and several optimisation methodologies that we devised.

The first step in this process is to quantify the acuity of human colour vision. We obtained a large corpus of colour matching data from a mobile video game called *Specimen*. We examine what questions about human vision this dataset allows us to answer and explore global statistics about colour vision based on this data on 41,000 players from 175 countries. We show that we can use the information in this dataset to infer potential candidate functions for the spectral sensitivities of each person in the dataset.

The human eye acts like a many to one function; quantifiably different spectra can look like the same colour. This is referred to as metamerism. From a device perspective, different spectra consume different amounts of energy to generate. We show that we can use these two properties to elicit the same colour sensation using less energy. In the colour samples we evaluated, we show that we can achieve up to 10 times less power consumption while achieving a colour match. Given that one cannot change the emission spectrum of a display after fabrication, we propose the use of a multi-primary colour display to achieve this. We present two indices for quantifying the metameric capacity of such a display and its ability to save energy. The emission spectrum of a quantum dot (QD) based device is very narrow. Previous work in the literature suggested that narrow bandwidth spectra can lead to observer metameric breakdown; different observers disagreeing on the perceived 'colour' of a spectrum. We show that this might not be the case, using modern colour science tools, and show how metameric breakdown in a display could be minimised by carefully choosing the primary emission wavelengths.

The limited colour acuity of human vision implies that people cannot notice small differences in colour. This fact has been used to create approximate colour transformation algorithms that subtly change colours in images such that they consume less energy when displayed on an emissive pixel display without causing unacceptable visual artefacts. We conducted a user study to gather information about the effect of one such colour transform called *Crayon*. We present a method for effectively picking the optimal transform parameters for *Crayon*, based on the user study results. The method presented calculates these parameters based on the properties of the image being transformed such that the power saving can be maximised while minimising the loss of image quality. The user study results show that we can achieve up to 50% power saving with a majority of the study participants reporting a negligible degradation in image quality in the transformed images.

We additionally investigate a hypothesis that was presented stating that images with large amounts of highly luminous pixels cause increased power consumption in OLED displays due to localised display

heating. We show that this hypothesis is wrong. We also investigate if sub-pixel rendering in *Pentile* displays can be used to reduce display power consumption by intentionally turning off random sub-pixels. However, we present a negative result showing that even single-pixel artefacts are observable on the test platform and thus, this cannot be used to improve display power efficiency.

The narrow-band optical emissions of QD based devices mixed with their ability to be fabricated through solution processing can be used to mix multiple QDs together to build devices that generate arbitrary spectral shapes. We show how to use this property in an numerical optimisation based design framework to create lighting devices with a high colour rendering index (CRI). We evaluate the effects of different cost functions and initialisation strategies, and show that, we are able to design devices with a CRI > 96 using only four different QD primaries. We use a charge-transport based simulator to asses the electric properties of the designed devices. We also showcase initial work done on a modular software interface and a material library we developed for this simulator.

To my parents... for teaching me the value of knowledge;

my grandparents... for teaching me the value of wisdom;

and my brother... for giving me a reason to strive for more of both.

Acknowledgements

This work was supported by the *EPSRC* under the DTP studentship award RG84040:EP/N509620/1. The work outlined in chapters 3 and 5 was partly funded by the Alan Turing Institute award TU/B/000096 and the EPSRC grant EP/N510129/1. The work presented in chapter 6 was carried out under the guidance of Dr. Sung-Min Jung. The charge-transport based quantum dot simulator described and used in this chapter was designed and implemented by Dr. Jung [1]. Figures 2.7 and 3.1 were made by Prof. Stanley-Marbell and have been used in this thesis with his permission. The LaTeX template used to format this thesis was based on the *CUED PhD Thesis Template* by Krishna Kumar et al. [2].

Finally, I would like to acknowledge everyone who helped me along this journey and to give them my most sincere gratitude. In particular...

- My parents and grandparents; for nurturing my curiosity, facilitating my undergraduate studies, and keeping me grounded.
- *Prof. Amaratunga and Prof. Kim*; for guiding my research and for providing financial support.
- *Prof. Stanley-Marbell*; for lending/recommending books, providing moral support, helping distil ideas, and keeping me focussed on graduating.
- *Dr. Richard Hopper*; for the frequent excursions to the Cambridge countryside and reminding me what the sun looks like.
- *Dr. Young-Tea Chun*, *Dr. Sung-Min Jung*, and *Dr. Felix Mocanu*; for the guidance and support they provided over the years.
- *Janith*; for being my intellectual punching bag... that occasionally punches back.
- *My brother*; for helping me digitise data tables, and keeping me up to date with the memes.
- *Hannah*; for providing a non-technical perspective on the writing and for telling me off when I was being an idiot about my work.
- *Imani, Josie, Eshana, Faria*, and the rest of my friends; for putting up with my complaining about being perpetually tired and lost, and the never-ending long nights writing.
- *Ruchini*; for believing in me, and pushing me to start this journey and pursue my dreams.
- To everyone else, that read through my messy writing and pointed out confusing points, and supported me in numerous other ways along this journey.
- Finally, the Sainsbury's frozen food section; for sustaining me through the long nights of simulations and writing.

To the reader: I hope you have an easier time reading this than I had writing it.

Publications and Preprints

Modelling Charge Transport and Electro-Optical Characteristics of Quantum Dot Light-Emitting Diodes. *2021*

S.M. Jung, T.H. Lee, S.Y. Bang, S.D. Han, D.W. Shin, S. Lee, H.W. Choi, Y.H. Suh, X.B. Fan, J.W. Jo, S. Zhan, J. Yang, C. Samarakoon, Y. Kim, L. Occhipinti, G. Amaratunga, & J.M. Kim.

Appears in npj Computational Materials, Vol 7, 2021.

Technology progress on quantum dot light-emitting diodes for next-generation displays. *2021*

S.Y. Bang , Y.H. Suh , X.B. Fan , D.W. Shin , S.H. Lee , H.W. Choi , T.H. Lee , J. Yang , S. Zhan , W. Harden-Chaters , C. Samarakoon , L.G. Occhipinti , S.D. Han , S.M. Jung, & J.M. Kim.

Appears in Nanoscale Horizons, Vol 6, 2021.

Inferring Human Observer Spectral Sensitivities from Video Game Data. *2020*

C. Samarakoon, G. Amaratunga, & P. Stanley-Marbell.

Work in progress. ArXiv preprint: 2007.00490 [q-bio.QM].

Content-Aware Automated Parameter Tuning for Approximate Color Transforms. *2020*

C. Samarakoon, G. Amaratunga, & P. Stanley-Marbell.

Appears in the proceedings of *Mobile HCI, 2020 (Extended Abstracts)*. ArXiv preprint: 2007.00494 [cs.HC].

Table of Contents

List of Figures **xvii**

List of Tables **xxvii**

List of Abbreviations **xxix**

1 Introduction **1**

2 Human Colour Vision, Colourimetry, and Display Technologies **5**

2.1 Introduction 5
2.2 What is Light? 6
2.3 What is Colour? 6
2.4 Transmission of Light 7
2.5 Biology of Vision 8
2.5.1 The Eye 9
2.5.2 Visual Processing in the Brain 10
2.6 Radiometry and Photometry 11
2.6.1 Key Radiometric Quantities 11
2.6.2 Key Photometric Quantities 12
2.7 Colourimetry and Colour Vision Models 13
2.7.1 Colour Matching Experiments 13
2.7.2 CIE-1931 2° Standard Observer and Associated Colour Spaces 14
2.7.3 CIE-1931 xyY Chromaticity Diagram 15
2.7.4 CIE-1964 XYZ Colour Space 16
2.7.5 CIE-2006 Physiologically-Relevant CMFs 16
2.7.6 Uniform Colour Spaces: CIE-UCS, CIELAB, and CIEUVW 17
2.7.7 CIE Standard Illuminants 19
2.7.8 Metamerism and Metameric Breakdown 19
2.8 Non-CIE Colour Models 20
2.9 Semiconductor-based Lighting and Display Technologies 21
2.9.1 Bulk-Semiconductor Light Emitting Diodes (LEDs) 21
2.9.2 Organic Light Emitting Diodes (OLEDs) 23

2.9.3 Quantum-Dot Light Emitting Diodes (QDs) . . . 24
2.9.4 Display Types . . . 26
2.10 Contextual Summary . . . 28

3 Individual CMF Inference from a Global Colour Matching Dataset 29
3.1 Introduction . . . 29
3.1.1 Contributions . . . 30
3.1.2 The *Specimen Dataset* . . . 30
3.2 Global Statistics . . . 32
3.2.1 Population Adjusted Player Counts . . . 32
3.2.2 Mean Play Time . . . 33
3.3 Player Behaviour (Players from the UK) . . . 34
3.3.1 Time Dependency . . . 34
3.4 Colour Acuity . . . 36
3.4.1 Sensitivity to *hue* . . . 36
3.4.2 Sensitivity to *saturation* and *Value* . . . 38
3.4.3 Colour Confusion Variation Over Time . . . 39
3.5 Individual CMF Inference from Specimen Data . . . 43
3.5.1 Motivation . . . 43
3.5.2 Problem Formulation . . . 45
3.5.3 Methodology . . . 46
3.5.4 Sensitivity to Regularisation . . . 48
3.5.5 Effect of the CMF Prior . . . 49
3.6 CMF Inference with a Psychologically-Relevant Observer Prior . . . 51
3.6.1 Methodology . . . 53
3.6.2 Results . . . 54
3.7 Related work . . . 55
3.8 Conclusions and Future Work . . . 56

4 Metamerism and Metameric Breakdown 59
4.1 Introduction . . . 59
4.1.1 Contributions . . . 60
4.2 Metamere Search on Hypothetical Three-primary Displays . . . 60
4.2.1 Experiment Setup . . . 61
4.2.2 Simulation Results . . . 62
4.3 A Mathematical Framework for Multi-Primary Displays . . . 72
4.3.1 Mathematical Framework . . . 73
4.3.2 Measure of Metameric Capacity and Energy Saving Potential . . . 73
4.3.3 Advantages of Multi-Primary Displays . . . 74
4.4 Metameric Breakdown in QD-LED Displays . . . 75

4.4.1 Metameric Breakdown Sensitivity to FWHM . . . 75
4.5 Conclusions and Future Work . . . 78

5 Approximate Colour Transformations 81
5.1 Introduction . . . 81
5.1.1 Operating Principles of Approximate Colour Transforms . . . 81
5.1.2 Contributions . . . 81
5.2 What is *Crayon*? . . . 82
5.3 How does the Crayon Colour Transform work? . . . 82
5.3.1 Mathematical Formulation of Crayon . . . 83
5.3.2 Least-Squares (ℓ_2^2) Solution . . . 83
5.3.3 Euclidean-distance (ℓ_2) Solution . . . 84
5.3.4 Image Colour Spaces . . . 85
5.3.5 Multi-Primary Displays . . . 85
5.3.6 Controlling the Transform . . . 86
5.4 The Original *Crayon* User Study . . . 86
5.5 Contemporary *Crayon* user studies . . . 88
5.6 User Study 1: Internal User Study . . . 89
5.6.1 Experimental Setup . . . 89
5.6.2 User Study Results . . . 90
5.6.3 Score Correlation with Lambda (Λ) . . . 90
5.6.4 Image Power Saving Correlation with Score . . . 90
5.6.5 Correlation with Simple 'Image Quality Metrics' . . . 93
5.6.6 Subjective Assessment of the 'Best' and 'Worst' Images . . . 94
5.6.7 Sensitivity of Power to Image content . . . 98
5.6.8 Factoring Out User Bias . . . 99
5.6.9 Predicting Score . . . 100
5.7 User Study 2: Redesigned Public-study . . . 102
5.7.1 Lambda (Λ) Normalisation . . . 103
5.7.2 Adjusted Perceptual Quality Scales . . . 104
5.7.3 Experiment Setup . . . 104
5.7.4 User Study Results . . . 105
5.8 Predicting Optimal Transform parameters . . . 106
5.8.1 Lambda (Λ) Lower Bound Calculation . . . 108
5.9 Applying the modified Crayon framework . . . 113
5.10 Spatial Dependence of Display Power . . . 116
5.10.1 Methodology . . . 116
5.10.2 Results . . . 117
5.11 Conclusions and Future Work . . . 121

6 Optimal Design of QD based High CRI Smart Lights **123**
6.1 Introduction . . . 123
6.1.1 What is CRI? . . . 124
6.1.2 Contributions . . . 125
6.2 Design Framework . . . 125
6.2.1 Numerical Optimisation . . . 125
6.2.2 Device and Process Simulation . . . 126
6.2.3 Validation Through Device Fabrication . . . 126
6.3 Finding Optimal QD Primary Combinations for Maximum CRI . . . 126
6.3.1 Methodology . . . 127
6.3.2 Results . . . 130
6.3.3 Optimised Primary Spectrum . . . 130
6.4 Charge-Transport Model Simulation . . . 133
6.4.1 Charge-Transport Model . . . 134
6.4.2 Calculating QD Device Parameters . . . 136
6.4.3 Band Gap Calculation . . . 136
6.4.4 Diameter Calculation . . . 136
6.4.5 Simulation Results . . . 136
6.5 QD Simulator Interface and Material Library . . . 139
6.6 Conclusions and Future Work . . . 142

7 Conclusions and Future Work **143**

References **147**

Appendix A OLED Power Characterisation and Modelling **155**
A.1 Introduction . . . 155
A.2 Power measurement . . . 155
A.2.1 Methodology . . . 156
A.2.2 Results . . . 157
A.2.3 Finding *Crayon* Model parameters . . . 161
A.2.4 Temperature dependence of power . . . 162

Appendix B Supplimentary Results from the Crayon User Studies **165**
B.1 Predicting user score . . . 165

Appendix C Tabulated results of the CRI optimisation process **169**

Glossary **173**

List of Figures

2.1 Light spectra can have arbitrary shapes depending on the physics underlying the light generating processes (data from [3, 4]). 7
2.2 Light from the source travels through a medium to the observer. The medium and objects in the light path can absorb and reflect different segments of the spectrum in different ways, affecting the light spectrum arriving at the observer. 7
2.3 The structure of the human eye captures light from the environment focusses it onto the retina (from [5]). 8
2.4 Photoreceptive cells and Ganglion cells in the retina converts the light into electrical signals and sends it to the brain (from [6]). 8
2.5 Human *scotopic* vision is nearly three times more sensitive to light than *photopic* vision (after [7]). 9
2.6 The three types of cone cells have overlapping bandpass sensitivities to different spectral ranges (after [8]). 9
2.7 Traditional colour matching experiments involve changing the intensities of 'long' (L), 'medium' (M), and 'short' (S) wavelength primaries till the mixture (❷) is perceptually identical to the monochromatic light (❶). 13
2.8 CIE-1931 2° RGB CMFs were obtained by averaging the results of colour matching experiments with multiple participants. 14
2.9 CIE-1931 2° XYZ CMFs were calculated by transforming the RGB CMFs to get rid of the negative values. 14
2.10 Projection of the CIE-1931 RGB colour space onto the r-g plane shows that some colours require a negative proportion of the red primary to achieve a colour match (from [9]). . 15
2.11 CIE-1931 xy colour space encapsulates all the chromatic sensations the human eye can perceive, in a strictly positive coordinate space. 15
2.12 MacAdam Ellipses (shown as 10x *just noticeable difference*) correspond to the set of colours around a colour where observers cannot discern the difference [10, p. 81]. . . . 17
2.13 In the CIE-1976 UCS the *MacAdam ellipses* are *more* uniform in eccentricity [10, p. 82]. 17
2.14 *RGB* spaces represent the colours that can be achieved by an additive mix of red, green, and blue coloured primaries (from [11]). 21

2.15 The *HSV* space is a colour representation where the different coordinates have an intuitive relationship to the appearance of a colour (from [12]). 21
2.16 Although most of the transitions in direct band gap semiconductors occur at $h\nu_1$, there are transition states around it that also lead to radiative recombination (e.g., $h\nu_2$) (after [13, p. 350]). 22
2.17 A P-N junction has a built-in potential (top, V_{bi}). This potential can be overcome by an external voltage (V) which can inject carriers into the depletion region (middle regions with sloped lines) where radiative recombination can occur (bottom, $h\nu$). The vertical axis is energy (after [13, p. 350]). 22
2.18 The internal structure of an LED. 23
2.19 The internal structure of an OLED. 23
2.20 The internal structure of a QD-LED. 23
2.21 QDs are semiconducting nanoparticles that can confine electrons in three dimensions, comprising a molecular core surrounded by ligands (from [14]). 25
2.22 The band structure of a QD is quantised, which restricts the allowed electron transitions and the band gap is a function of the QD size (after [15, p. 3] and [16]). 25
2.23 QDs can excited (a) optically or (b) electrically. 25
2.24 LCDs are composed of a backlight, layers of optical polarisers, colour filters, and an electrically controllable layer of liquid crystals (LC) (after [17, 18]). 26
2.25 Unlike LCDs, OLED/QD-LED displays have a simpler structure as they do not require a complex optical assembly to control the light emission (EL= Emissive layer) (after [17, 18]). 26

3.1 In *Specimen*, the players pick which coloured blob (*'specimen'*) matches the background colour. 29
3.2 Number of players by Country ((a) on a linear scale, (b) on a logarithmic scale, (c) player counts adjusted by country population). 35
3.3 Mean Playtime by country shows that most players played the game for around 2 minutes at a stretch. 36
3.4 Frequency of selection events follows a daily cycle with the player activity peaking around midnight. 36
3.5 The rate of incorrect colour matches, as a proportion of the number of selections, is minimally variant over the course of the day. 36
3.6 Mean session length is around 2 minutes, with players playing after midnight playing for approximately 1 minute longer than average. 37
3.7 Median session length shows that the mean session length variation after midnight is biased by a few players playing significantly longer than average. 37

3.8 Overall, the players more easily confused ‘green’ and ‘blue’ hues compared to others (right). Each bar spans 5 *hue* units and the dashed vertical lines indicate the hues of the *sRGB* primaries. . . . 37
3.9 The ‘second-order’ confusion distribution shows that ‘green’ hues are more likely to be confused repeatedly. Each bar spans 5 *hue* units and the dashed vertical lines indicate the hues of the *sRGB* primaries. . . . 38
3.10 The distribution of the colour *value* and *saturation* of the colours presented by the *Specimen* game. . . . 38
3.11 Colour confusions normalised by how frequently each colour *value*/*saturation* was presented. . . . 39
3.12 The relative proportion of confused hues is not substantially affected by the time of day (each line spans 20 units of *hue* from the specified value). . . . 39
3.13 Top 10 most confused hues are consistently confused regardless of the time of day (each point spans 5 units of *hue*). . . . 40
3.14 The relative proportion of confused colours by *saturation* is not substantially affected by the time of day. Each line spans 20 units of *saturation*, except the line for ‘230’ which represents the values between 230 and 255. . . . 41
3.15 Top 10 most confused colours by *saturation* does not vary substantially over the day (each point spans 5 units of *saturation*). . . . 41
3.16 The relative proportion of confused colours by *value* is not substantially affected by the time of day. Each line spans 20 units of *value*, except the line for ‘230’ which represents the values between 230 and 255. . . . 42
3.17 Top 10 most confused colours by *value* does not vary substantially over the day (each point spans 5 units of *value*). . . . 42
3.18 The shapes of the ‘red’, ‘green’, and ‘blue’ primary emissions of the *iPhone X* display were extracted from its white point spectrum [19]. . . . 44
3.19 Due to the narrow emission bandwidths there is limited information at some wavelengths (e.g., around 500 nm and 580 nm). . . . 44
3.20 Flow chart of the initial architecture used to learn the spectral correction term through a numerical optimisation process. . . . 44
3.21 Flow chart showing how the initial architecture is modified to incorporate a physiologically relevant model of the eye. . . . 44
3.22 Without regularization the optimization overfitted and resulted in calculated values that are physically impossible. . . . 48
3.23 *Sum of Absolute Values* regularization led to localised peaky artefacts. . . . 49
3.24 *Mean absolute* regularization caused *peaky* artefacts. . . . 49
3.25 *Max absolute* regularization led to slightly smoother curves than non-regularized optimization. . . . 49

3.26 *Mean Squares* regulariser resulted in a smooth set of CMFs but showed more fluctuation around the CMF prior compared to the *sum of squares* regulariser. 49
3.27 Optimizing with *sum of squares* regularizer show that both 2° and 10° standard observer CMF were adjusted to explain the observed colour confusions better. 50
3.28 The 2° standard observer CMFs were approximated with three Gaussians. 51
3.29 The CMFs learnt with the gaussian approximation were similar to those obtained when using the 2° standard observer prior. 51
3.30 The optimisation process is able to learn spectral sensitivities that better explain the observed colour confusions. 54
3.31 Density plots of the colour confusions by players C and D show significant overlap. . . 56

4.1 The metamere search experiment was conducted on the set of all possible three-primary spectral shapes with fixed FWHM and Gaussian shaped primary emissions. 60
4.2 The simple LED energy approximation follows a quadratic shape (based on data from [20]). 60
4.3 The chromaticities of the metameres closely matched the corresponding target colours as can be seen on the CIE-1931 xy chromaticity diagram. 62
4.4 Mean spectrum (blue) and variance (red) plot for the *Turquoise* colour sample show that some frequencies are more fundamental to generating a colour. 65
4.5 Energy plots for the *Turquoise* colour sample show that there is a five-fold difference between the maximum and the minimum energies required to generate this colour (red = power of the initial target spectrum). 65
4.6 Metamere energy is correlated most strongly with the medium primary intensity for the *Turquoise* colour metameres. Energy is in arbitrary units. 66
4.7 Mean spectrum (blue) and variance (red) plots for the *Yellow* colour sample show that there are two 'fundamental' frequencies used in generating this colour. 66
4.8 Metamere energy is correlated most strongly with the medium primary intensity for the *Yellow* colour metameres. Energy is in arbitrary units. 66
4.9 Energy plots for the *Yellow* colour sample show that there is a ten-fold difference between the maximum and the minimum energies required to generate this colour (Red = power of the initial target spectrum). 67
4.10 Mean spectrum (blue) and variance (red) plots for the *Pink* colour sample show that there are two 'fundamental' frequencies used in generating this colour and the overall shape looks like a mix of the mean spectra for the *Turquoise* and *Yellow* colours. 67
4.11 Metamere energy is correlated most strongly with the medium primary intensity for the *Pink* colour metameres. Energy is in arbitrary units. 68
4.12 Energy plots for the *Pink* colour sample show that there is a seven-fold difference between the maximum and the minimum energies required to generate this colour (red = power of the initial target spectrum). 68

4.13 Scatter matrix of metamere tristimulus values and energy for the *Turquoise* colour sample shows that the tristimulus values nor the spectral energies are uniformly distributed. 69
4.14 Scatter matrix of the spectral properties of the *Turquoise* metameres show that there are modes in this space where metameres preferentially occupy. 69
4.15 Scatter matrix of metamere tristimulus values and energy for the *Yellow* colour sample shows that by increasing the quantisation resolution, the tristimulus values can be made to be more uniformly distributed. 70
4.16 Scatter matrix of the spectral properties of the *Yellow* metameres even with a finer quantisation, some properties (e.g., *mIntensity* and *mMean*) show a non-uniform metamere distribution. 70
4.17 Scatter matrix of metamere tristimulus values and energy for the *Pink* colour sample shows that the tristimulus values nor the spectral energies are uniformly distributed. . . 71
4.18 Scatter matrix of the spectral properties of the *Pink* metameres are quite similar to that of the *Turquoise* sample indicating that the observed structure is biased by the coarseness of the quantisation. 71
4.19 Two Possible ways (blue and yellow paths) of achieving a colour c (red dot) in an 8-primary display. White dots are the chromaticities of the display's primary emissions 72
4.20 Common display colour gamut standards plotted on the CIE-1931 chromaticity diagram. 72
4.21 At narrow FWHMs a substantial amount of observer metameric breakdown can be observed (left), which can be reduced by increasing the FWHM (right). 76
4.22 In the extreme case, narrow wavelengths cause substantial deviation in the individual observers' tristimulus values, implying a high degree of observer metameric breakdown. 77
4.23 The metameric breakdown can be mitigated be carefully picking the location of the primary wavelengths. 77
4.24 In the extreme case, a substantial degree of metameric breakdown can be seen even up to FWHM=50. When primaries are carefully chosen, this deviation is minimal across the range of FWHMs tested. 78
4.25 When using primaries with narrow FWHMs, the individual observer deviation is extremely sensitive to the primary wavelength. The sensitivity decreases with increasing primary FWHM. 79

5.1 The original *Crayon* study used three bitmap images and two vector images as the baseline [21]. 86
5.2 A collage of the images used in the internal study showing the diversity in the images' colour, source, and style. 89
5.3 The Λ value required to achieve a given score depends on the distance metric and colour space. 91
5.4 The relative power saving is most strongly correlated with user score in the images transformed using the ℓ_2 distance. 92

5.5 The correlation of root mean squared error (RMSE) and structural similarity index measure (SSIM) metrics with Lambda (Λ) is heteroscedastic for images transformed using the ℓ_2 distance in *sRGB* colour space. 93
5.6 The correlation of RMSE and SSIM metrics with score for images transformed using the ℓ_2 distance in *sRGB* colour space is similar the variation of the image quality metrics (IQMs) with Lambda (Λ), but discretised. 93
5.7 The correlation of RMSE and SSIM for images transformed using the ℓ_2^2 distance in the *sRGB* and *CIEUVW* colour spaces show more homoscedastic variation. 94
5.8 Colour swatches for the best and worst images show that the 'best' transformed images show a bias towards darker colours. 95
5.9 Some of users' scores are inconsistent with the observed perceptual quality of images. . 95
5.10 The majority of the 'best' transformed images were transformed using the ℓ_2^2 distance and contain more darker colours. 96
5.11 All of the 'worst' transformed images were transformed using the ℓ_2 distance. 97
5.12 Variation of the single-pixel power over a range of Lambda (Λ) values with the *Huawei P30 Pro* model. The black horizontal line indicates 85% relative power. 98
5.13 Variation of the single-pixel power over a range of Lambda (Λ) values with the model used in the original *Crayon* work. The black horizontal line indicates 85% relative power. 98
5.14 *Huawei P30 Pro* power model. 99
5.15 Original *Crayon* power model. 99
5.16 The *SUREAL* method aims to factor out inconsistencies and biases imposed by the users and the images to achieve more stable estimates for the user score. 100
5.17 Simply training models using Lambda (Λ) to predict score led to predictions that were close numerically, but were unrealistic (in the *sRGB* space using ℓ_2^2 distance). 102
5.18 Normalising Lambda (Λ) enables the support vector machine (SVM) model to achieve better performance but still with a non-monotonic-increasing trend as required. 103
5.19 Collage of images used in the second contemporary study to highlight the diversity of images. 105
5.20 We used a *Binder MK56* thermal chamber (left) to measure display parameters for the *Huawei P30 Pro* devices (inside the chamber, right) at 25°C. 105
5.21 Only the results from the above subset of images are used in figs. 5.23 and 5.27 to make the graphs of the study results more readable. The images are referenced by the author's last name; (a) Hojo, (b) Reis, (c) Owens, (d) Rajendharkumar, (e) Mu, (f) Johnson. . . 105
5.22 The strength of the *Crayon* transform can be controlled by the parameter Lambda (Λ). Lower Λ achieves improved power savings at the expense of image quality. Images were transformed in the RGB colour space using the ℓ_2^2 distance. 106

5.23 MOS variation with Lambda (Λ) for different images follow a similar trend when using ℓ_2^2 distance (figs. 5.23a, 5.23c and 5.23e) in contrast to ℓ_2 distance (figs. 5.23b, 5.23d and 5.23f) in all the colour spaces (*sRGB* (figs. 5.23a and 5.23b), *CIELAB* (figs. 5.23c and 5.23d), and *CIEUVW* (figs. 5.23e and 5.23f)). 107
5.24 Image quality is preserved for a larger range of Lambda (Λ) when using the ℓ_2^2 distance (a) in contrast to ℓ_2 distance (b). 108
5.25 The relative display power variation with Lambda (Λ) for different images also follow a similar trend when using ℓ_2^2 distance (figs. 5.25a, 5.25c and 5.25e) in contrast to ℓ_2 distance (figs. 5.25b, 5.25d and 5.25f) in all the colour spaces (*sRGB* (figs. 5.25a and 5.25b), *CIELAB* (figs. 5.25c and 5.25d), and *CIEUVW* (figs. 5.25e and 5.25f)). . . 109
5.26 Correlation coefficients for different heuristics (ℓ_2^2 distance). Top to bottom, *sRGB*, *CIELAB*, *CIEUVW*. 111
5.27 The exponential model can fit the lower bound for Lambda (Λ) against mean opinion score (MOS) when using ℓ_2^2 distance for all three colour spaces; *sRGB* (fig. 5.27a), *CIELAB* (fig. 5.27b), and *CIEUVW* (fig. 5.27c). 114
5.28 Predicted lower bounds for the image in fig. 5.21b using the *leave-out* SVM model follow the trend of the data points better than the calculated fit for all three colour spaces; *sRGB* (fig. 5.28a), *CIELAB* (fig. 5.28b), and *CIEUVW* (fig. 5.28c). 114
5.29 Given a target image quality (MOS), the proposed method calculates the most optimal Lambda (Λ) for each image instead of globally choosing Λ which can be too high, causing increased image artefacting or be too conservative, leading to sub-optimal power saving. ('Image index' arbitrarily indexes the images used in the first study (fig. 5.2), without the images used to train the models (fig. 5.21).) 116
5.30 Images of chequerboard patterns along with vertical and horizontal black and white stripes were used as test patterns in the spatial dependency tests. 117
5.31 Display power decreases as finer chequerboard patterns are displayed. 118
5.32 Display power is mostly invariant to the fineness of the horizontal pattern. 118
5.33 Display power decreases as finer vertical patterns are displayed. 118
5.34 Close up images of the display show the black and white regions are not aligned. . . . 120
5.35 The pixel misalignment is worsened when finer patterns are displayed (fineness = 371). 120
5.36 The pattern where the pixel colour should change every other pixel, shows all white pixels. 120

6.1 The optimisation stage aims to find the optimal spectral shape as function of the primary wavelengths ($\overline{\lambda^*}$) and intensities ($\overline{\beta^*}$). 123
6.2 The CTM simulates a six-layer QD device. 123
6.3 The optimisation process used two initialisation schemes for the primaries: ② uniformly placed primaries in the range 330 nm to 700 nm (pale green shaded area) and ③ primaries placed uniformly in a narrow range centred around the LMS cone fundamental peaks (420 nm to 599 nm) (pale orange shaded area). 128

6.4 Plots show the results of the optimisation process with the *reciprocal CRI* cost, for an increasing number of primaries from 5 to 8. Note that the peak at 400 nm in the 5-primary case and 370 nm in the 8-primary case are actually two primaries overlaid (as evident by the increased FWHM). 131
6.5 Sequential optimisation using *sum of squares* cost and the *reciprocal CRI* cost does not reliably result in spectra with both chromaticity match and high colour rendering index (CRI). 132
6.6 The spectral shapes of the 4, 5, and 6 primary devices show that they match the intensity of the CIE-D65 spectrum as designed. 134
6.7 The variation of the spectral properties of the chosen devices as a function of the number of primaries. 134
6.8 The chromaticities of the calculated spectra closely match that of CIE-D65. 135
6.9 QD diameter model was obtained by fitting an exponential to measurements in the literature. 135
6.10 The relative radiances of the simulated primaries do not match the required radiance ratio if the emission area was the same for all the primaries. 137
6.11 Simulated optoelectronic properties of the 4-primary device. 138
6.12 Simulated optoelectronic properties of the 6-primary device. 139
6.13 Calculated sub-pixel areas for the two devices so that they achieve a colour match at 10 V. 139
6.14 A code snippet showing how a complete QD-LED can be defined in with the python interface. 140
6.15 Interface layers can be easily defined with its properties arranged in a hierarchical manner. 140
6.16 Material definitions can be re-used and are agnostic to the application. 140
6.17 The spectral properties of the simulated 4-primary and 6-primary devices match closely with their calculated spectral properties. Left to right, emission spectrum, chromaticity coordinates on the CIE-xy space, and the individual test colour sample (TCS) CRIs. . . 141

A.1 We used a *Binder MK56* thermal chamber (left) to measure display parameters for the Huawei P30 Pro devices (inside the chamber, right) at 25°C. 156
A.2 Raw current readings from the devices are very noisy and needs filtering. 157
A.3 The measurement variations are more apparent when zoomed in. 158
A.4 By correlating the display current with the display state, the variation of current as a function of pixel value for the three channels can be calculated. 158
A.5 Zero-filtering removes the baseline power readings and make the current variation more apparent. 159
A.6 'Neighbourhood averaging' removes the observed variation at a given display state and results in a clear current trend. 159

A.7 Pariwise re-normalisation uses the average baseline reading (blue) immediately before the active display state (red) to isolate the current draw of the display from the total system current draw. 160
A.8 The pairwise re-normalisation process isolates the display power as a function of the pixel RGB value. 160
A.9 The baseline current extracted by the 'neighbourhood averaging' process is highly correlated with the display state. This indicates that the rise and fall times associated with the display sate changes are causing the baseline current to be overestimated. . . . 161
A.10 Marking the timestamps for display state changes makes it clear that the calculated display current is biased by the rise and fall times of the system current. 161
A.11 Step-filtering leads to a more accurate estimate of the system current draw compared to the neighbourhood averaging and shows the clear trend even before the pairwise re-normalisation process. 162
A.12 Baseline current draw extracted by the step-filtering method is more stable and shows no correlation with the display state. 162
A.13 Step-filtered and pairwise re-normalised values show a clear quadratic relationship between the display current and the RGB value. 163
A.14 The results of simply zero-filtering the raw current measurements are inadequate for calculating the display power model. 163
A.15 Due to the effect of rise/fall times, the neighbourhood averaging process significantly underestimates the overall system current draw. 163
A.16 The error caused in the neighbourhood averaging propagates to the pairwise re-normalisation step where it is exacerbated by the overestimated baseline current reading leading to a calculation error $> 100\mu A$ (for the blue pixel at 255). 164
A.17 The display power model is calculated by fitting a triple of quadratics to the step-filtered and pairwise re-normalised data. 164
A.18 The current draw of the display is not significantly affected by the ambient temperature. 164

List of Tables

3.1 Top 25 countries by *player count*. 33
3.2 Top 25 countries by *average play time*. 33
3.3 Top 25 countries by *players per 100k people*. 34
3.4 Standard Deviation of the physiological parameters from Asano et al. [22]. 54
3.5 Individual player's optical properties learned by the optimiser. 55

4.1 Search Parameters and results for Brute Force search for metameres. 62

5.1 Normalised Lambda (Λ) values allow comparison between different display power models. 104
5.2 Different heuristics are correlated to varying degrees with the parameter k. 112
5.3 The model is able calculate the lower bound parameter k to a mean percentage mean squared error (MSE) of 7%, with 5-fold cross-validation. 113
5.4 When using 13 of the 14 original images, the SVM model is able to further reduce the error to 1.6% showing that the accuracy could be increased with more data. 113

6.1 Optimised primary wavelengths for 3, 4, 5, and 6 primary devices. 133
6.2 Chromatic properties of the calculated 3, 4, 5, and 6 primary devices. 133
6.3 Properties of the simulated 4 and 6 primary devices. 137

B.1 Performance of the models using Λ to predict the user adjusted scores (UAS). 165
B.2 Performance of the models using Λ to predict the mean opinion scores (MOS). 165
B.3 Performance of the models using normalised Λ to predict the mean opinion scores (MOS). 166
B.4 Performance of the models using normalised Λ, and mean and variance of HSV values to predict the mean opinion scores (MOS). 166
B.5 Performance of the models using normalised Λ, and mean and variance of CIE-Lab values to predict the mean opinion scores (MOS). 167

C.1 Summary of abbreviations used in the data tables 169
C.2 The properties of the calculated optimal spectra for the different cost functions and optimisation methods. 170

C.3 Primary wavelengths and the relative intensities of the primaries for the calculated optimal spectra using the different cost functions and optimisation methods. 171
C.4 Primary wavelengths and the relative intensities of the primaries for the calculated optimal spectra using the different cost functions and optimisation methods (continued). 172

List of Abbreviations

AMT Amazon Mechanical Turk.
API Application Programming Interface.
CCFL Cold Cathode Fluorescent Lamp.
CCT Correlated Colour Temperature.
CE Current Efficiency.
CIE International Commission on Illumination.
CMF Colour Matching Function.
CRI Colour Rendering Index.
CTM Charge-Transport Model.
DSIS Double-Stimulus Impairment Scale.
EBL Electron Blocking Layer.
EML Emissive Layer.
EQE External Quantum Efficiency.
ETL Electron Transfer Layer.
FWHM Full Width At Half Maximum.
GP Gaussian Process.
GPU Graphics Processing Unit.
HBL Hole Blocking Layer.
HDR High Dynamic Range.
HIL Hole Injection Layer.
HOMO Highest Occupied Molecular Orbital.
HTL Hole Transfer Layer.
HVS Human Vision System.
IQM Image Quality Metric.
LC Liquid Crystal.
LCD Liquid Crystal Display.
LED Light Emitting Diode.
LEF Luminous Efficiency Function.
LGN Lateral Geniculate Nucleus.
LUMO Lowest Unoccupied Molecular Orbital.
MOCVD Metal-Organic Chemical Vapour Deposition.
MOS Mean Opinion Score.
MPD Multi-Primary Display.
MSE Mean Squared Error.
OLED Organic Light Emitting Diode.
PPI Pixels per Inch.
QD Quantum Dot.
QD-LED Quantum Dot Light Emitting Diode.
RMSE Root Mean Squared Error.
SDR Standard Dynamic Range.
SPD Spectral Power Distribution.
SSIM Structural Similarity Index Measure.
SVM Support Vector Machine.
TCAD Technology-CAD.
TCS Test Colour Sample.
TFT Thin Film Transistor.
UAS User Adjusted Score.
UHD Ultra High Definition.
VTE Vacuum Thermal Evaporation.
WCG Wide Colour Gamut.
XGA Extended Graphics Array.

Hofstadter's Law: It always takes longer than you expect, even when you take into account Hofstadter's Law.

— Dr. Douglas Hofstadter.
Gödel, Escher, Bach: An Eternal Golden Braid [33, p. 152].

Chapter 1

Introduction

Following Moore's law, there has been a continuous push for the miniaturisation of solid-state fabrication technologies to create faster and more power-efficient compute devices [23]. This has allowed for faster portable computers and mobile phones, but has also led to the advent of increasingly mobile computing systems such as smart watches and other wearable devices. These leverage the increase in energy efficiency facilitated by the aforementioned fabrication process optimisations to allow their users to carry out meaningful computing tasks within the constraints imposed by the battery technology limits (e.g. activity tracking, audio streaming, etc.).

Most of these mobile systems also incorporate one or more displays. Unlike solid-state logic devices, displays rely on optoelectronic processes that do not miniaturise well, which poses an upper bound for the amount of energy saving that could be obtained through process optimisations [24–26]. Thus, increasingly, the overall system power is dominated by these displays and this will continue to worsen for the foreseeable future without radical technological advances.

There is an increasing shift towards using emissive pixel technologies like organic light emitting diodes (OLEDs) and quantum dot light emitting diodes (QD-LEDs), in these mobile devices, due to their increased performance over traditional liquid crystal displays (LCDs). As these displays do not include a separate backlight, they are thinner and lighter than LCDs, and exhibit better image contrast [27]. OLED displays can support up to three orders of magnitude higher refresh rates than LCDs [28], while QD-LEDs can produce colours from a much larger colour gamut [29].

The power demands placed by displays are further exacerbated by the consumer demand for larger, brighter, and more pixel-dense displays. The recent shift towards 4K ultra high definition (UHD) and high dynamic range (HDR) content require new displays to be compatible with these technologies to be competitive in the market [30, 31]. For context standard dynamic range (SDR) displays usually have a peak luminance of around 400 cd/m^2 [32]. In contrast, the HDR standard set out by *ITU Rec BT-2100* states that for displays to be HDR compliant, they need to be able to achieve a peak luminance of 1000 cd/m^2 [31].

Due to the lack of an *always-on* backlight, emissive pixel displays can also operate at near-zero power (when displaying black), and the power consumption of the display is a function of the content

being displayed. This presents an opportunity for optimising the display power by modulating the optical emission spectrum of the display.

In this thesis we explore how to leverage the limits of human colour perception and the opportunities presented by QD-LED based emissive pixel displays to devise optimisation schemes to improve their energy efficiency. We also evaluate how to use QD-LEDs to design high-performance lighting devices and present an optimisation based framework for quantum dot (QD) device design.

The thesis is arranged into seven chapters; the introduction (this chapter), a background chapter, four core chapters, and the conclusions.

Chapter 2 provides a summary of the core theories the rest of the work presented in this thesis is based on while also providing context for the discussions in the subsequent chapters. It starts with a primer on the mechanics of light and the human perception of it. This is supplemented by a discussion on how light is quantified. The core of this chapter discusses colourimetry, the study of how colour sensations can be quantified. It summarises the key definitions and practices used in colourimetry and while also giving an introduction to common colour spaces used to store and reproduce colour sensations. Finally, the chapter gives details on the operational principles of the main semiconductor technologies used to generate light in the context of both displays and general lighting devices. The chapter concludes with a summary of how the ideas presented in it relate to the chapters that follow.

Chapter 3 details how to use the information from a mobile video game about colour matching to calculate the spectral sensitivities of the players' eyes. The chapter summarises the information contained in the anonymised data logs collected by the game and present statistics about the colour acuity of players from around the world. We present a model for how to infer the spectral sensitivities based on the players' interactions with the game and evaluate whether they can be learnt through an optimisation process. The final part of the chapter presents a physiologically relevant vision model and show that it is possible to infer potential candidate values for a player's ocular properties using this model and the video game data.

Chapter 4 is concerned with metamerism; the effect where quantifiably different spectra are perceived to be the same colour. The chapter outlines how different spectra can take different amounts of energy to generate and through a brute-force search through the space of all three-primary colour spectra (with Gaussian-shaped primary emissions), we show that up to an order of magnitude power saving can be achieved by leveraging metamerism. The chapter also presents a mathematical framework for describing multi-primary displays (MPDs) and derives two indices for measuring the metameric capacity of an MPD. Finally, this chapter illustrates the potential issue of metameric-breakdown in QD based displays and highlights a negative result in trying to use MPDs to minimise metameric-breakdown.

Chapter 5 approaches the idea of using metamerism to achieve display power saving from a software algorithmic point of view. It presents an extension to prior work in the literature that uses colour approximation in images to minimise display energy. The chapter outlines the results of a user study we conducted to quantify the effect of this colour approximation on the perceived quality of images. We show that up to 50% power saving can be achieved with the participants reporting acceptable quality degradation. Based on these results, we present a normalisation scheme for the colour transform

parameters and a method for calculating optimal transform parameters based on the properties of the image being transformed. This chapter also presents a negative result of an investigation we carried out into evaluating whether the display power is affected by the spatial arrangement of pixels and the use of sub-pixel rendering to minimise display power consumption.

Chapter 6, in contrast to the others, deals with the use of QDs in smart lighting; in particular, lights with high colour rendering index (CRI). In this chapter, we present a numerical-optimisation based method for designing QD based devices. We show how this method can be used with an MPD-like architecture to design QD based white lights with a CRI exceeding 97. The chapter also presents the results of device simulations we carried out using a *charge-transport model* based simulator.

Finally, chapter 7 provides a summary of the key findings presented in this thesis, and recommendations for future work.

'Begin at the beginning', the King said gravely, 'and go on till you come to the end: then stop'.

— Lewis Carroll.
Alice in Wonderland [38, p. 182].

Chapter 2

Human Colour Vision, Colourimetry, and Display Technologies

2.1 Introduction

In the late 1600s, Isaac Newton discovered that *white* light is composed of seven primary colours; red, orange, yellow, green, blue, indigo and violet [34, p. 8]. Newton also observed a correlation between the wavelength of the monochromatic light and its perceived hue, and incorrectly concluded that there is a necessary correlation for all hues [35, p. 13]. However, we now know that this quantisation into seven primary colours is somewhat arbitrary and real white light is a continuous spectrum of different wavelengths of electromagnetic radiation. Although a white light spectrum can be decomposed into a sum of monochromatic beams, it is fallacious to say that white light in general is *composed* of these beams, since the decomposition is not necessarily unique [35, p. 14]. This idea was pioneered by Thomas Young with his trichromatic colour theory and later expanded on by James Clerk Maxwell through his colour matching experiments [35, p. 20][36]. These ideas form the foundation of colourimetry to this day.

This chapter aims to provide a primer on the fundamentals of light, human vision, and colourimetry. As a key aspect of the project is on light and the human perception of it, it is important to understand how humans perceive light. This chapter will also summarise the key aspects of modern semiconductor-based light generation methods and provide context to the work in the chapters that follow. Note that, the study of human vision, colourimetry, and devices that generate colour sensations encompass a large body of disjointed work that cannot be summarised in enough detail in a single chapter. As such, it should be understood that the work presented in this chapter is not exhaustive and is not meant to serve as a survey of the fields, but rather present, in simple terms, the concepts relevant to the discussions contained in the later chapters of this thesis.

2.2 What is Light?

A beam of light can be defined as an incoherent collection of electromagnetic waves [35, p. 55]. When the wavelengths of the waves constituting the beam are in the range of wavelengths the human eye is sensitive to (approximately 350 nm to 800 nm), it is referred to as *visible light* [35, p. 55]. The collection of individual wavelengths is called the *spectrum* of the light, and is described according the relative intensities of the different wavelengths (see fig. 2.1 and section 2.6). From the perspective of quantum mechanics, this beam of light can be described as being composed of discrete *packets of light*, termed *photons*. The energy of a single photon is governed by the Einstein relation $E = h\nu$, where h is the Plank's constant and ν is the frequency of the light [37, p. 59]. This key relationship between the energy of light and its frequency (and consequently, its wavelength) would be instrumental in discussing how light is generated in optoelectronic devices, especially those based on quantum dots (QDs).

In the writing that follows, the word 'light' would be used when referring to an incoherent mixture of wavelengths in the visible range. We would use the phrase 'monochromatic light' when referring to light with a smaller range of contiguous wavelengths (e.g., '500 nm monochromatic light' implies that the most of the wavelengths comprising this light are around 500nm). Finally, when the quantised nature of light needs to be highlighted, 'photon' would be used. In rare instances, we will refer to a light source as 'a light', but the meaning would be clear from the context (primarily used in chapter 6).

2.3 What is Colour?

Colour can be defined in many ways depending on the context it is used. In an everyday context, the meaning of colour is so fundamental to ones' day to day experience that it is hard to describe it to someone who does not have any first-hand experience of it. However, to study colour, it needs to be defined at a level of abstraction appropriate for the field of study. In our discussions, we would use three definitions of 'colour'. The first is from a psycho-physical perspective to have a similar meaning to everyday use of the word *colour*. A *colour* is the appearance of a particular beam of light as sensed by the (human) eye and interpreted by the (human) visual processing system in the brain. However, as this is definition is somewhat abstract, without a more formal definition of *colour* it is difficult to discuss it more accurately. As such, we will also use a second definition, in terms of *coordinates in colour spaces* defined by the International Commission on Illumination (CIE) to quantify what a *colour* is. A *colour* in this sense is defined as a coordinate triple in a chosen CIE *colour space*. This will be discussed in more detail in section 2.7. The third definition of *colour* is from the perspective of (display) *colour models*, which are standardised representations used to record and reproduce *colour sensations*. Similar to the coordinates in CIE colour spaces, a *colour* as defined through these models refer to a coordinate triple with respect to a specified *colour models*. Common *colour models* include *sRGB* (commonly used to represent colours on the internet), and *CMYK* (used in print media). Refer to section 2.8 for more details.

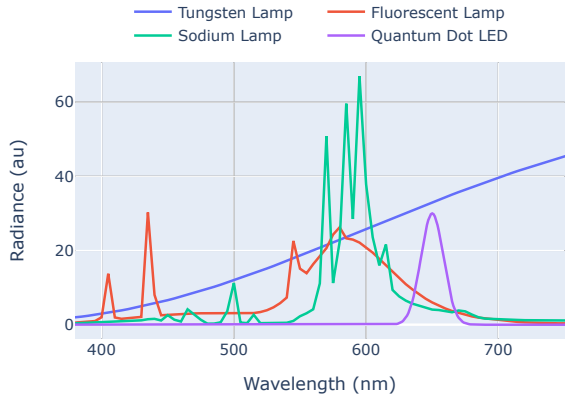


Fig. 2.1 Light spectra can have arbitrary shapes depending on the physics underlying the light generating processes (data from [3, 4]).

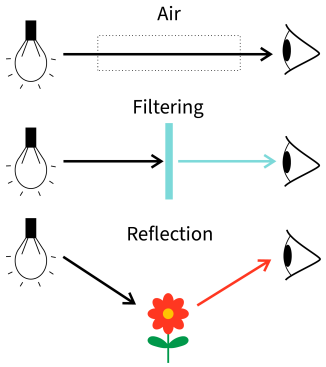


Fig. 2.2 Light from the source travels through a medium to the observer. The medium and objects in the light path can absorb and reflect different segments of the spectrum in different ways, affecting the light spectrum arriving at the observer.

2.4 Transmission of Light

There are three main components in the process of light perception; the light source, the medium of transmission, and the observer system (see fig. 2.2). The light source determines the properties of the initial light spectrum; frequency components, spectral radiance, emission angle, etc. The shape of the light spectrum is determined by the physics governing the light-generating processes in the source (e.g., black body radiation in the sun, radiative recombination in QDs, etc. (see fig. 2.1)). The medium of transmission determines how the source spectrum gets attenuated between the source and the eye. Typically, the transmission medium is air, but in general represents all the processes the spectrum is subjected to in its path from the source to the observer (e.g., fog in the air, polarising lenses on sunglasses, surfaces that the light reflects off of, etc.). The degree of attenuation is dependent on the path the light beam takes in getting from the source to the eye, and in general, can also be dependent on properties of the light itself (e.g., colour filters attenuating only some frequencies of light).

In our investigations, especially in chapter 6, we are particularly concerned about the effects of intermediate processes on the spectral components of the light (i.e., how are the radiances of the different spectral components changed by the transmission processes?). This is of particular importance when light reflects off surfaces. The effect of a reflective surface on an incident light spectrum is generally to attenuate different spectral components in the light to different degrees based on the chemical/physical structure of the surface.

A distinction is generally made between light that is reflected off surfaces and light that is not. When light is reflected off surfaces, the perceived colour of the spectrum is considered to be the colour of the surface. Otherwise, the perceived colour is taken as the colour of the light source. For example, consider a ‘green’ grassy field under (midday) daylight. The source in this situation is the sun, and the

spectrum is generally perceived to be 'white' and is considered to be the colour of the sunlight. The grass is illuminated by the same spectrum. However, the grass looks 'green' because it absorbs the spectral components in the 'white' sunlight in such a way that the human vision system (HVS) perceives it to be 'green'. Although, the spectrum has the same source, in this case one would refer to spectrum reflecting off the grass as if the grass is emitting that spectrum, when one refers to that spectrum as 'green' and refers to the grass as being 'green'. The importance of this distinction would be apparent in chapter 6. This is further complicated by processes like thin-film interference and diffraction which results in iridescence where the perceived colour of an object changes based on the angle that you look at it.

Finally, the light spectrum, after being attenuated by the transmission medium, is incident on the observer system. Where the observer system is the HVS, the light is incident on the eye, where it is transduced into a collection of electrical signals, which is then interpreted by the visual cortex in the brain. In general, the *observer* can be any system capable of responding to a light spectrum (e.g., a camera, spectrometer, etc.).

2.5 Biology of Vision

From the most simplistic point of view, the human vision system consists of three components: the eye, the optic nerve, and the brain. Light from the environment enters the eye through the pupil and is incident on the retina at the back of the eye. Photosensitive cells in the retina captures the energy in the light and converts it into electrical impulses. The optic nerve carries these impulses to the brain and the brain interprets them to produce visual sensations.

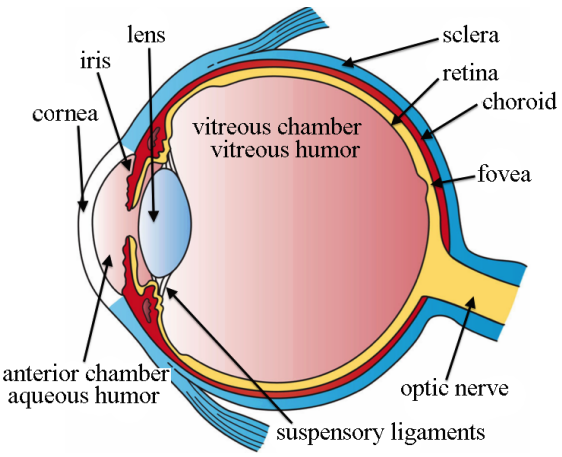


Fig. 2.3 The structure of the human eye captures light from the environment focusses it onto the retina (from [5]).

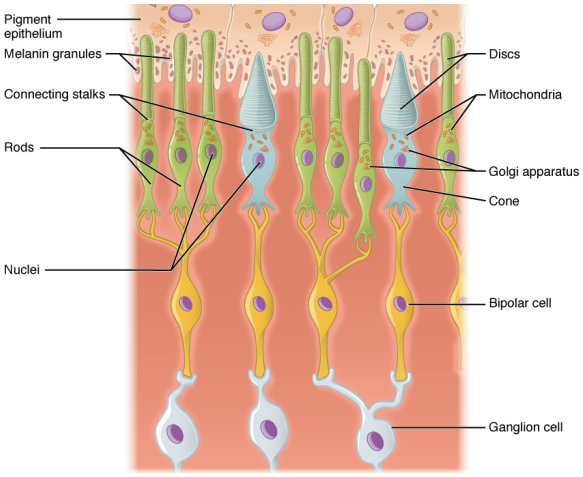


Fig. 2.4 Photoreceptive cells and Ganglion cells in the retina converts the light into electrical signals and sends it to the brain (from [6]).

2.5.1 The Eye

At the front of the eye is the lens. The lens is responsible for focusing light from the environment onto the retina where there are photosensitive cells (see fig. 2.3). As humans age, the lens' ability to transmit short-wavelength light decreases, which leads to a yellowing of the lens (i.e., an increase in the lens pigment) [39, p. 3]. The volume at the centre of the eye, referred to as the *vitreous humour*, is full of a transmissive liquid that has a positive pressure relative to air pressure and is responsible for maintaining the shape of the eye.

There are roughly 127 million photosensitive cells in the retina; referred to as rods and cones due to their shapes [39, p. 11]. These are responsible for converting the energy in the incident light to electrical signals. Cones are specialised in detecting details and colour in well-lit environments ($> 100\,cd/m^2$), and are concentrated around the centre of the retina (referred to as the fovea), where the image of the object the person is focusing on, would form. Rods are found mostly in the peripheral regions of the retina and are more numerous than cones; roughly 120 million rods to 7 million cones [39, p. 8]. Rods lack the ability to discern colours but are more sensitive to low light [34, p. 20-21]. The rods and cones have different spectral sensitivities (see fig. 2.5). In low light ($< 1\,cd/m^2$), where the response of rod cells dominate, the vision is said to be *scotopic*. In normal lighting conditions, the eye has *photopic* vision, where the optical response of the cone cells dominate. Vision at intermediate luminances that incorporate the responses from both rods and cones are referred to as *mesopic* vision [39, p. 8]. The CIE standard luminous efficiency functions (LEFs) for *scotopic* and *photopic* vision (also referred to as $V'(\lambda)$ and $V(\lambda)$ respectively) are plotted in fig. 2.5 [34, p. 22].

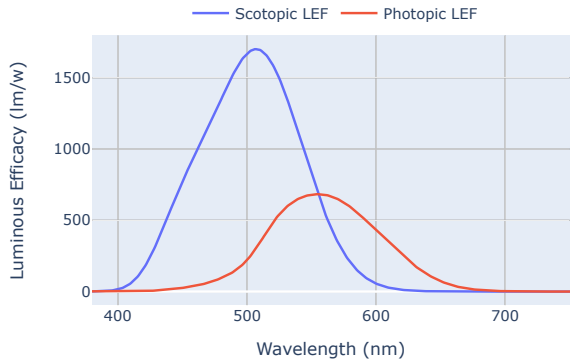


Fig. 2.5 Human *scotopic* vision is nearly three times more sensitive to light than *photopic* vision (after [7]).

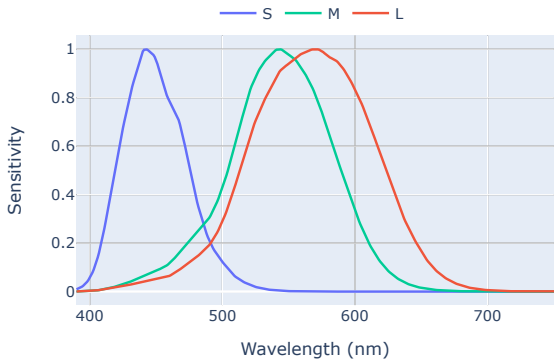


Fig. 2.6 The three types of cone cells have overlapping bandpass sensitivities to different spectral ranges (after [8]).

Looking at cone cells in more detail, there are three types, referred to as L, M, and S cones respectively. Each type has a band-pass sensitivity to a range of overlapping wavelengths in the, 'long', 'medium', and 'short' wavelength ranges. Figure 2.6 shows a plot of the spectral sensitivities of each cone cell type, where the sensitivity corresponds to the probability that a photon of a given wavelength is absorbed by the cone [36]. The relative strength of the electrical signal generated by the cone cell

would thus depend on the intensity of the incident light as well as its constituent wavelengths. Note that the aforementioned $V(\lambda)$ function represents the aggregate response of all 3 types of cone cells [39, p. 9]. The fovea has the highest density of cone cells. There are no rods in the fovea, and in the region $12^\circ - 15^\circ$ from the fovea has no photoreceptors at all, as it is where the optic nerve is formed (commonly referred to as the 'blind spot'). To prevent high-intensity short-wavelength light from damaging the fovea, there is a yellow filter in front of it, called the *macula*. The variation of lens pigment and macula pigment are considered to be the main causes colour vision variability between observers with normal colour vision (i.e., no colour blindness or anomalous colour vision (e.g., protanomaly/deuteranomaly, etc)) [39, p. 6]. The relative proportions of L, M, and S cones in the eye are approximately in the ratio 40:20:1 and outside the fovea the density of cone cells is much lower [39, p. 11].

The photoreceptors are connected to the optic nerve through a hierarchical network of cells (see fig. 2.4). There are four types of cells in this network; bipolar, ganglion, horizontal, and amacrine. The ganglion cells are at the top of the hierarchy, connecting the optic nerve to the network of cells in the retina. The bipolar cells aggregate responses from multiple photoreceptors and pass them to the ganglion cells. The horizontal cells on the other hand connects photoreceptors and bipolar cells laterally to other photoreceptors and bipolar cells. The amacrine cells create similar connections among bipolar and ganglion cells.

This system is highly complex and the specifics of the processing that occurs in each cell type is not yet understood completely [39, p. 14]. However, the overall effect of this network of cells is to convert the responses from nearly 130 million photoreceptors such that the information contained in their responses can be represented by the responses from approximately 1 million ganglion cells without loss of meaningful information [39, p. 6].

2.5.2 Visual Processing in the Brain

The processing of the visual signals in the brain is itself quite a lot more complex to the extent that the exact mechanisms involved in devising the high-level semantic representations from the visual signals are still debated [40, p. 54].

Once the optical signals arriving at the retina have been converted to chemical and electrical signals by the photoreceptors, they are converted to a low-dimensional representation at the ganglion cells and are projected directly into the lateral geniculate nucleus (LGN) in the brain. The ganglion cells themselves act in a similar way to an edge detection kernel in that their receptive field is sensitive to the *change* in the photoreceptor signals [39, p. 14]. There is a one-to-one mapping between the ganglion cells and the cells in the LGN. The LGN modulates the response from the ganglion cells based on feedback from higher levels in the visual cortex and projects this information into the *visual area one* (V1) in the occipital lobe [39, p. 14].

In V1, the output from multiple LGN cells are compared and combined to create higher-level representations of the visual stimulus [39, p. 15]. It has been shown that the receptive fields in V1 are sensitive to oriented edges, differences between the input from the two eyes, and spatial/temporal

frequency variations in the stimulus [41]. The responses from V1 are fed into a matrix of inter-connected higher-order visual processing areas that are responsible for responding to more complex spatial and temporal properties of the incident light [40, p. 54]. The net effect of these systems is to produce a visual sensation that is a comprehensible representation of the external world.

2.6 Radiometry and Photometry

Before analysing the *effects* of light on human perception, we need to quantify the *properties* of light. Radiometry and photometry are the fields of study focused on quantifying measurable properties of electromagnetic radiation [35, p. 59]. Primarily this involves measuring the distribution of spectral power in the space through which the radiation propagates. In the context of the light transmission model presented in section 2.4, radiometry uses 'objective' measurement equipment as the observer and aims to measure all the radiation arriving at the observer and not just the radiation the HVS is sensitive to. In contrast, photometry measures the radiation in the context of the HVS, weighting the radiation arriving at the observer by the (photopic) LEF of the human eye ($V(\lambda)$) to simulate the perceived 'brightness' of the radiation [42, p. 24.7].

2.6.1 Key Radiometric Quantities

The following are the key radiometric quantities relevant to the investigations that follow [42, p. 24.8].

- **Radiant Energy** Q_e (J) : The total electromagnetic energy emitted by the source.
- **Radiant Flux** Φ_e (W) : The energy in the electromagnetic radiation being emitted by the source, per second.
$$\Phi_e = \frac{dQ_e}{dt}$$
- **Irradiance** $E_{e,\Omega}$ (W/m^2) : The radiant flux incident on a unit area (dA_d). Radiant exitance is a related quantity with the same units that measures the radiant flux *leaving* a unit area.
$$E_{e,\Omega} = \frac{d\Phi_e}{cos\phi_d dA_d}$$
- **Radiant Intensity** $I_{e,\Omega}$ (W/Sr) : The amount of radiant flux emitted or received per unit solid angle.
$$I_{e,\Omega} = \frac{d\Phi_e}{d\Omega}$$
- **Radiance** $L_{e,\Omega}$ $(W/Sr \cdot m^2)$: The radiant flux emitted or received per unit solid angle per unit area.
$$L_{e,\Omega} = \frac{d\Phi_e}{cos\theta_s dA_s d\Omega}$$

These quantities represent the total value across the whole wavelength range. When these measurements are given with respect to individual wavelength bands, they are given as 'spectral' quantities [43, p. 2].

For example, *radiance* ($L_{e,\Omega}$) becomes *spectral radiance* ($L_{e,\Omega,\lambda}$) with units of $W/Sr \cdot m^2 \cdot nm$ with the following relationship [42, p. 24.11]:

$$L_{e,\Omega} = \int_0^\infty L_{e,\Omega,\lambda} d\lambda$$

Note that, spectral irradiance ($E_{e,\Omega,\lambda}$) is referred to as the spectral power distribution (SPD) of a light source, but the phrase 'spectral power distribution' is also sometimes (erroneously) used to refer to any of the above quantities when they are expressed as a function of wavelength.

2.6.2 Key Photometric Quantities

Generally, the photometric quantities are obtained by weighting the radiometric quantities by the CIE standard observer photopic LEF, $V(\lambda)$ [42, p. 24.11]. The factor K_m, in the following equations, is the *luminous efficacy* for $V(\lambda)$ and sets the scale between radiometric quantities and photometric quantities. K_m is defined at the peak of $V(\lambda)$, and for photopic vision, $K_m \approx 683\ lm/W$ [42, p. 24.12].

- **Luminous Energy** Q_v ($cd \cdot Sr \cdot s$) : The equivalent photometric quantity to *Radiant Energy*.
- **Luminous Flux** Φ_v ($cd \cdot Sr$) : The equivalent photometric quantity to *Radiant Flux*. Sometimes given in units of lumens (lm) which is equivalent to $cd \cdot Sr$.

  $$\Phi_v = K_m \int \Phi_{e,\lambda} V(\lambda) d\lambda$$

  where $\Phi_{e,\lambda}$ is the radiant flux as defined in section 2.6.1.
- **Illuminance** $E_{v,\Omega}$ ($cd \cdot Sr/m^2$) : The equivalent photometric quantity to *Irradiance*. Sometimes given in units of *lux* which is equivalent to lm/m^2.

  $$E_{v,\Omega} = \frac{d\Phi_v}{cos\phi_d dA_d} = \frac{d\left[K_m \int \Phi_{e,\lambda} V(\lambda) d\lambda\right]}{cos\phi_d dA_d}$$

- **Luminous Intensity** $I_{v,\Omega}$ (cd) : The equivalent photometric quantity to *Radiant Intensity* and are given in units of candela (cd).

  $$I_{v,\Omega} = \frac{d\Phi_v}{d\Omega} = \frac{d\left[K_m \int \Phi_{e,\lambda} V(\lambda) d\lambda\right]}{d\Omega}$$

- **Luminance** $L_{v,\Omega}$ (cd/m^2) : The equivalent photometric quantity to *Radiance*. Sometimes given in non-SI units of 'nits', where 1 nit = 1 cd/m^2 [7].

  $$L_{v,\Omega} = \frac{d\Phi_v}{cos\theta_s dA_s d\Omega} = \frac{d\left[K_m \int \Phi_{e,\lambda} V(\lambda) d\lambda\right]}{cos\theta_s dA_s d\Omega}$$

Similar to the radiometric quantities, the photometric quantities can also be given as a function of wavelength (e.g., Spectral Luminance ($L_{v,\Omega,\lambda}$) in units of $cd/m^2 \cdot nm$).

2.7 Colourimetry and Colour Vision Models

Colourimetry is the study of numerically specifying the perceptual properties of spectral stimuli [43, p. 117]. While photometry is concerned with the perceived brightness of a light spectrum, colourimetry is concerned with its perceptual colour. The standards and practices used in modern-day colourimetry are regulated by the International Commission on Illumination (abbreviated to CIE after its French title).

2.7.1 Colour Matching Experiments

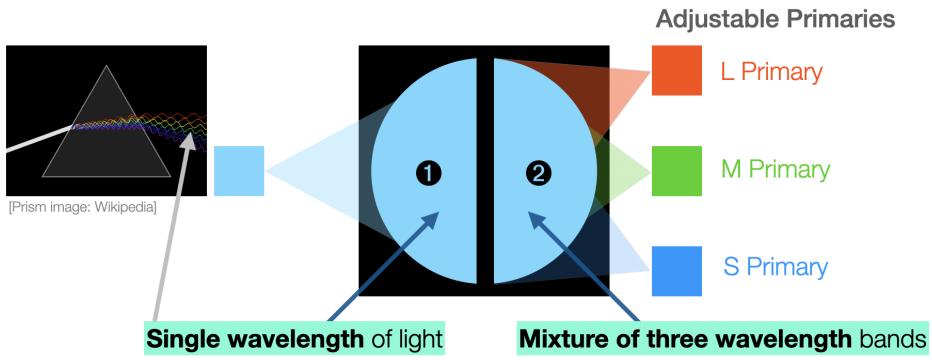


Fig. 2.7 Traditional colour matching experiments involve changing the intensities of 'long' (L), 'medium' (M), and 'short' (S) wavelength primaries till the mixture (❷) is perceptually identical to the monochromatic light (❶).

At the heart of colourimetry are colour matching experiments which themselves are based on the *trichromatic generalisation*, which states that many colour stimuli can be matched completely by an additive mixture of three fixed primary stimuli, provided that any one of the primaries cannot be obtained by an additive mixture of the other two [43, p. 117]. This was first demonstrated in 19th century, independently by Maxwell and Helmholtz, and forms the basis for most modern colour displays, where three fixed primary stimuli (red, green, and blue) are used to generate all the colour sensations a display can produce [36].

In colour matching experiments, the participants are presented with a split field of view, where half of the visual field is illuminated by a monochromatic light source, while the other half is illuminated by a mixture of the different primaries (see fig. 2.7). Generally, the three primaries are chosen to have a narrow band emission and relegated to 'long', 'medium', and 'short' wavelength ranges to maximise the coverage of the visible spectrum [36]. During the experiment, the participants are requested to adjust the mixture of the three primaries till the mixture (❷ in fig. 2.7) looks identical to the monochromatic light (❶ in fig. 2.7). The relative intensities of the three primaries that give a colour match are recorded, the wavelength of the monochromatic light is changed, and the process is repeated for all the wavelengths in the visible range. The resulting set of curves, as a function of the wavelength of the monochromatic light, is referred to as the colour matching functions (CMFs) (see fig. 2.8). Chapter 3 explores the statistics of

a global digital colour matching dataset and evaluate how to extract the individual observer CMFs from it.

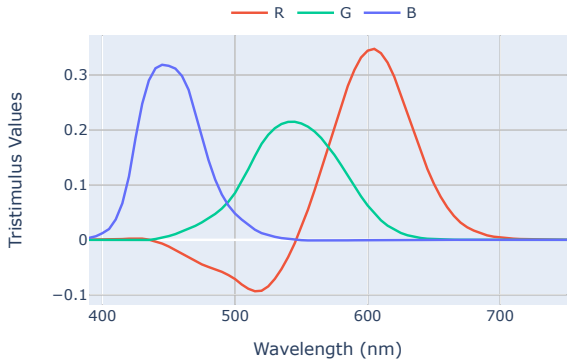


Fig. 2.8 CIE-1931 2° RGB CMFs were obtained by averaging the results of colour matching experiments with multiple participants.

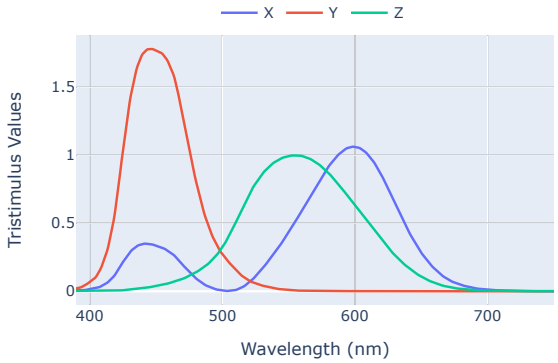


Fig. 2.9 CIE-1931 2° XYZ CMFs were calculated by transforming the RGB CMFs to get rid of the negative values.

2.7.2 CIE-1931 2° Standard Observer and Associated Colour Spaces

Pioneering work in the area of formalising vision through colour matching was done by Wright and Guild around the 1930s [44, 45]. Based on their results, a standard was devised by the CIE, referred to as the CIE-1931 RGB 2° standard observer; called as such due to the 2° visual field [46]. Figure 2.8 shows the CIE-1931 RGB 2° standard observer CMFs. At a given wavelength, the value of the CMFs represent the relative luminances of the different primaries required to match a monochromatic stimulus of that wavelength and are referred to its *tristimulus values*. The primaries used in the CIE-1931 RGB 2° std. observer were centred at 700 nm, 546.1 nm, and 435.8 nm for the 'red', 'green', and 'blue' primaries respectively. The intensities of the primaries were normalised such that the spectrum at colour match also matched in energy [10, p. 29]. Thus, for a wavelength with tristimulus values, R, G, and B, the luminance, L, is given by $L = 1.0R + 4.5907G + 0.0601B$.

Note that, a negative tristimulus value indicates that, to achieve a colour match at that wavelength, the given proportion of the primary with the negative tristimulus value had to be mixed with the monochromatic light (i.e., ❶ in fig. 2.7). To circumvent these negative tristimulus values in the CIE-1931 RGB colour space, the CIE devised a new colour space called the *CIE-1931 XYZ colour space* composed of three imaginary primaries, $[X, Y, Z]$, that could be defined as a linear transformation of the *CIE-1931 RGB colour space* (see fig. 2.9). The corresponding CMFs are usually referred to as the *CIE-1931* 2° *std. observer CMFs*.

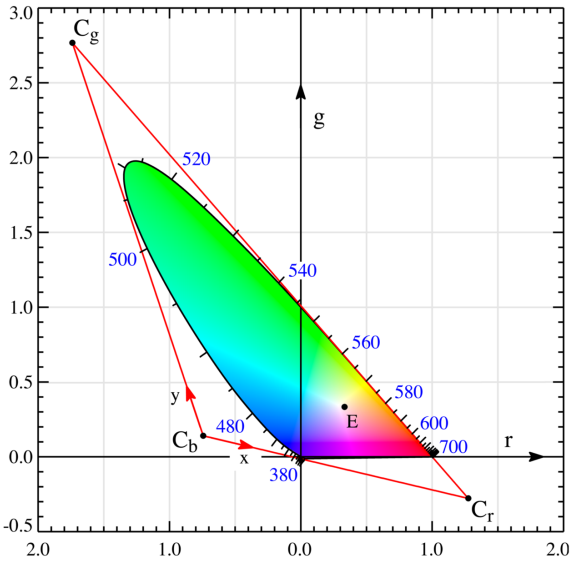


Fig. 2.10 Projection of the CIE-1931 RGB colour space onto the r-g plane shows that some colours require a negative proportion of the red primary to achieve a colour match (from [9]).

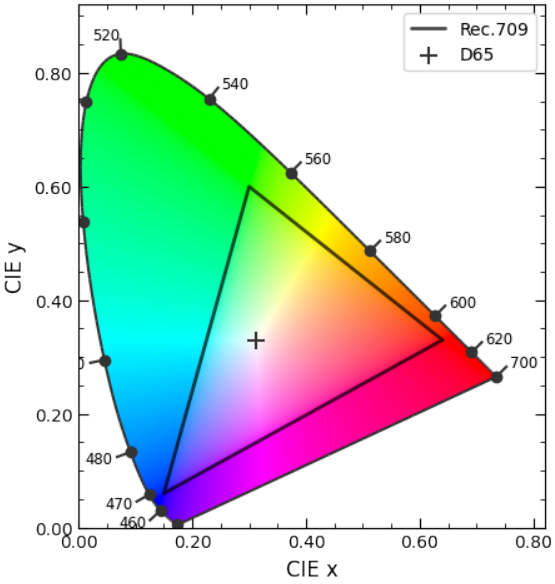


Fig. 2.11 CIE-1931 xy colour space encapsulates all the chromatic sensations the human eye can perceive, in a strictly positive coordinate space.

As the CIE-1931 XYZ CMFs are spectral quantities, they can be used to calculate the $[X,Y,Z]$ tristimulus values for a given spectrum using the following equations [10, p. 32]:

$$X = K_m \int_\lambda L_{e,\Omega,\lambda}(\lambda)\,\bar{x}(\lambda)\,d\lambda \tag{2.1}$$

$$Y = K_m \int_\lambda L_{e,\Omega,\lambda}(\lambda)\,\bar{y}(\lambda)\,d\lambda \tag{2.2}$$

$$Z = K_m \int_\lambda L_{e,\Omega,\lambda}(\lambda)\,\bar{z}(\lambda)\,d\lambda \tag{2.3}$$

where $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are the CMFs and K_m is the peak luminous efficacy as given in section 2.6.2. In general, any radiometric quantity can be used in place of $L_{e,\Omega,\lambda}(\lambda)$ [10, p. 32]. However, as $\bar{y}(\lambda)$ is equivalent to the photopic LEF, when using $L_{e,\Omega,\lambda}(\lambda)$, the tristimulus value Y is identical to the luminance of the spectrum [10, p. 32]. Based on the definition of the CMFs above, two spectra with identical tristimulus values will be indistinguishable when viewed under the same external conditions (e.g., light adaptation in the eye) [10, p. 33].

2.7.3 CIE-1931 xyY Chromaticity Diagram

Usually, it is important to discern the chromaticity (the 'colour') of a given spectrum. The CIE-1931 xyY chromaticity diagram was devised for this. The diagram is derived by transforming the CIE-1931 XYZ coordinates using the following normalisation:

$$x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z} \qquad z = \frac{Z}{X+Y+Z} \tag{2.4}$$

As $x+y+z=1$ the coordinates in this space are usually given as the triple $[x,y,Y]$. The CIE-1931 xy chromaticity diagram is a projection of this colour space, plotted on a cartesian plane [10, p. 33]. Figure 2.11 shows a plot of this colour space. The perimeter of the chromaticity diagram represents the chromaticities of monochromatic light, the wavelengths of which are indicated on the perimeter. The colours within this boundary represent all the (chromatic) colour sensations the human eye can perceive. Note that this space is a 3D space with luminance axis, Y, extending out the page.

Given two spectra with chromaticity coordinates of A and B, the line joining A and B represent all the colours that can be generated by an additive mix of those two spectra. Similarly, for three primaries A, B, and C, the area bounded by the lines joining their chromaticity coordinates represent all the colours that can be generated with additive mixtures of the three primaries. This area is called the *colour gamut* of the primaries.

2.7.4 CIE-1964 XYZ Colour Space

In 1964, the CIE 2° std. observer was supplemented by the 10° std. observer, based on work by Styles and Burch [47], and Speranskaya [48]. The key difference between the 10° std. observer compared to the 2° std. observer is that the new model represents the visual acuity for stimuli which lands within an approximately 10° angle on the retina. A 10° visual field corresponds to an area with a diameter of 90 mm at a distance of 0.5 m [10, p. 35]. An RGB to XYZ transformation similar to the CIE-1931 2° std. observer was also derived, as well as a chromaticity diagram similar to CIE-1931 xyY [10, p. 35]. Process for calculating the XYZ and xyY coordinates is identical to before, but with the 10° std. observer CMFs and a different transformation matrix for converting RGB CMFs to XYZ CMFs.

2.7.5 CIE-2006 Physiologically-Relevant CMFs

The CMFs presented thus far have been based on colour matching experiments. These serve as a proxy for directly measuring the spectral sensitivities of the cones in the human eye. However, it is these cone spectral sensitivities that ultimately define the fundamental processes involved in how the eye reacts to spectral energy. Around the year 2000, Stockman et al. presented results of measurements of individual cone spectral sensitivities through a mixture of in-person studies using colour vision deficient participants, and the results from past colour matching experiments and in-vitro retinal sample measurements reported in the literature [49, 50]. In this work, they were able to estimate the individual cone spectral sensitivities of their participants using this data. These spectral sensitivities are generally referred to as the *Stockman and Sharpe cone fundamentals*. Figure 2.6 shows these results graphically. Based on these results, the CIE presented a standard methodology for calculating the CMFs from physiological measurements [51]. The details of this process will be detailed in section 3.6, where it will be more relevant in context.

2.7.6 Uniform Colour Spaces: CIE-UCS, CIELAB, and CIEUVW

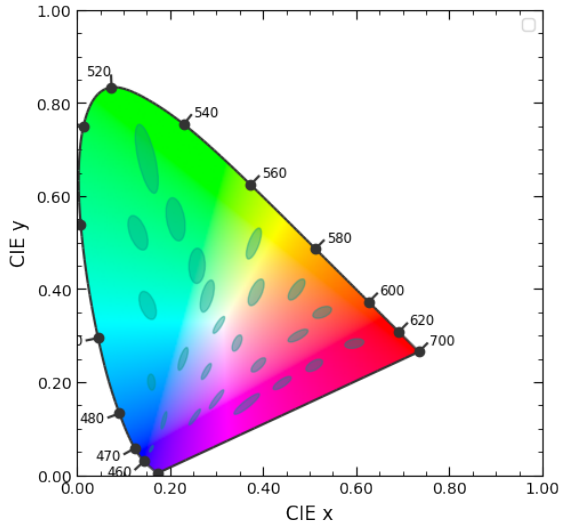


Fig. 2.12 MacAdam Ellipses (shown as 10x *just noticeable difference*) correspond to the set of colours around a colour where observers cannot discern the difference [10, p. 81].

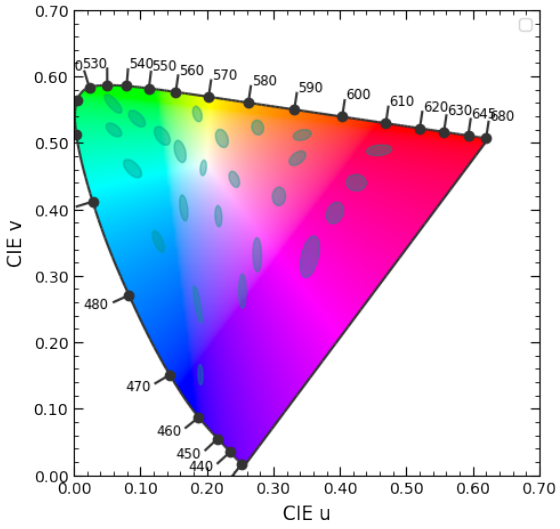


Fig. 2.13 In the CIE-1976 UCS the *MacAdam ellipses* are *more* uniform in eccentricity [10, p. 82].

CIE-1976 UCS

The CIE-xy chromaticity diagrams and the associated colour spaces are good for describing identity in colour stimuli, but fails to express how similar two colour stimuli are. In 1942, David MacAdam showed that the regions of *just noticeable colour difference* in the CIE-xy chromaticity diagram vary depending on the chromaticity of the stimulus (see fig. 2.12) [52]. Thus, simple distance metrics like Euclidean distance between the x and y coordinates cannot be used to represent the perceived colour difference between two stimuli. To mitigate this, in 1976, the CIE presented the *CIE Uniform Chromaticity Scale diagram* (CIE-1976 UCS or simply CIE-UCS) [10, p. 59]. The coordinates in this space were specified as $[u', v', w']$ and are calculated from the CIE-XYZ coordinates using eqs. 2.5, or from the CIE-xyY coordinates using eqs. 2.6.

$$u' = \frac{4X}{X+15Y+3Z} \qquad v' = \frac{9Y}{X+15Y+3Z} \tag{2.5}$$

$$u' = \frac{4x}{-2x+12y+3} \qquad v' = \frac{9y}{-2x+12y+3} \tag{2.6}$$

The XYZ coordinates can be calculated using either the CIE-1931 2° std. observer or the CIE-1964 10° std. observer, based on the visual angle. Similar to the CIE-xyz coordinates, $u' + v' + w' = 1$. Figure 2.13 shows a plot of the CIE-1976 UCS with MacAdam ellipses overlaid. Note that, although

the eccentricities of the MacAdam ellipses are smaller in CIE-UCS, the correlation between coordinate differences and colour differences is still low due to the variation in the ellipses' orientation. Related to CIE-UCS is the CIEUVW colour space with coordinates $[U^*, V^*, W^*]$ defined with the following equations,

$$U^* = 13W^*(u' - u'_n) \qquad V^* = 13W^*(v' - v'_n) \qquad W^* = 25Y^{1/3} - 17 \tag{2.7}$$

where $[u', v']$ and $[u'_n, v'_n]$ are CIE-1976 UCS coordinates of the test stimulus and the reference white colour stimulus respectively, and Y is the luminance component of the CIE-XYZ tristimulus values of the test stimulus.

CIE-1976 $L^*a^*b^*$ Colour Space (CIELAB)

Alongside the CIE-UCS recommendation, the CIE recommended two other colour spaces in the proceedings of the 1975 CIE technical committee sessions; CIELAB and CIELUV [10, p. 61]. As with the CIE-1976 UCS coordinate calculation, they use the CIE-XYZ tristimulus values as the basis for calculate their coordinates; namely $[L^*, a^*, b^*]$ and $[L^*, u^*, v^*]$ (frequently given without the asterisks). The CIELAB coordinates are defined according to the following equations,

$$L^* = 116f(Y/Y_n) - 16 \tag{2.8}$$

$$a^* = 500\,[f(X/X_n) - f(Y/Y_n)] \tag{2.9}$$

$$b^* = 200\,[f(Y/Y_n) - f(Z/Z_n)] \tag{2.10}$$

where $[X, Y, Z]$ are the tristimulus values of the test colour stimulus, and $[X_n, Y_n, Z_n]$ are the tristimulus values of a reference, white, colour stimulus and f as defined in eq. (2.11).

$$f(k) = \begin{cases} k^{1/3} & if k > (24/16)^3 \\ (841/108)k + 16/116 & if k \leq (24/16)^3 \end{cases} \tag{2.11}$$

CIE-1976 $L^*u^*v^*$ Colour Space (CIELUV)

The coordinates of the CIELUV colour space are defined in a similar way to the CIELAB colour space, using the XYZ tristimulus values. The L^* coordinate is identical in both colour spaces, and the u^* and v^* coordinates are defined as follows,

$$u^* = 13L^*(u' - u'_n) \qquad v^* = 13L^*(v' - v'_n) \tag{2.12}$$

where $[u', v']$ and $[u'_n, v'_n]$ are the CIE-1976 UCS coordinates of the test stimulus and the reference white colour stimulus respectively.

As CIELAB and CIELUV colour spaces are defined to be uniform, the Euclidean distance in this space can be used to quantify the approximate magnitude of the perceived colour deviations. These

are defined as the colour difference metrics ΔE^*_{ab} and ΔE^*_{uv} respectively and are given by eqs. (2.13) and (2.14).

$$\Delta E^*_{uv} = \left[(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2\right]^{1/2} \tag{2.13}$$

$$\Delta E^*_{ab} = \left[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2\right]^{1/2} \tag{2.14}$$

2.7.7 CIE Standard Illuminants

Alongside the colour spaces for quantifying colour stimuli, the CIE also provides definitions for standard light sources. The *CIE standard illuminants* are a set of standardised light sources with well defined SPDs. These are used as benchmarks and reference points, respect to which other colour phenomena are described. There are different illuminant types defined by the CIE (e.g., 'A', 'B', 'F', etc.) for simulating different lighting conditions. For example, 'illuminant A' simulates the SPD of a Tungsten-filament ('incandescent') bulb [10, p. 44-45]. In the studies that follow, we will be particularly concerned with the D series sources, more specifically *D*65. The SPD of the *D*65 source represents the average daylight conditions on the surface of the earth [10, p. 43].

2.7.8 Metamerism and Metameric Breakdown

The key concept underlying the trichromatic generalisation is metamerism. Metamerism is the effect where two quantifiably different spectra are perceived to be identical [10, p. 70]. In the case of the colour matching experiments outlined in section 2.7.1, each monochromatic spectrum was matched by an additive mixture of three different spectra, and there was no requirement for these three primary spectra to contain the spectral components of the monochromatic spectrum. These matching spectral pairs are said to be metameric to each other.

On the flip side, situations where metamerism is broken is referred to as 'metameric breakdown'. This occurs in two key ways; due to variations in the observer or due to variations in the light source [10, p. 70-72]. As outlined in section 2.5.1, the spectral sensitivities of peoples eyes vary due to a variety of factors. In cases where there is a substantial variation between the spectral sensitivities of two observers, a condition can arise where they do not produce the same colour matches (i.e., they perceive the same spectrum as different 'colours'). This is referred to as *observer metameric breakdown*. This is explored in detail in chapter 4.

A similar effect can be observed under a change in illuminant. As outlined in section 2.4, the spectral content of the source affects the perceived colour of objects illuminated by it. Consequently, a change in the illuminant could lead to the spectrum being reflected off objects to be different enough that the brain perceives it to be a different colour. This is termed *illuminant metameric breakdown*, and is relevant to the work presented in chapter 6.

2.8 Non-CIE Colour Models

The colour spaces outlined in section 2.7, allow one to quantify the similarity between different spectral stimuli, through tristimulus values and colour coordinates. However, they give no information about how to generate the stimuli from the colour coordinates. The role of a colour model is to specify colours in a way that they can be stored, transmitted, and reliably reproduced. Although we make a distinction between colour models and the CIE colour spaces, from a mathematical point of view they are identical to each other, in that and rely on the trichromatic generalisation to represent colour stimuli in a three-dimensional coordinate space.

The most commonly used colour models rely on the RGB system [53]. These colour spaces bare the most similarity to the CIE-RGB colour space, not only because they use 'red', 'green', and 'blue' primaries, but also because at the heart of the CIE-RGB colour space is the idea that, by mixing different proportions of these primary colours, a whole range of colour sensations can be achieved. In that sense, the RGB spaces are not concerned with the spectral properties of the light, but rather the perceived chromaticity. Just like how the meaning of the CIE-RGB colour space is bound to the exact wavelengths of the 'red', 'green', and 'blue' primaries used in the colour matching experiments, the RGB colour spaces are also bound to a standard triple of 'red', 'green', and 'blue' primary colours, but they are not defined using their emission wavelengths, but rather their CIE-xy chromaticity coordinates [54].

As outlined in section 2.7.3, given three points on the chromaticity diagram, the area bounded by the lines joining those points represent all the colours that can be achieved by an additive mixture of the spectra with those individual chromaticity coordinates. Consequently, these colour spaces are agnostic to the actual spectra of the primaries, and the system designers are free to pick any spectral shape, as long as they are well-calibrated (i.e., they produce the required chromaticity coordinates and these values are stable over a range of intensities).

The most common RGB colour space in use is the HDTV standard, *ITU-R BT.709* (or simply *rec.709*) [54]. *sRGB* is another colour space that uses the same primaries as *rec.709*, and is used widely in web content. Generally, colours in the *sRGB* colour space are represented as three 8-bit values (i.e., three values, each in the range 0 to 255), representing the relative intensities of the three primaries, with 255 representing the maximum luminance of a given primary, according to the standard. It also uses the chromaticity of the D65 spectrum as its reference 'white point'. Figure 2.14 shows the *sRGB*/*rec.709* colour space graphically and fig. 2.11 shows the colour gamut of the *sRGB*/*rec.709* colour space on the CIE-xy chromaticity diagram. The end result is that, it is possible to encode colours digitally in a way that is reproducible. Consider an image being transmitted over the internet from the creator to a recipient. The image is composed of a set of *sRGB* values for each pixel, and provided that the recipient's display is calibrated the same way as the creator's (as the display standard ensures), the creator can be sure that the recipient sees the image the same way as they do (assuming they both have 'normal' colour vision).

We would mainly be concerned with colour representations in *sRGB*. However, in some cases colours will be given in *HSV*. *HSV* is a transformation of the RGB colour model where the three coordinates represent the colour's *hue*, *saturation*, and *value*. The *hue* coordinate represents the 'colour'

of the object in the colloquial sense (e.g., 'red', 'green', 'yellow', etc.) commonly given as an angle between 0 and 360 degrees [39, p. 88]. The saturation of the colour measures the 'colourfulness' of the hue. For example, a colour with low saturation will look more pale compared to a colour with a high saturation [39, p. 91]. Value measures the brightness of the colour [39, p. 88]. High 'value' colours emit more light than low 'value' colours. Sometimes 'lightness' is used instead of 'value', where it represents normalised brightness relative to white [39, p. 88]. Figure 2.15 shows the *HSV* colour space graphically. This colour space is useful when the properties of a colour needs to be isolated into *perceptually meaningful* dimensions.

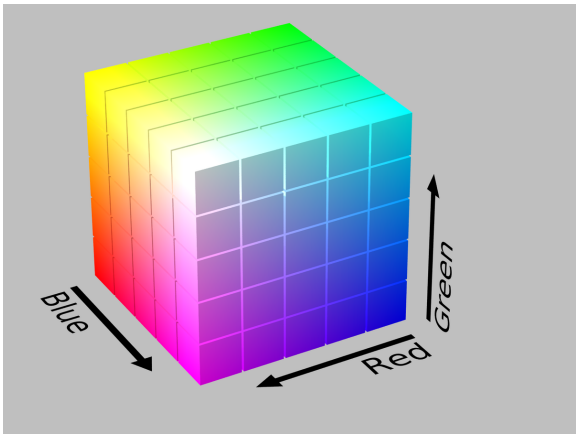


Fig. 2.14 *RGB* spaces represent the colours that can be achieved by an additive mix of red, green, and blue coloured primaries (from [11]).

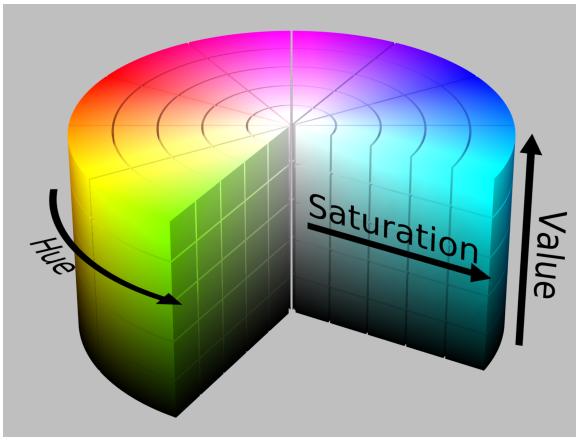


Fig. 2.15 The *HSV* space is a colour representation where the different coordinates have an intuitive relationship to the appearance of a colour (from [12]).

2.9 Semiconductor-based Lighting and Display Technologies

The focus of this work is on efficiently generating light through semiconductor technologies; displays and solid-state lighting devices (e.g., light bulbs made from light emitting diodes (LEDs)). As such a basic understanding of the internal processes occurring in these devices would be required to engage with the content that is presented in the chapters that follow.

2.9.1 Bulk-Semiconductor Light Emitting Diodes (LEDs)

The most common type of semiconductor-based light sources are p-n junction LEDs. They comprise two semiconductor layers, one doped to be n-type (i.e., have an abundance of electrons) and the other doped to be p-type (i.e., have an abundance of 'holes') [13, p. 350]. Holes are virtual particles that represent the absence of an electron, but behave as if they are real particles with a charge equal and opposite to the charge of an electron.

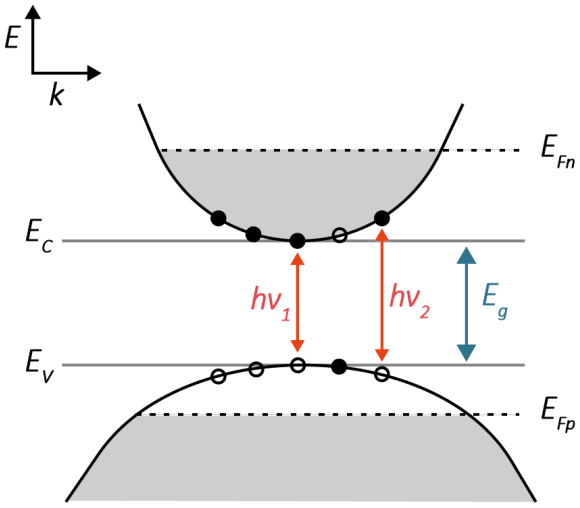


Fig. 2.16 Although most of the transitions in direct band gap semiconductors occur at $h\nu_1$, there are transition states around it that also lead to radiative recombination (e.g., $h\nu_2$) (after [13, p. 350]).

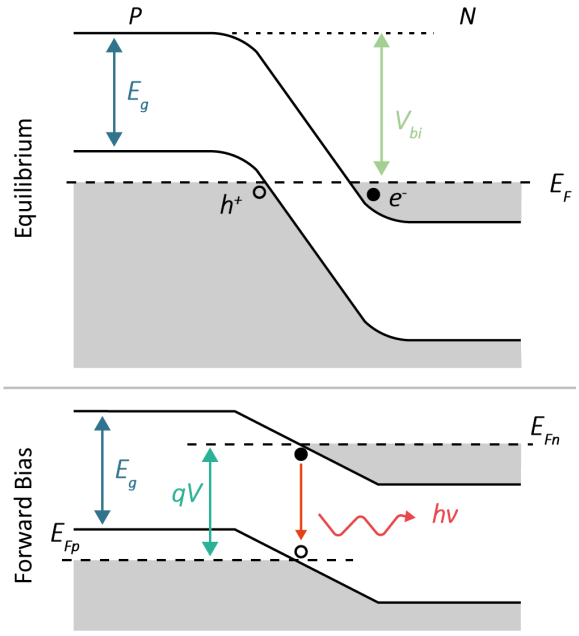


Fig. 2.17 A P-N junction has a built-in potential (top, V_{bi}). This potential can be overcome by an external voltage (V) which can inject carriers into the depletion region (middle regions with sloped lines) where radiative recombination can occur (bottom, $h\nu$). The vertical axis is energy (after [13, p. 350]).

When these layers are brought together, due to the imbalance of electrons, an electric field is built up between these regions, referred to as the 'built-in potential' with the region across which it exist being referred to as the 'depletion region' (see fig. 2.17 top). When a positive voltage is applied to the p-side (the 'anode') with a magnitude greater than the built-in potential, holes and electrons are injected from the valence band on the p-side and the conduction band on the n-side, to the depletion region, respectively. In the depletion region, the electrons and holes can *recombine* with a given probability. When that happens, the electrons transition from the conduction band to the valance band and releases the difference in energy (E_g) as light (see fig. 2.17 bottom). The probability of recombination can be increased by introducing an intrinsic (or lightly doped) semiconductor layer between the P and N regions to make a P-I-N junction. This increases the width of the depletion region and consequently the residence time of electrons and holes, which increases the recombination probability.

As outlined in section 2.2, the energy of the light is related to its frequency through the Einstein relation $E = h\nu$. Thus, a device with a band gap of E_g, will emit light of with a frequency around $\nu = E_g/h$. However, fig. 2.16 shows what the band alignment looks like in a *direct band gap* p-n junction (i.e., a junction capable of emitting light; *radiative recombination*). As can be seen, there are many states around E_g, with different energy gaps, where electrons can recombine. Consequently, the actual spectral emission from the device will have a wide range of wavelengths centred around E_g/h.

Generally, LEDs are made with direct band gap semiconductors such as GaAs and GaAlAs because they have a higher recombination efficiency than indirect band gap semiconductors. In indirect bandgap semiconductors, the lowest energy state in the conduction band has a different crystal momentum to the highest energy state of the valence band. As a result, when electron-hole recombination occurs, some of the energy is lost as phonons (heat) to account for the difference in the crystal momentum, which leads to a lower recombination efficiency.

The band gap is defined by the electrochemical properties of the semiconductor material. However, it can be altered to a certain extent by alloying different semiconductor materials together, which in turn will change the emission wavelength [55].

LED Fabrication

Typical LED fabrication relies on epitaxial growth. This is the process of growing layers of crystals on top of crystalline substrate [55]. This starts out with the substrate wafer on which the device will be built (e.g., sapphire), which also provides mechanical strength for the device [55]. The device is built up by growing the different materials comprising the device, layer by layer through epitaxial growth techniques (e.g., metal-organic chemical vapour deposition (MOCVD)). Typically masking patterns are used when growing these layers to control the deposition areas [55]. This is important to keep in mind as we contrast these techniques with solution processing involved in QD based devices. Usually, a single wafer will have multiple individual LEDs which are then cut into individual devices and packaged with externally accessible electrical contacts.

2.9.2 Organic Light Emitting Diodes (OLEDs)

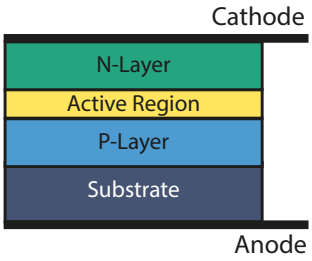


Fig. 2.18 The internal structure of an LED.

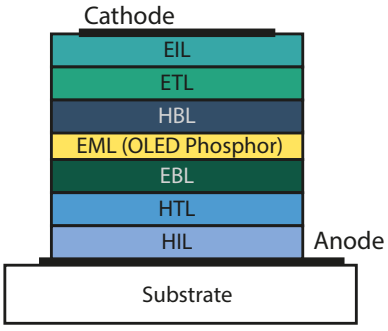


Fig. 2.19 The internal structure of an OLED.

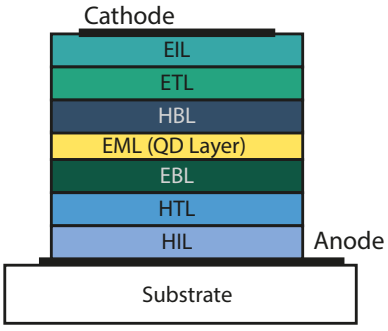


Fig. 2.20 The internal structure of a QD-LED.

Organic light emitting diodes (OLEDs) are LEDs that use organic material layers to convert electricity into light. In LEDs, the band gap is defined by the material properties of bulk semiconductors. However, organic molecules do not have a well-defined band gap. Instead, the electron occupancy of the different orbital levels act in an analogous way to the conduction and valence bands in bulk semiconductors; referred to as lowest unoccupied molecular orbital (LUMO) and highest occupied molecular orbital

(HOMO) respectively [56]. Analogous to the p and n type regions in an LEDs, OLEDs use hole transfer layers (HTLs) and electron transfer layers (ETLs) that are made from materials that are better at conducting holes and electrons respectively (see fig. 2.19). Similar to an LED, when a positive potential is applied across the OLED, electrons and holes flow into the device and recombine at the HTL-ETL interface and emit light. Modern OLEDs use a phosphorescent organic layer between the HTL and ETL that increases the efficiency of electro-optical transduction and allow for better control of the emission spectrum (see emissive layer (EML) in fig. 2.19) [56]. Finally, a hole blocking layer (HBL) and an electron blocking layer (EBL) are added to keep the electrons and holes in the EML for longer and improve the recombination probability [56].

OLED Fabrication

The fabrication starts in a similar way to LED fabrication, with the substrate (typically glass). Onto this substrate, the anode contacts, made up of a transparent conductive oxide (e.g., Indium Tin Oxide (ITO)), are patterned using photolithography. The organic layers can be deposited onto this prepared substrate through vacuum thermal evaporation (VTE) (with shadow masks) or solution processing techniques (e.g., spin coating, transfer printing) [56]. Finally, the cathode contact is deposited through thermal evaporation (with a shadow mask). The finished device usually needs to be hermetically sealed to minimise degradation over time.

Advantages and Disadvantages

In contrast to LEDs, OLEDs are better suited for producing devices with large emissive areas. They also exhibit fast switching speeds, low power consumption, and vivid colour output. As a result they have seen extensive use in current-generation high-end display devices. The ‘thin-film’ architecture of OLEDs also makes them suitable for novel applications like transparent and flexible displays [56]. However, for general lighting, the technology of choice has thus far been LEDs due to the limited lifetime, brightness, and luminous efficiency of OLEDs [57].

2.9.3 Quantum-Dot Light Emitting Diodes (QDs)

QDs are a class of 0-D semiconducting nanoparticles that exhibit 3-D quantum confinement of electrons. At very small diameters, the electronic band structure in nanoparticles become discretised, due to quantum mechanical effects (see fig. 2.22) [15, p. 2]. This discretisation restricts the allowed electron transition states. As outlined before, the emission wavelength from radiative recombination depends on the energy difference in electron transitions. Consequently, this leads to QDs having well defined, narrow emission spectra [58]. Contrary to LEDs and OLEDs, the allowed electron transitions in QDs are a function of their diameter. Thus, the optoelectronic properties of QDs can be tuned in a continuous fashion by controlling their diameters during fabrication [59]. In addition to this, the 0-D nature of QDs makes them *pseudo–molecular* structures without long range lattice order. This allows recombination to

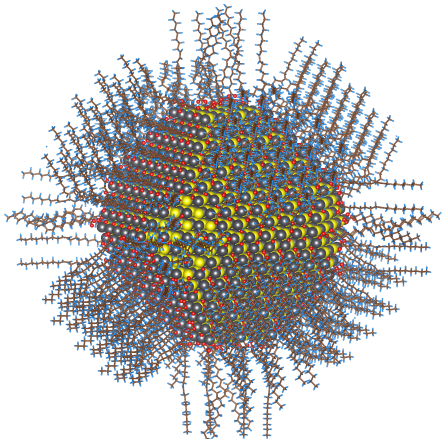


Fig. 2.21 QDs are semiconducting nanoparticles that can confine electrons in three dimensions, comprising a molecular core surrounded by ligands (from [14]).

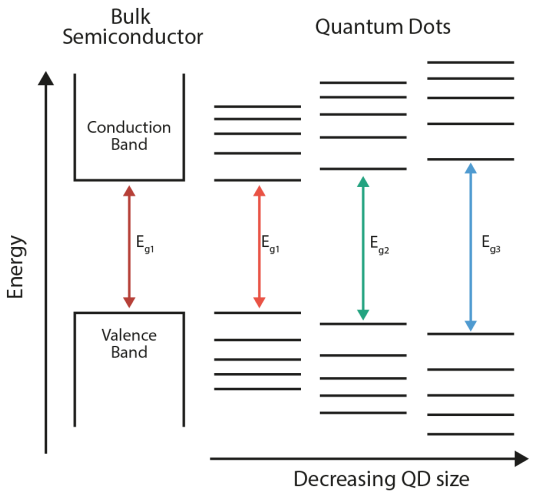


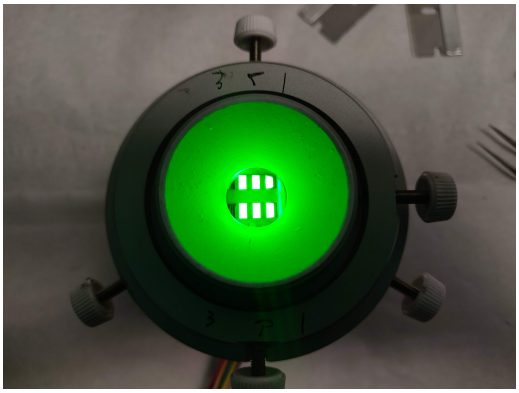
Fig. 2.22 The band structure of a QD is quantised, which restricts the allowed electron transitions and the band gap is a function of the QD size (after [15, p. 3] and [16]).

occur without the need for crystal momentum matching, which makes QDs have high recombination efficiencies. Common QD materials include, CdSe, ZnS, and InP.

The structure of quantum dot light emitting diodes (QD-LEDs) are similar to OLEDs (see fig. 2.20). In the case of QD-LEDs, the EML is replaced by a layer of QDs. This layer acts as the electro-optical transducer, encouraging radiative recombination in the electrons and holes that enter it.



(a) When colloidal QDs are illuminated with UV radiation, they emit visible light (photoluminescence)(from [16]).



(b) QDs can also be made to emit light when an electric potential is applied across them (electro-luminescence).

Fig. 2.23 QDs can excited (a) optically or (b) electrically.

QD-LED Fabrication

QDs themselves can be fabricated in a myriad of ways, however, to be compatible with solution processing techniques, QDs are usually required in a colloidal state. This can be done through traditional colloidal synthesis techniques, which results in a solution with QDs suspended in it [15, 59].

Due to the similarity to OLEDs, QD-LEDs can be fabricated in a similar way through solution processing techniques (e.g., spin coating). In addition to this, it has been shown that QD-LEDs can be fabricated through photolithography, transfer printing, and inkjet printing processes [60–64].

Advantages and Disadvantages

Many of the advantages of OLEDs applies to QD-LEDs. The key advantage QD provides is the tunability of emission wavelengths [65]. Figure 2.23a show a range of QDs in colloidal state emitting light through photoluminescence. Furthermore, the narrow emission spectrum from a QD-LED results in higher colour purity, which is particularly important for next-generation wide colour gamut (WCG) displays that require primaries with high colour purity [30]. QDs have already seen use as optical down-converters in place of RGB colour filters in consumer TVs [66]. One of the key challenges that has limited the adoption of QDs based devices has been the low emission efficiency of non-toxic QDs. Many of the highly efficient QD structures use toxic-heavy metals such as Cadmium which makes them unsuitable for consumer products [67]. This is compounded by the short lifetimes of the current state-of-the-art QDs [67].

2.9.4 Display Types

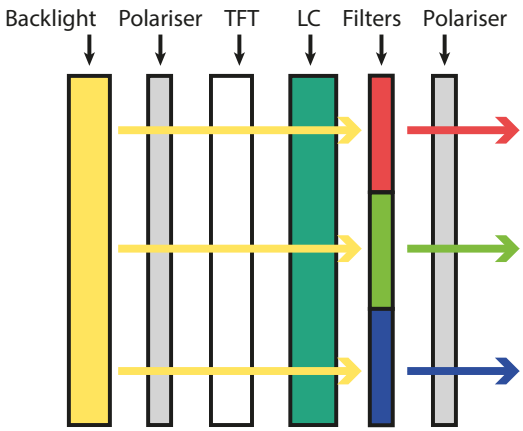


Fig. 2.24 LCDs are composed of a backlight, layers of optical polarisers, colour filters, and an electrically controllable layer of liquid crystals (LC) (after [17, 18]).

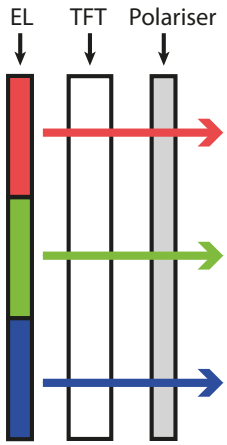


Fig. 2.25 Unlike LCDs, OLED/QD-LED displays have a simpler structure as they do not require a complex optical assembly to control the light emission (EL= Emissive layer) (after [17, 18]).

Liquid Crystal Displays

Liquid crystal displays (LCDs) are arguably the most common type of flat-panel displays in use currently. They comprise a white backlight, a set of polarising filters, a set of light diffusers, an array of colour filters, and an array of individually controllable liquid crystals (LCs) (see fig. 2.24). Each pixel on the display comprises three sub-pixels (usually red, green, and blue), and the region on LC corresponding to each sub-pixel can be individually controlled through a layer of thin film transistors (TFTs). The exact number/order of filters and the relative orientation of the top and bottom polarisers depends on the design of the display and the liquid crystal technology used [68–70].

The light from the backlight is initially polarised and homogenised through a sequence of polarising filters and light diffusers, before passing through the liquid crystal layer and the colour filters. The structure of these LCs are such that they can change the polarisation of the light and they can respond to an externally applied electric field ('dielectric anisotropy') [71]. Through the TFTs, the orientation of these crystals, and thus the polarisation of the light, can be electronically controlled. When a given sub-pixel is off the LCs are aligned such that the polarisation of the light exiting the LC layer is orthogonal to the polarisation of the polarising filter that follows the LC layer (i.e., the polarising layer will block all the light exiting the LC layer). This changes accordingly for when the sub-pixel is fully lit (polarisation aligned with the filter, letting all the light through), or when the sub-pixel has an intermediate brightness (polarisation angle between 0° and 90°, letting only part of the light through). This allows LCDs to control the brightness of individual sub-pixels and control the colours being displayed on-screen. Note that the backlight is always turned on and the output brightness is controlled by blocking a portion of that light. Traditional LCDs tended to use cold cathode fluorescent lamps (CCFLs) and newer displays that are marketed as 'LED displays' generally use LEDs for the backlight and do not use LEDs to generate the coloured light directly [72].

Emissive Pixel Displays: LEDs, OLEDs, and QD-LEDs

The solid-state processing techniques and the optical properties of the semiconductor materials used in LED fabrication, limits their ability to be miniaturised which in turn places a lower limit on their power demands [24–26]. As a result, LED based emissive displays have usually been restricted to the domain of large area, high brightness displays where the pixels per inch (PPI) resolution is not as crucial (e.g., road signs and large advertisement panels) [73].

Figure 2.25 outlines a typical structure of a OLED/QD-LED display. The key difference between these displays and LCDs is the absence of a backlight. These displays leverage the emissive properties of OLEDs/QD-LEDs and their capacity to be processed into thin-films, to create pixel arrays that can emit light by itself (hence the phrase *emissive pixel display*). As before a TFT backplane is used to control the individual sub-pixels. Consequently, the power consumption of the display is a function of the content being displayed. In the extreme case, when the display is showing a black image, its power consumption can be reduced all the way to zero. This makes these displays much more energy efficient. In chapter 5 we show how to leverage this property to achieve display power savings by optimising the

content on the screen. Furthermore, these displays have a much simpler structure compared to LCDs, which allows OLED/QD-LED displays to be made much lighter and thinner than LCDs.

2.10 Contextual Summary

This chapter provided an outline of the basic theories and ideas that underline the work presented in the chapters that follow. Looking ahead; chapter 3 explores the statistics of a global digital colour matching dataset and evaluates how to extract individual observer CMFs from this data. Chapter 4 explores how to leverage metamerism to achieve display power savings and how this can be implemented using a multi-primary display (MPD) architecture. It will also evaluate the efficacy of using MPDs to minimise *observer metameric breakdown.* Chapter 5 details how to leverage the dynamic power consumption of emissive pixel displays to achieve display power savings by optimising the content on the screen. Finally, chapter 6 presents a set of designs for QD based lights that aim to maximise the CRIs by matching the SPD of daylight.

All models are wrong, but some are useful.

— Dr. George Box [77].

Chapter 3

Individual CMF Inference from a Global Colour Matching Dataset

3.1 Introduction

Colour matching experiments have been the foundation for modern colourimetry since the pioneering work by Maxwell, Young, and Helmholtz [10]. In these experiments, the participants try to match the appearance of a monochromatic light source with a spectrum composed of a mixture of controllable primary spectral emissions, as detailed in section 2.7.1 and illustrated in fig. 2.7 [74, 44]. However, as might be expected, these in-person user studies are expensive and time consuming to conduct.

In 2015, the design studio *PepRally* released *Specimen*, a colour matching game for iOS [75]. In the game, the players are presented with coloured 'blobs' (referred to as 'specimens') on a coloured background and are asked to pick the blob with the colour that matches the background (see fig. 3.1). If the player correctly picks the matching 'specimen', the chosen 'specimen' is removed from the screen and they would be presented with a different background colour. The process is repeated till there are no 'specimens' left to match.

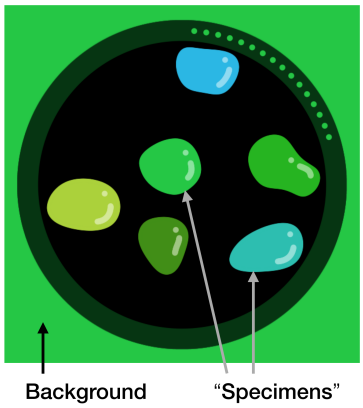


Fig. 3.1 In *Specimen*, the players pick which coloured blob (*'specimen'*) matches the background colour.

Although it might not seem so at first glance, the structure of the *Specimen* game has striking similarities to traditional colour matching experiments. Each correct or incorrect match a player makes gives an insight into the way they see colour. According to the designers, *Specimen* was designed with the auxiliary goal of investigating human colour perception [76]. To that end, *Specimen* was designed to anonymously log the interactions the players made with the game. The entirety of the collected player logs has information on 28.6 million colour selection events.

This chapter outlines the results of an analysis carried out on this data, henceforth referred to as the '*Specimen dataset*'. The primary focus of the first half of the chapter is on assessing what kind of inferences can be made using the results of this *online* colour matching experiment. We present a summary of the data contained within the *Specimen dataset* and explore variations in colour acuity among players from the United Kingdom. The second half of the chapter is concerned with using this data to infer the individual players' spectral sensitivities. We present a potential method for extracting individual CMFs and evaluate the results obtained.

3.1.1 Contributions

In this chapter we present the following contributions:

- Provide details of the information contained in the *Specimen dataset*, a dataset of 28 million+ colour matches.
- Present Statistics about colour confusion rates of players from around the world based on this data.
- Evaluate in detail the correlates to colour misidentification in the players from the UK.
- Provide a mathematical framework for extracting individual CMFs from this colour matching data.
- Implementation of the mathematical framework using *TensorFlow*'s optimizer backend.
- Demonstration that the implementation of the framework, applied to data from the *Specimen* dataset, and across four different priors (uniform, gaussian triple, 2° std. observer, and 10° std. observer), is able to find CMFs that better explain the observed colour confusions.
- Present a detailed, physiologically relevant mathematical model for predicting the ocular properties of the players in the *Specimen dataset*.
- Demonstration that it is possible to potentially learn the ocular properties and the CMFs from the presented frameworks.

3.1.2 The *Specimen Dataset*

The *Specimen dataset* is the entirety of all the anonymised player activity logs collected by *Specimen*, the game. The logs store details about the game state whenever an 'event' occurs in the game (e.g., the start of a new game session, selection events, player purchasing microtransactions, etc.). For each 'event' occurring in the app, it records the following information.

- Time of event : as a UNIX timestamp.

- Type of event : selection event, purchase event, 'game start' event, etc.
- Play session start time and duration.
- For selection events, the chosen colour and the background colour : As *sRGB*, *CIELAB*, and HSV values.
- Anonymised player ID.
- Device model.
- Location : as an ISO country code.

The dataset contains event logs from 21:33:00 (GMT) on Monday, 15^{th} of June 2015, to 16:06:02 (GMT) on Saturday, 9^{th} of December 2017. The complete dataset contains details on 28.6 million colour matches from 41,000 players from 175 countries using 96 different iOS-based devices, split across 489,580 play sessions.

Uses of the *Specimen* Data in the Literature

There are two reported uses of this dataset in the literature. In 2018, Cambronero et al. devised an incremental colour quantisation algorithm for image compression for colour vision deficient observers [78]. The authors looked at individual players' colour confusions to identify potential colour vision deficient players. By considering the confused colours as potential candidates for colour mergers, the authors showed that approximately 22%–29% extra space saving could be achieved over the state of the PNG image compressors, `TinyPNG` and `pngquant` [79, 80].

The other use is by Stanley-Marbell et al., also in 2018, where the authors presented *Ishihara*, a language for 2D graphics with a notion of perceptual equivalence [81]. Through *Ishihara*, programs can define perceptually equivalent colours for a particular application, which is then used to transform the colours in the program to reduce the power dissipation of the display. The *Specimen dataset* is used in this work to create an empirical model of the perceived-colour equivalence classes, which forms the core of *Ishihara's perceived-colour semantics*.

Meta of the Dataset

Before an analysis can be done on the dataset, the following can be said about the information contained within it. This is to ensure that the reader is aware of the limitations of the data and the potential thresholds for interpretation, after which, extrapolation of the results will be meaningless and potentially dishonest. It should be stated that the sample constituting the players of *Specimen* is likely biased and is unlikely to be representative of the population of a country, thus the results (and inferences made on them) are in no way indicative of the properties of this population.

The obvious commonality between the players is that they all have access to an *iOS* device and as well as to the internet. In 'developed' countries, given the geographic variation of the market share of iOS-based devices [82], the sample is more likely to be representative of the population as a whole. However, the sample is also restricted to the people that are interested in playing *casual* mobile games.

Thus, one should be weary of using the dataset to make general statements about the population of a particular demographic. Consequently, the analysis that follows will be intentionally limited to only making qualified observations about the trends apparent in the data. However, considering these restrictions, due to the large number of colour match observations spanning multiple geographic regions, first-order inferences about the global statistics of the colour acuity variability in human vision can be made.

On a macro level, the information contained in the dataset can be segmented into three key areas;

1. Details about perceived colour and chosen colour.
2. Details about micro-transaction events.
3. Information about where (country, and device) and when (local time) those events occurred.

The first gives details about the perceptual acuity of the players and how it varies with the properties of the colour but also with other measurable properties, like time of day and the type of device. The second can be used to identify when players are most susceptible micro-transactions. As this information can be misused to manipulate players we will not be exploring it in this work, on ethical grounds. Finally, the geographic data will be useful in normalising the statistics (e.g., by population or to adjust the time of the events to be in the local time zones).

Given that the luminous environment is not well controlled, it can be said that the data from *Specimen* is probably not a good enough substitute for in-person experiments. However, given the large number of samples, we believe that it might be possible to make approximate predictions about the colour acuity of individual players, and we present a proof of concept of a potential method for doing so.

With this in mind, the answers to the following questions are explored in this chapter:

1. How does the prevalence of players and playtime vary between countries?
2. How does player behaviour change throughout the day?
3. Does the colour acuity of players change throughout the day?
4. How does the colour acuity vary with the *Specimen* colour?

3.2 Global Statistics

As summarised before, the dataset contains information from 41,000 players from 175 countries. Focusing on the geographic distribution of the players, the data shows that China had the most number of players (26,087 players), followed by the United Kingdom (14,953 players), the United States of America (14,725 players), France (9,786 players), and Italy (7,081 players). Table 3.1 shows the player counts for the 25 countries with most number of players. Figures 3.2a and 3.2b shows the player counts for the top 100 countries by player count.

3.2.1 Population Adjusted Player Counts

Figure 3.2c shows the top 100 countries by the number of players, adjusted for the population of the country, *players per 100,000 people*. This figure does not include details about Andorra, Montenegro,

Table 3.1 Top 25 countries by *player count*.

Country	Players
China	26,087
United Kingdom	14,953
United States of America	14,725
France	9,786
Italy	7,081
Germany	7,050
Russian Federation	4,667
Mexico	3,964
Spain	3,482
Turkey	3,022
Brazil	2,807
Netherlands	2,476
India	2,418
Poland	2,205
Australia	1,932
Romania	1,911
Taiwan	1,733
Switzerland	1,623
Ukraine	1,524
Argentina	1,510
Canada	1,437
Belgium	1,367
Denmark	1,313
Philippines	1,286
Sweden	1,187

Table 3.2 Top 25 countries by *average play time*.

Country	Avg. Play Time
Bangladesh	6 min 41.4 s
Spain	6 min 16.2 s
Saint Lucia	5 min 39.8 s
Yemen	5 min 27.3 s
Trinidad and Tobago	5 min 18.9 s
Republic of Macedonia	4 min 32.6 s
Cayman Islands	4 min 29.2 s
Fiji	4 min 6.6 s
Cambodia	3 min 55.9 s
Barbados	3 min 55.5 s
Wallis and Futuna	3 min 46.3 s
Cape Verde	3 min 42.7 s
Guyana	3 min 37.8 s
Bermuda	3 min 31.8 s
Liechtenstein	3 min 30.6 s
Vietnam	3 min 24.6 s
Andorra	3 min 24.5 s
Uzbekistan	3 min 24.1 s
Monaco	3 min 19.8 s
Cuba	3 min 19.3 s
Sweden	3 min 18.8 s
Singapore	3 min 17.4 s
Guadeloupe	3 min 15.3 s
Canada	3 min 12.5 s
French Polynesia	3 min 11.4 s

Myanmar, and Palestine due to lack of reliable population data [83]. According to the data, the top 5 countries with most players per 100,000 people in the population were Wallis and Futuna, Greenland, Iceland, Anguilla, and Macau with 91.36, 55.07, 48.15, 44.60, and 39.62 players per 100,000 people, respectively. Table 3.3 summarises the results for the top 25 countries by players per 100,000 people.

3.2.2 Mean Play Time

Globally, the mean play session length was 2 minutes and 45.6 seconds and the median session length was 1 minutes and 38.2 seconds. Aggregating the play time data by country, the top 5 countries in order from the longest mean play time were, Bangladesh (6 minutes and 41.4 seconds), Spain (6 minutes and 16.2 seconds), Saint Lucia (5 minutes and 39.8 seconds), Yemen (5 minutes and 27.3 seconds), and Trinidad and Tobago (5 minutes and 18.9 seconds). Table 3.1 shows the mean play times for the 25 countries with the highest mean playtime per player. Figure 3.3 shows the distribution of the 100 countries with the longest mean play time per player.

Table 3.3 Top 25 countries by *players per 100k people*.

	Country	Players per 100k
1	Wallis and Futuna	91.36
2	Greenland	55.07
3	Iceland	48.15
4	Anguilla	44.60
5	Macau	39.62
6	San Marino	36.65
7	Luxembourg	26.38
8	Cayman Islands	23.44
9	United kingdom	23.33
10	Denmark	23.22
11	Malta	22.59
12	Liechtenstein	21.54
13	Gibraltar	20.00
14	Switzerland	19.83
15	Singapore	19.64
16	Bermuda	18.68
17	Faroe Islands	18.52
18	New Caledonia	17.49
19	France	14.81
20	Netherlands	14.67
21	Hong Kong	13.99
22	Sweden	12.19
23	Belgium	12.18
24	Cyprus	12.12
25	Finland	11.92

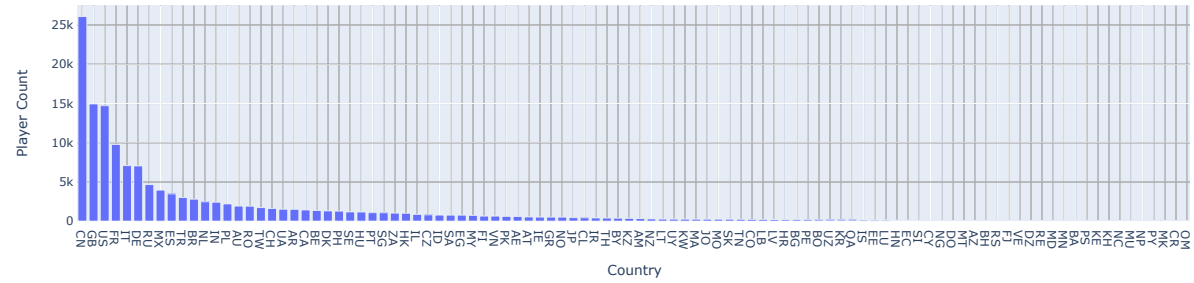
3.3 Player Behaviour (Players from the UK)

The dataset uses *UNIX time* or *milliseconds-since-epoch* notation, relative to the GMT/UCT time-zone, to record time of events. Since the geographical information is encoded as a string with the ISO country code, and many larger countries span multiple time-zones, extracting relationships related to the time of events becomes difficult and unreliable. In this initial analysis, we explored the behaviour of players from the UK as this circumvents the need for time conversion. Furthermore, there is a large amount of player data from UK based players (2^{nd} highest) as well as a high prevalence of players (9^{th} highest by number of players adjusted for population) which would increase the likelihood of the data being a more representative sample.

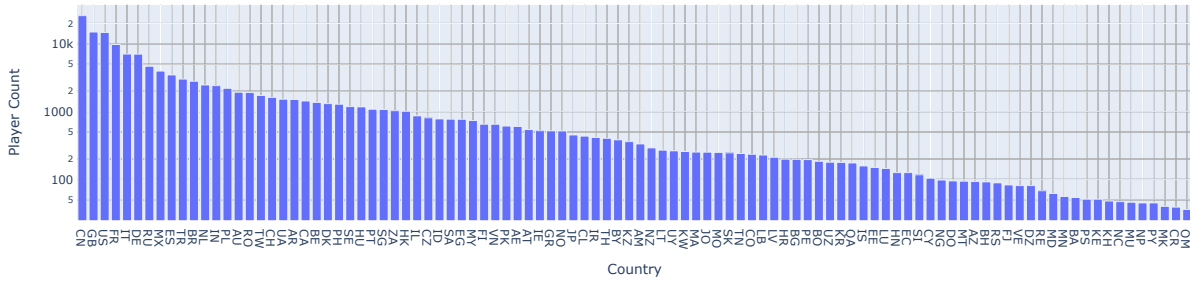
3.3.1 Time Dependency

There are 787,494 colour selection events from 14,953 play sessions attributed to 4,737 players from the UK. Out of the 787,494 colour selection events, 118,092 were incorrectly matched giving an average accuracy of 85.0%. The mean play session length was 2 minutes and 30 seconds and the median session length was 1 minutes and 35 seconds.

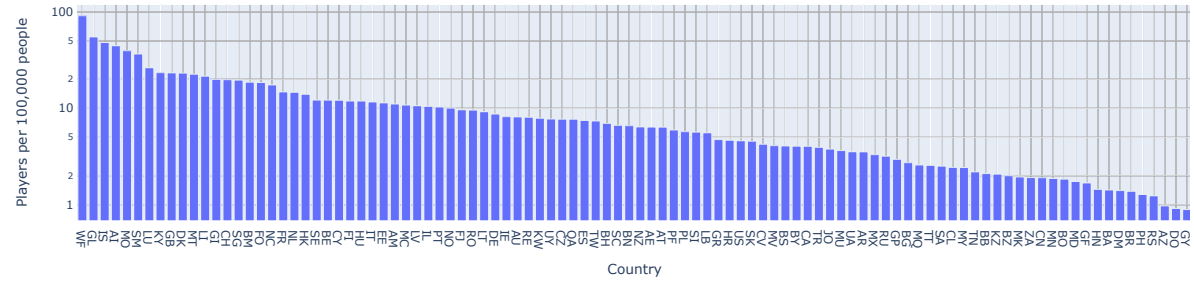
The data shows that the frequency of colour selection events have a cyclic relationship with the time of day (see fig. 3.4). From 0500h onwards, the number of selection events gradually increase (almost linearly from 0800h onwards) from around 2,500 per hour to a maximum of 63,000 per hour around 2200h. Following this the selection events drop off rapidly till 0500h. We believe that this drop off is correlated with the time when players go to sleep. This could be used as a heuristic infer approximately when the players go to bed around the world. The number of selection events, total session length and the raw number of incorrect matches follow the same trend, as they have the same underlying cause; the frequency of play.



(a) Linear vertical axis shows that most players were from the top 5-10 countries



(b) Logarithmic vertical axis shows that most of the top 100 countries had 100-1000 players.



(c) Adjusting player count for population shows that the prevalence of players is less than 100 per 100,000 people for all the countries (logarithmic vertical axis).

Fig. 3.2 Number of players by Country ((a) on a linear scale, (b) on a logarithmic scale, (c) player counts adjusted by country population).

Figure 3.5 shows the incorrect match frequency normalised to the player activity (i.e., frequency of selection events) to remove the variation due to change in play frequency. The plots show that the frequency of mismatches does not change substantially over the course of the day. Although, there is a slight downward trend from 16.04% at the start of the cycle (0600h) to 14.52% by the end of the cycle (0300h). The proportion of mismatches in the hours following 0400h and 0500h are outliers to this trend, with 16.07% and 13.44% mismatches.

As outlined before, the mean play session length is about 2 minutes and 30 seconds and is quite consistent throughout the day (see fig. 3.6). However, towards the end of the day (2200h to 0300h) the session length increases gradually to nearly 3 minutes and 20 seconds. The median play time distribution further illustrates the time-invariant nature of play time (see fig. 3.7). This indicates that, for *Specimen*, the observed player behaviour is to play the game in short bursts throughout the day and not engage with the game for extended periods of time. This could explain why the play activity distribution (see fig. 3.4)

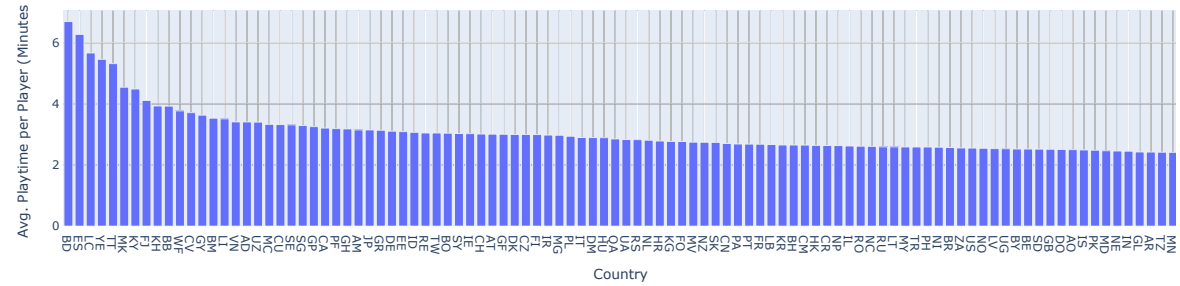


Fig. 3.3 Mean Playtime by country shows that most players played the game for around 2 minutes at a stretch.

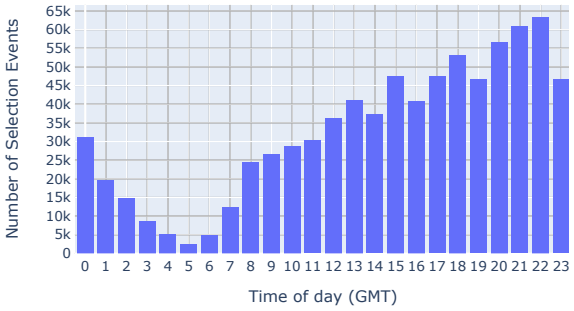


Fig. 3.4 Frequency of selection events follows a daily cycle with the player activity peaking around midnight.

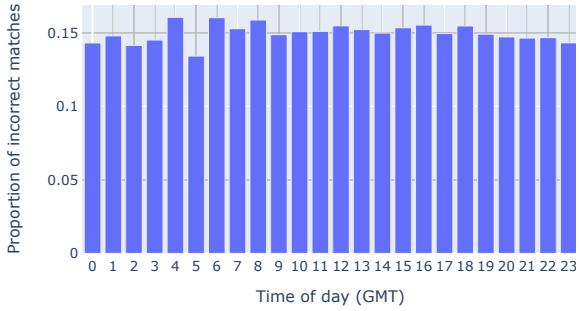


Fig. 3.5 The rate of incorrect colour matches, as a proportion of the number of selections, is minimally variant over the course of the day.

shows a substantial amount of activity even throughout the workday. As before, the play time during the hour commencing 0400h show a substantially shorter mean and median play time. We believe that this might be due to sampling bias, given the (relatively) small number of selection events in that time period.

3.4 Colour Acuity

One of the questions posed at the start of the chapter is on whether the colour acuity of players vary with the colour of the 'specimen'. Selection accuracy can be used as a proxy for this, as incorrect selections correspond to colour confusions, and thus, indicate where the players' ability to discriminate colours is limited.

3.4.1 Sensitivity to *hue*

Figure 3.8a shows a histogram of confusion frequency against the colour *hue*, binned at intervals of 5 *hue* steps. The vertical dotted lines indicate the hues of the three *sRGB* primaries. The peaks between *hue* values of 50 and 100 reveal that the sensitivity is lowest for 'green' (i.e., 'green' hues are confused more frequently than other hues). However, when looking at how frequently each hue was presented and

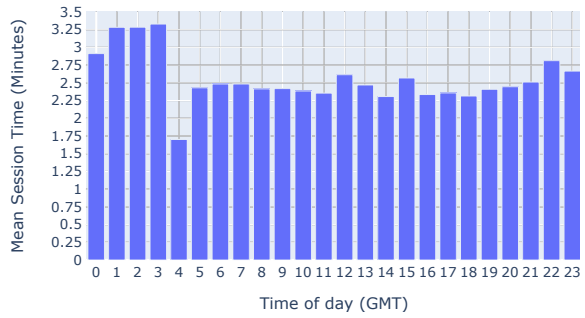


Fig. 3.6 Mean session length is around 2 minutes, with players playing after midnight playing for approximately 1 minute longer than average.

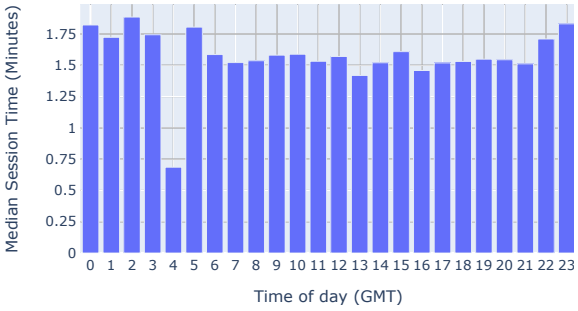
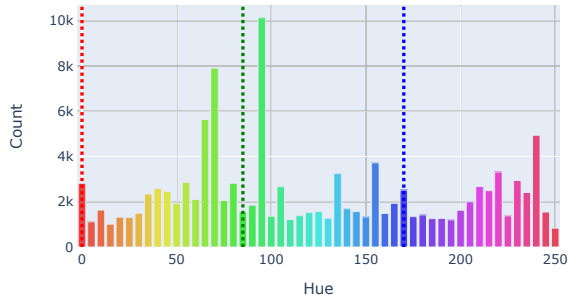


Fig. 3.7 Median session length shows that the mean session length variation after midnight is biased by a few players playing significantly longer than average.

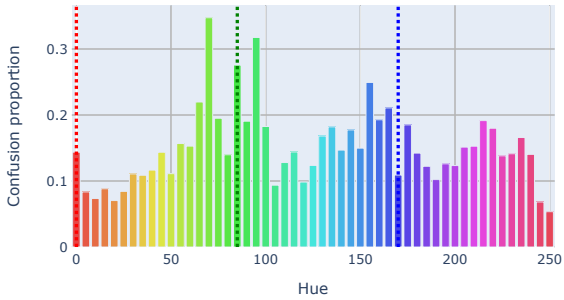
by normalising the values to factor out the variations due to this, we observed that this causes the overall deviation to be not as extreme as was initially apparent (see fig. 3.8b). The normalised plot shows that the peaks in the 50–100 *hue* range are followed by a region of confusions that primarily constitutes 'blue' hues.

When an incorrect choice is made, the *Specimen* game allows the players further attempts to try and pick the matching 'specimen'. Incorrect choices the player makes after initially picking the wrong colour are referred to as *second-order confusions*. The information contained within the *Specimen dataset* allows us to identify these second-order confusions. Curiously, when looking these second-order confusions, the disproportionate difficultly in correctly matching 'green' hues becomes even more substantial (see fig. 3.9).

A potential hypothesis, beyond the inherent variation in the colour acuity of human vision, is that the regions of peak confusion correspond to hues that are dominated by a single primary. The primary *hue* lines (the dashed lines in fig. 3.8) shows some evidence of this for the 'blue' and 'green' primaries, but not for the 'red' primary.

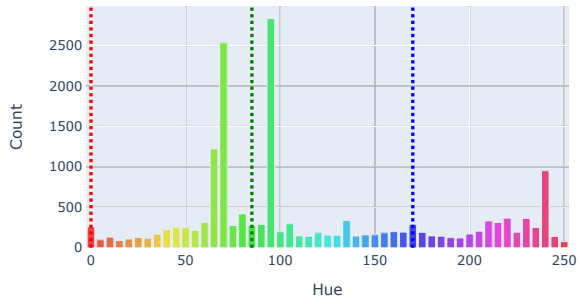


(a) The distribution of the raw number of colour confusions by *hue*.

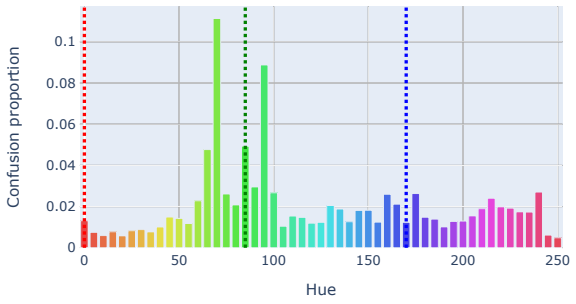


(b) Number of colour confusions as a proportion of the how often the *hue* was presented.

Fig. 3.8 Overall, the players more easily confused 'green' and 'blue' hues compared to others (right). Each bar spans 5 *hue* units and the dashed vertical lines indicate the hues of the *sRGB* primaries.



(a) The distribution of the raw number of 'second-order' colour confusions.

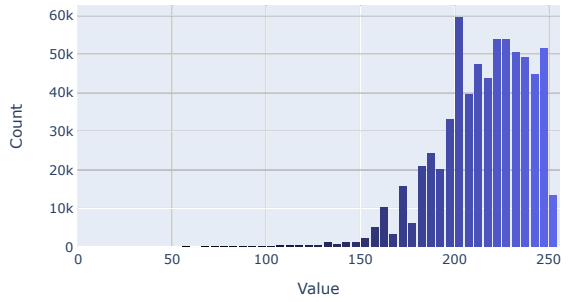


(b) Number of 'second-order' colour confusions as a proportion of the how often the *hue* was presented.

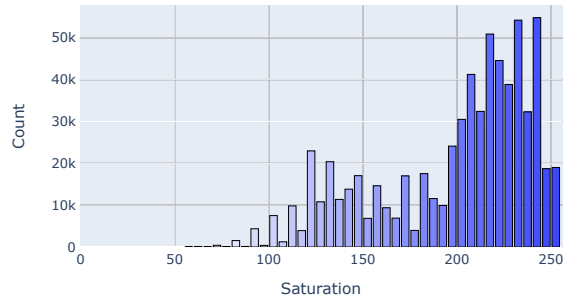
Fig. 3.9 The 'second-order' confusion distribution shows that 'green' hues are more likely to be confused repeatedly. Each bar spans 5 *hue* units and the dashed vertical lines indicate the hues of the *sRGB* primaries.

3.4.2 Sensitivity to *saturation* and *Value*

A similar analysis can be done on the sensitivity of player accuracy to *saturation* and *value* of the colours. The game data shows a bias in the algorithm used to pick the presented colours, towards highly saturated colours with higher *value* with a lower limit of 50 for both properties (see fig. 3.10). After normalising the colour confusions to account for this bias, the data shows that the likelihood of a colour confusion by a player is inversely correlated with that colour's *value*; darker colours were harder to correctly match (see fig. 3.11a). The overall effect of *saturation* on colour confusions appear to be minimal but colours with very low *saturation* (notably in the range 55-60) are disproportionately more easily confused (see fig. 3.11b). However, given that there are only four observations in that *saturation* range, this extreme value is quite likely due to sampling bias.

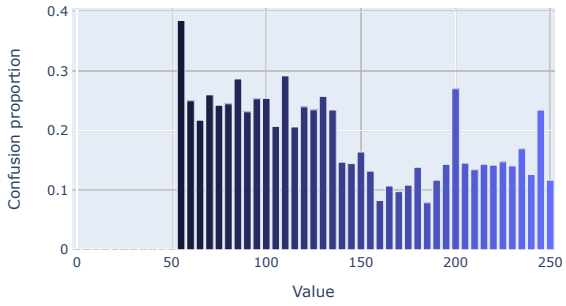


(a) Most of the colours presented had a *value* greater than 200.

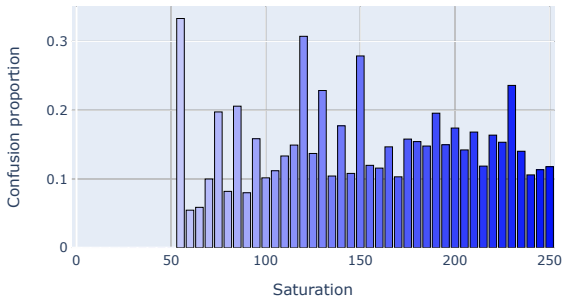


(b) The algorithm used by *Specimen* to pick colours show a bias towards highly saturated colours.

Fig. 3.10 The distribution of the colour *value* and *saturation* of the colours presented by the *Specimen* game.



(a) Players had a harder time correctly matching colours with low *value*.



(b) The difficulty of classifying colours show a slight negative correlation with the colours' *saturation*.

Fig. 3.11 Colour confusions normalised by how frequently each colour *value*/*saturation* was presented.

3.4.3 Colour Confusion Variation Over Time

As shown before, the mean selection accuracy of the players do not vary substantially over the course of the day (see fig. 3.5). However, a more granular question that could be explored with this data is whether the sensitivity to different colour properties change over the course of the day? For example, are some colours more difficult to discriminate in the evening compared to the morning?

Sensitivity to *Hue*

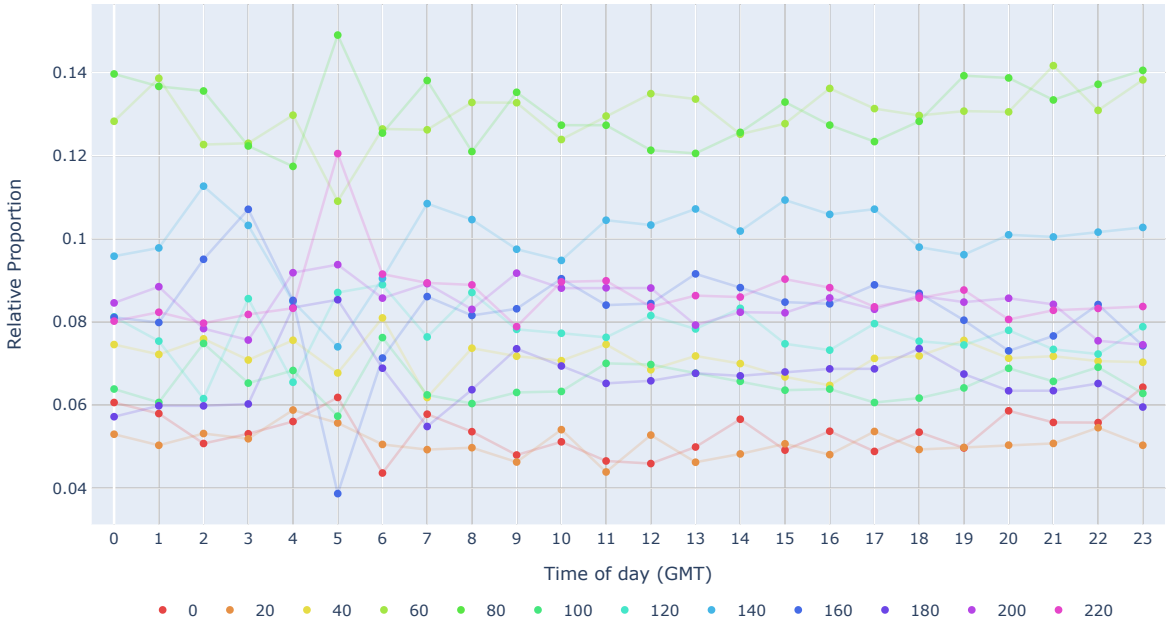


Fig. 3.12 The relative proportion of confused hues is not substantially affected by the time of day (each line spans 20 units of *hue* from the specified value).

Figure 3.12 shows the relative proportion of confused hues as a function of time marked by the one hour period starting from the time given on the x-axis. The individual lines represent a band of hues, 20

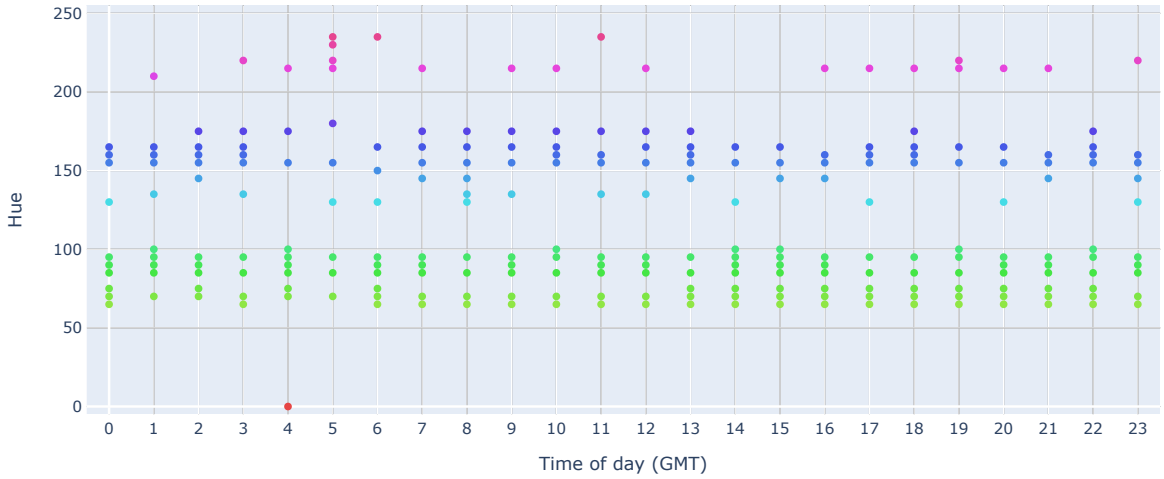


Fig. 3.13 Top 10 most confused hues are consistently confused regardless of the time of day (each point spans 5 units of *hue*).

hue units wide, starting from the value given on the legend. Figure 3.13 shows the 10 most frequently confused hues as a function of when the colour confusion was made (each point represents a range of 5 *hue* units). Both figures show that the likelihoods of confusion for different *hues* are not significantly affected by the time of day. However, incorrect matches made between 1300h and 1600h show a slightly increased bias towards 'greenish' hues. As the plot is showing the top-10 most frequent confusions, the increase in 'green' colours could be an artefact of the hard cut-off imposed by picking only the top-N colours.

Sensitivity to *saturation* and *value*

A similar analysis can be done with regard to *saturation* and *value* sensitivity. However, as with *hue*, the frequency of confusions does not seem to change significantly over time (see figs. 3.14 and 3.16). The variation of *saturation* follows a similar distribution to fig. 3.11b (see fig. 3.14). The most commonly confused *saturation* values are spread almost uniformly between the range of presented *saturation* values and do not vary substantially throughout the day (see fig. 3.15). The most commonly confused *values*, fig. 3.17, span the range from 50–150 detailing the same trend as shown in fig. 3.11a. Note that discontinuities in the lines represent ranges where there are no colour mismatch observations.

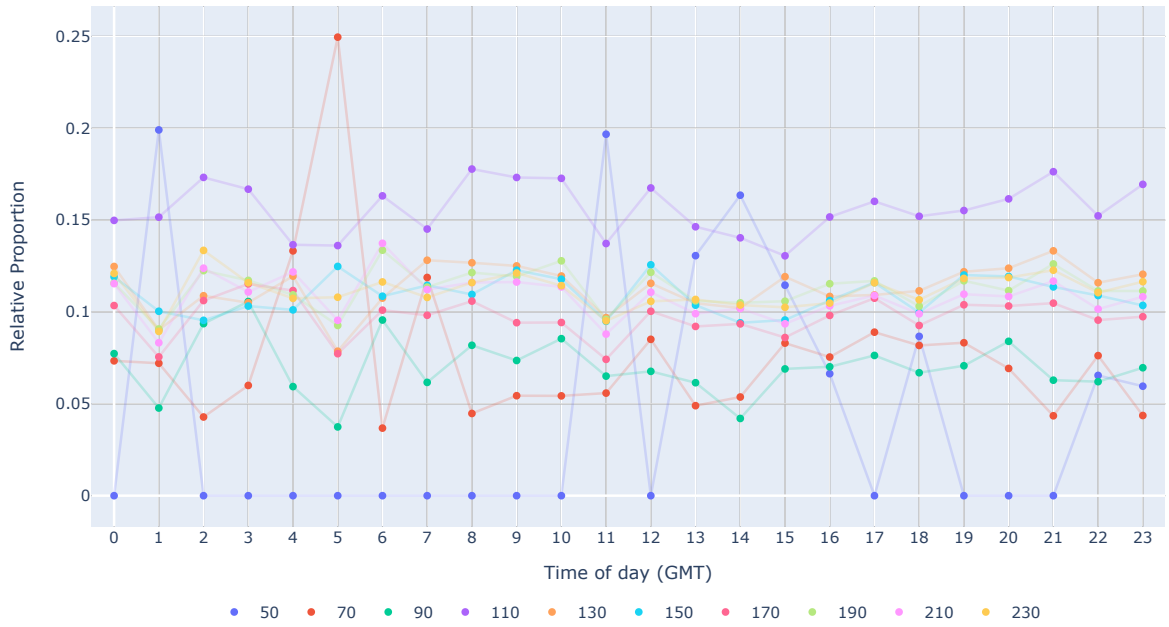


Fig. 3.14 The relative proportion of confused colours by *saturation* is not substantially affected by the time of day. Each line spans 20 units of *saturation*, except the line for '230' which represents the values between 230 and 255.

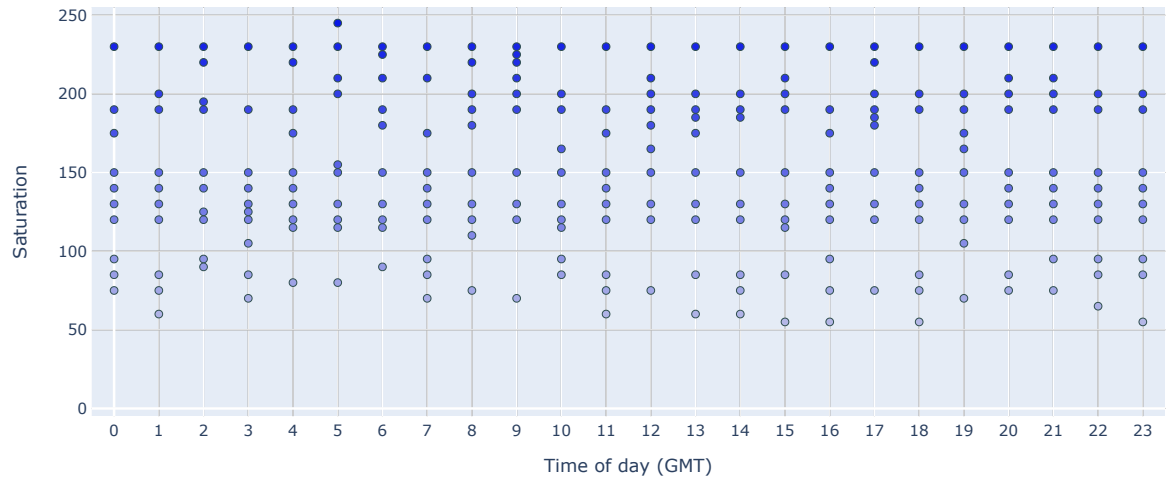


Fig. 3.15 Top 10 most confused colours by *saturation* does not vary substantially over the day (each point spans 5 units of *saturation*).

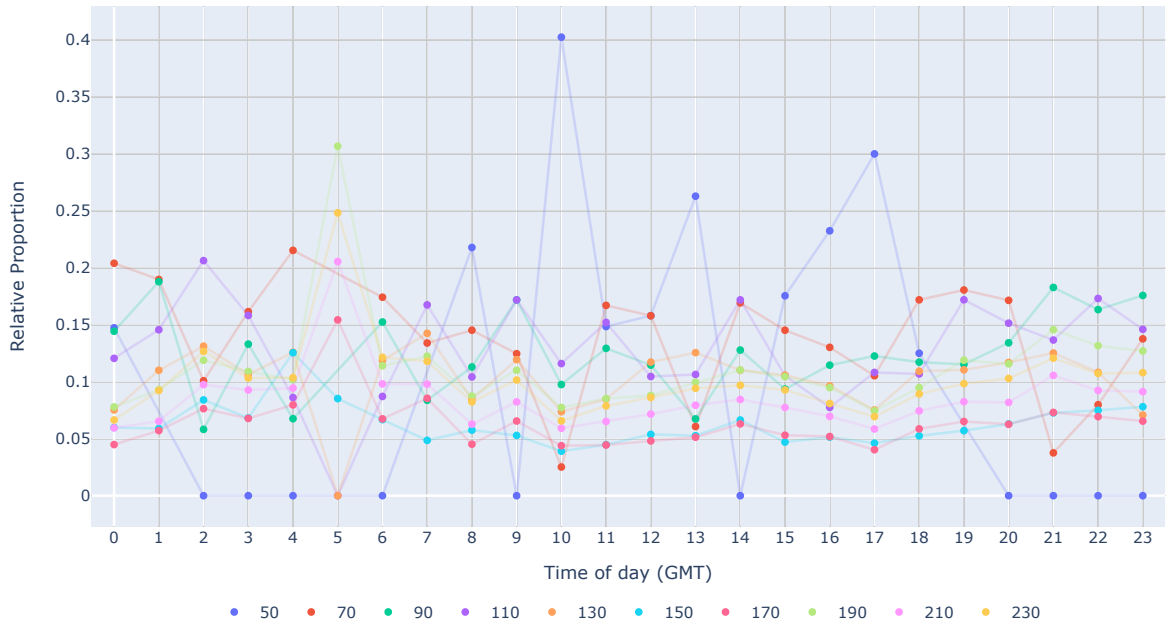


Fig. 3.16 The relative proportion of confused colours by *value* is not substantially affected by the time of day. Each line spans 20 units of *value*, except the line for '230' which represents the values between 230 and 255.

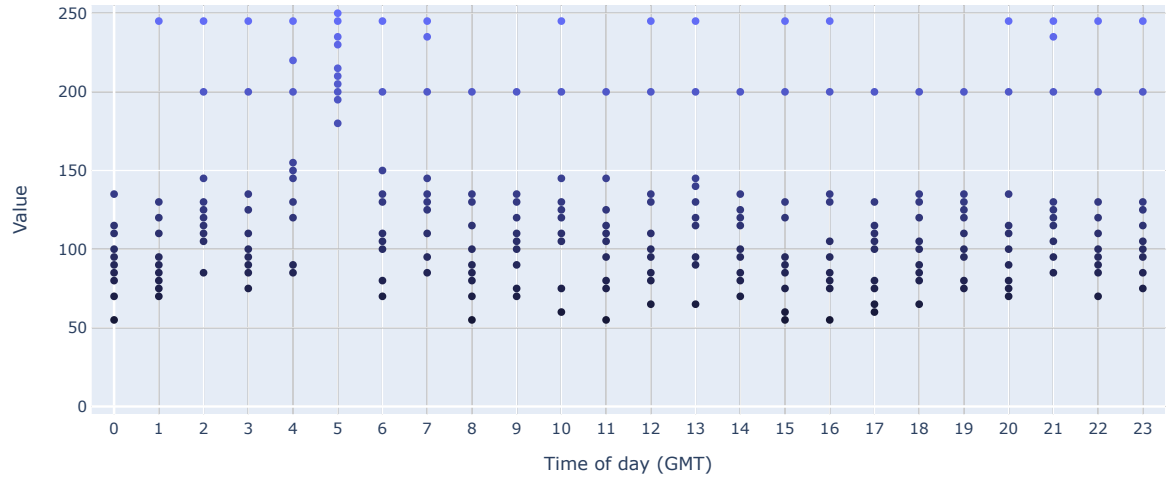


Fig. 3.17 Top 10 most confused colours by *value* does not vary substantially over the day (each point spans 5 units of *value*).

3.5 Individual CMF Inference from Specimen Data

As shown before, the *Specimen dataset* is a large corpus of colour matching data and this provides an insight into the colour acuity of 41,000 individual observers from around the world. In particular, given the ability to identify single player using their anonymised *player id*, it is possible to make a list of all the colours each player has incorrectly identified, as well as the corresponding correct colour. This is somewhat analogous to the traditional colour matching experiments used in colour science.

Each colour mismatch in the *Specimen dataset* shows locations where the observer's colour vision is unable to discern the difference between the target colour x and the chosen colour y. From a colour science perspective, this means that if one were to calculate the chromaticity coordinates of the colours x and y in a uniform colour space like *CIELAB* using the player's individual CMFs, the distance between them should be small. This fact should hold true to all the mismatched colours for a particular player, in the *Specimen dataset*. Thus, starting from a meaningful prior estimate of the individual's CMFs, like the CIE 2° std. observer, the process of finding the *individual CMFs* of a player can be posed as an optimisation problem where the objective is to find an estimate for the CMFs that minimises the distance between all the observed x and y pairs in the chosen colour space. This process is formalised in section 3.5.2.

3.5.1 Motivation

One of the key advantages of QD based optoelectronic devices is their narrow emission spectra. As narrow primaries allow displays to have access to a wide colour gamut, modern displays are increasingly using QDs and other optoelectronic materials that give narrow optical emissions. However, literature shows that, having primaries with narrow bandwidth spectra make displays more prone to *observer metameric breakdown* compared to those with wider bandwidth spectra [84]. Observer metameric breakdown, as outlined in section 2.7.8, is the situation where two different observers disagree on what 'colour' a particular spectrum is due to variations in their individual spectral sensitivities.

Most displays in the market today use primaries with sufficiently wide emissions that do not cause observer metameric breakdown to any practically problematic degree [84]. This has been supported, in part, by standardisation processes, like standard colour spaces and their associated primaries, which try to ensure colour constancy across different media [85, p. 67]. However, this will not be sufficient to prevent *observer metameric breakdown* when using narrow primaries like QDs as it is bounded by the variations in human vision. A possible solution for personal devices (e.g., mobile phones) would be to carry out a secondary display calibration to match the spectral sensitivities of the user's eyes. The method presented in the following sections is a potential avenue for measuring a user's spectral sensitivities unobtrusively (and remotely), to carry out secondary display calibration.

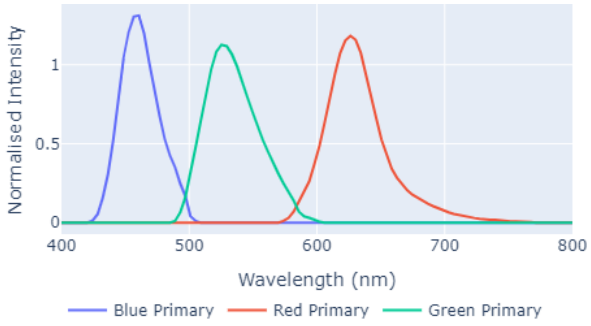


Fig. 3.18 The shapes of the 'red', 'green', and 'blue' primary emissions of the *iPhone X* display were extracted from its white point spectrum [19].

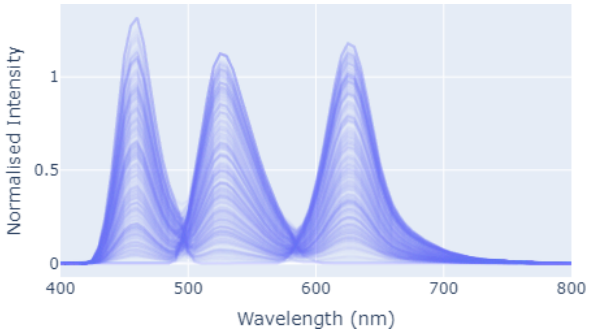


Fig. 3.19 Due to the narrow emission bandwidths there is limited information at some wavelengths (e.g., around 500 nm and 580 nm).

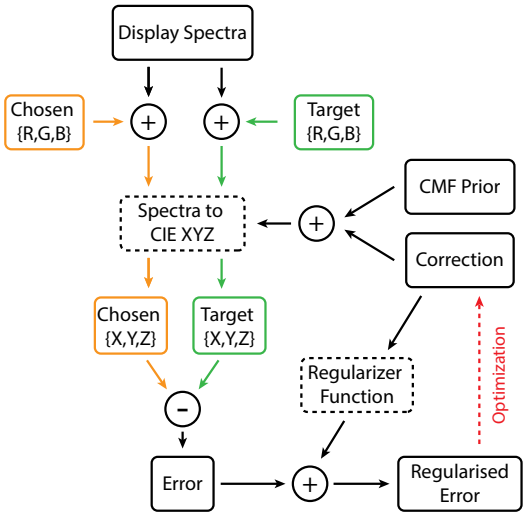


Fig. 3.20 Flow chart of the initial architecture used to learn the spectral correction term through a numerical optimisation process.

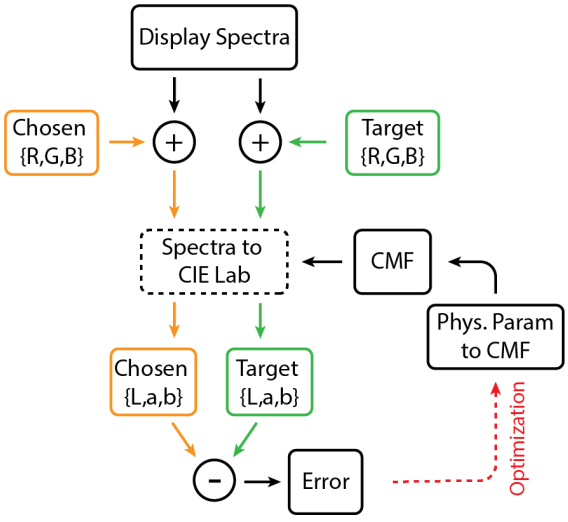


Fig. 3.21 Flow chart showing how the initial architecture is modified to incorporate a physiologically relevant model of the eye.

3.5.2 Problem Formulation

Let $\overline{C}_{rgb} = [C_r, C_g, C_b]$ be an incorrectly chosen colour and $\overline{T}_{rgb} = [T_r, T_g, T_b]$ be the correct target colour, both in the *sRGB* colour space with values normalized to the range $[0,1]$. Also, let $\overline{P} = [P_r, P_g, P_b]$ be the spectra for the chosen display's 'red', 'green', and 'blue' primaries. The spectrum of the chosen colour, S_c, and the target colour, S_t, can be calculated using eqs. (3.1) and (3.2) respectively.

$$S_c = \sum_{k \in \{r,g,b\}} P_k \cdot C_k \tag{3.1}$$

$$S_t = \sum_{k \in \{r,g,b\}} P_k \cdot T_k \tag{3.2}$$

Let $\overline{\phi}_{prior} = [\overline{x}_{prior}, \overline{y}_{prior}, \overline{z}_{prior}]$ be the CMFs prior used as the initial prediction of the user's individual CMFs. In this evaluation we use the CIE 2° and 10° std. observer. The CIE tristimulus values $[X,Y,Z]$ for the chosen spectrum S_c are calculated as,

$$X = \int_\lambda S_c \cdot \overline{x}_{prior}\, d\lambda \tag{3.3}$$

$$Y = \int_\lambda S_c \cdot \overline{y}_{prior}\, d\lambda \tag{3.4}$$

$$Z = \int_\lambda S_c \cdot \overline{z}_{prior}\, d\lambda \tag{3.5}$$

Similarly for the target spectrum S_t.

Let $\overline{\phi}_u$ be the individual CMFs for the user. Also, let $\overline{\delta_u}$ be a correction term with the same domain as $\overline{\phi}_{prior}$ such that $\overline{\phi}_u = \overline{\phi}_{prior} + \overline{\delta_u}$. The correction term represents the deviation of the user's vision from the CMFs prior. The user adjusted tristimulus values can be calculated using eqs. (3.3) to (3.5) with individual CMFs, $\overline{\phi}_u$, instead of the CMF priors, $\overline{\phi}_{prior}$. Let $\overline{\Phi}_c^{u_xyz} = [X_c^u, Y_c^u, Z_c^u]$ and $\overline{\Phi}_t^{u_xyz} = [X_t^u, Y_t^u, Z_t^u]$ be these user adjusted tristimulus values in the *CIE-XYZ* colour space, for the chosen spectrum and the target spectrum respectively. $\overline{\Phi}_c^{u_xyz}$ and $\overline{\Phi}_t^{u_xyz}$ is then transformed into the *CIELAB* colour space to get $\overline{\Phi}_c^{u_Lab} = [L_c^u, a_c^u, b_c^u]$ and $\overline{\Phi}_t^{u_Lab} = [L_t^u, a_t^u, b_t^u]$

In the *CIELAB* space, Euclidean distance corresponds to perceptual similarity. The colour difference is defined as $\Delta E_{ab}^* = \sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2}$. If $\Delta E_{ab}^* = 0$, the two colours are *perceptually identical*; $\overline{\Phi}_c^{u_Lab} = \overline{\Phi}_t^{u_Lab}$. Thus, the goal is to find the correction term $\overline{\delta_u}$ that minimises $\Delta E_{ab}^*(\overline{\Phi}_c^{u_Lab}, \overline{\Phi}_t^{u_Lab})$. This can be posed as the following minimisation problem:

$$\underset{\overline{\delta_u}}{\arg\min} \left[\Delta E_{ab}^* \left(\overline{\Phi}_c^{u_Lab}, \overline{\Phi}_t^{u_Lab} \right) \right] \tag{3.6}$$

To prevent the correction term from causing large deviations, we explored numerous regularisation functions. The regularised optimisation problem is given by the following equation:

$$\underset{\overline{\delta_u}}{\arg\min}\left[\Delta E^*_{ab}\left(\overline{\Phi}^{u_Lab}_c,\overline{\Phi}^{u_Lab}_t\right)+kR\left(\overline{\delta_u}\right)\right] \tag{3.7}$$

where R is a regularisation function, and k a positive real value defining the weight given to the regularisation function.

Equation (3.7) gives the optimisation process using a single colour selection event. However, the optimum value for $\overline{\delta_u}$ should be valid across all the colour selection events for a given user. Thus, for the optimisation, we consider all the user's incorrect selection events in a batch and calculate the mean value for $\Delta E^*_{ab}(\overline{\Phi}^{u_Lab}_c,\overline{\Phi}^{u_Lab}_t)+R(\overline{\delta_u})$ across this batch to optimise δ_u. By evaluating eq. (3.7), we can infer the individual CMFs from the *sRGB* values of the mismatched colour pairs.

3.5.3 Methodology

Device Selection

The process outlined in the previous section depends on knowledge of the shapes of the display's primary emissions. Therefore, it is primarily applicable for devices with emissive pixel displays like OLEDs and QD-LEDs. As the *Specimen* game stopped collecting data in 2018, the only iOS device with such a display at the time was the *iPhone X*. We used the measurements carried out by Raymond Soneira of *DisplayMate Technologies* on the white-point of the *iPhone X* display as the starting point to manually isolate the shapes of the three primary emissions from the combined white spectrum and used this as the display model (see fig. 3.18) [19].

Player Selection

In the *Specimen dataset*, there are 141 players using *iPhone X*s, totalling 21,250 colour matches. In the analysis that follow, we will primarily be using the colour mismatches of the player with the largest number of incorrect matches, to evaluate our method. This player (hitherto referred to as *player A*) had 344 incorrect matches from 2,042 total matches. We chose this player as the large number of mismatches maximises the amount of information the learning algorithms have to infer the player's spectral sensitivities from. This represents a best-case scenario, given the information available in the *Specimen* dataset.

Optimisation

The *Specimen dataset* provides the *sRGB* and *CIELAB* coordinates for the colours presented in the game. First, all the colour confusions for a given player was aggregated and made into a list of 3-tuple pairs, where each pair corresponded to the coordinates of the target and the chosen colours, in the *sRGB* colour space.

As outlined in the previous section, the aim was to find a correction term on the CMF prior that explains the observed colour confusions. The optimisation was done using the *Adam* optimiser (within *Tensorflow*) [86]. The Adam optimiser is a stochastic gradient descent algorithm for finding solutions to numerical optimisation problems. During each iteration, the optimiser aims to change the correction term such that the mean difference between the tristimulus values of the chosen and target colours, is minimised.

In this initial evaluation four CMF priors were evaluated; uniform prior, Gaussian triple prior, CIE 2° std. observer, and CIE 10° std. observer [87]. For the uniform prior, the peak value was chosen to be 1, while for the Gaussian mixture the peak wavelengths, intensities and the bandwidths were matched approximately to those of the 2° std. observer CMFs (see fig. 3.28).

For each mismatched, chosen and target colour pair, emission spectra were calculated by multiplying the normalised R, G, and B values with the corresponding primary emission of the display. The pairs of emission spectra were then converted into their corresponding *CIELAB* tristimulus values using the calculated CMFs.

Note that the *Specimen dataset* does not give any information about the absolute intensities of the display primaries. Thus, in this preliminary investigation, we adjusted the relative spectral intensities to match the display luminance at 50% display brightness. This was measured to be $700\,cd/m^2$; a radiance of $1.025\,w/m^2$. Let $L^w_{e\Omega\lambda}$ and $\ell^w_{e\Omega\lambda}$ be the peak and relative, white point spectral radiance of the display. $L^w_{e\Omega\lambda}$ can be calculated from $\ell^w_{e\Omega\lambda}$ using a scaling factor a as defined in eq. (3.8).

$$L^w_{e\Omega\lambda} = a \cdot \ell^w_{e\Omega\lambda} \tag{3.8}$$

The scaling factor, a, can be calculated using eq. (3.9).

$$a = \frac{L_v}{K_w \int_\lambda \ell^w_{e\Omega\lambda} \cdot V(\lambda)\, d\lambda} \tag{3.9}$$

where L_v is the display luminance ($700cd/m^2$), K_w is the photopic spectral luminous efficacy ($683lm/W$), and $V(\lambda)$ is the CIE-2006 *physiologically-relevant* 2° LEFs [51].

Regularisation

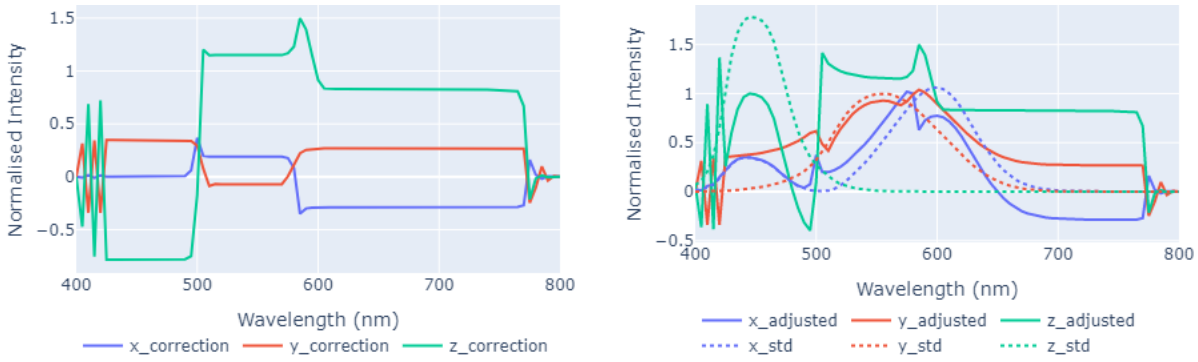
Given that the proposed method relies purely on numerical optimisation, the definition of the cost function (eq. (3.7)) plays a key role in the accuracy of the result. In particular, the cost function serves as the conduit for embedding physical properties of the eye into the optimisation. We explored simple regularisation functions (R in eq. (3.7)) as a first order approximation, to dictate a preference for the type of CMFs the optimiser learns. We evaluated the following regularisers;

1. Maximum of absolute values : $max(|\overline{\delta_u}|)$ - This penalises large positive or negative deviations with the dimension with the maximum deviation penalised the most.

2. Mean of Absolute values: $\Sigma(|\overline{\delta_u}|)/n$ - Similar to above, but the deviations are averaged, which smooths out the effects of the deviations across all the dimensions. Likely to lead to higher minimum cost than $max(|\overline{\delta_u}|)$.
3. Sum of Absolute values [88]: $\Sigma(|\overline{\delta_u}|)$ - This penalises the total absolute error thus giving equal weight to all dimensions.
4. Root Mean Squares : $\Sigma(\|\overline{\delta_u}\|_2)/n$ - Similar to mean of absolute values, but with errors weighted quadratically (i.e., errors greater than one get increased weight than those less than one).
5. Mean of Squares : $\Sigma(\|\overline{\delta_u}\|_2^2)/n$ - Similarly to Root Mean Squares, but the scale of the aggregated error is larger (squared) than the scale of the individual errors.
6. Sum of Squares [89]: $\Sigma(\|\overline{\delta_u}\|_2^2)$ - Similar to Mean of Squares but with errors weighted uniformly across all dimensions.

Note that, we are using *CIE-XYZ* colour matching functions in this initial investigation, instead of the cone fundamentals because it simplifies the optimisation process and because the cone fundamentals can be extracted from the CMFs using a known linear transformation if required [90]. In section 3.6, we will explicitly incorporate properties of the cone fundamentals into the optimisation process to achieve a stronger, physiologically relevant prior. This initial evaluation also calculates distance in the *CIE-XYZ* space instead of the *CIELAB* space as outlined in the ideal case (section 3.5.2) due to implementation issues (see fig. 3.20). We discuss this in more detail in section 3.6.2.

3.5.4 Sensitivity to Regularisation



(a) Correction term.

(b) Learned CMFs

Fig. 3.22 Without regularization the optimization overfitted and resulted in calculated values that are physically impossible.

Figure 3.22 shows that, without regularisation, the numerical optimisation leads to physically implausible results. Figures 3.23 to 3.26 show the qualitative effects of the different regularisers on the learned CMFs, when using the CIE 2° std. observer.

Due to *broken gradients* in the optimiser, we were unable to optimise using the *root mean squares* regulariser. The first three regularisers, $max(|\overline{\delta_u}|)$, $\Sigma(|\overline{\delta_u}|)/n$, and $\Sigma(|\overline{\delta_u}|)$, resulted in CMFs with negative

values, which are implausible; In the *CIE-XYZ* colour space the CMFs are defined to be strictly positive. In comparison, both $\Sigma(\|\overline{\delta_u}\|_2^2)/n$ and $\Sigma(\|\overline{\delta_u}\|_2^2)$ regularisers resulted in smooth CMFs with the former achieving smaller final cost (i.e., better match with the data).

Overall, the sum of squares regularisation ($\Sigma(\|\overline{\delta_u}\|_2^2)$) gave the smoothest and most qualitatively realistic looking CMFs (see fig. 3.27).

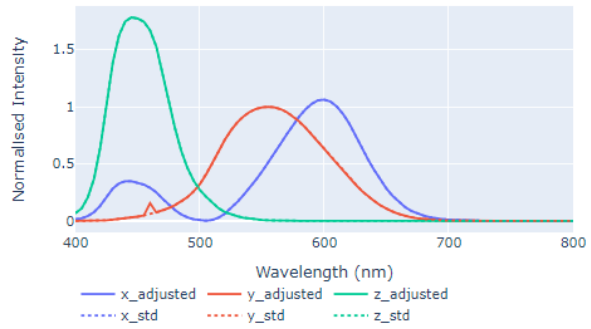


Fig. 3.23 *Sum of Absolute Values* regularization led to localised peaky artefacts.

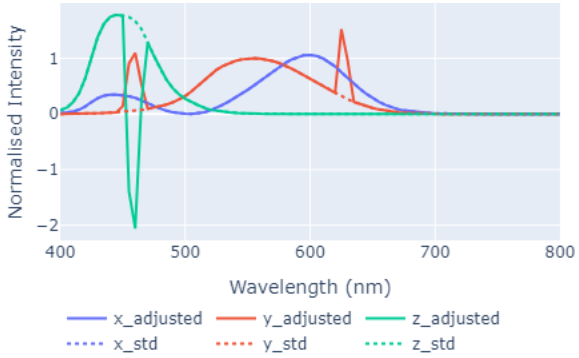


Fig. 3.24 *Mean absolute* regularization caused *peaky* artefacts.

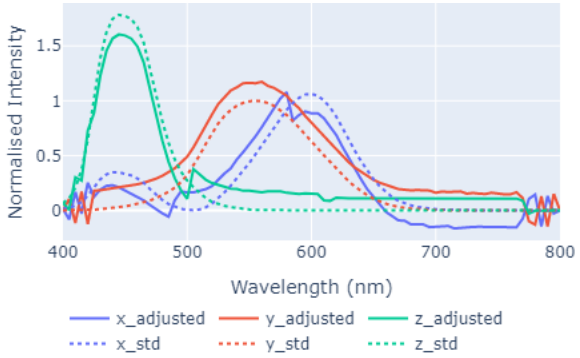


Fig. 3.25 *Max absolute* regularization led to slightly smoother curves than non-regularized optimization.

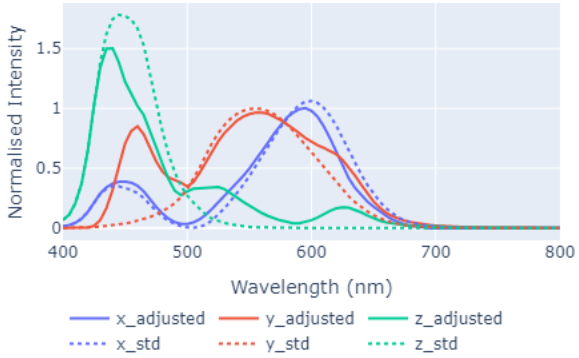
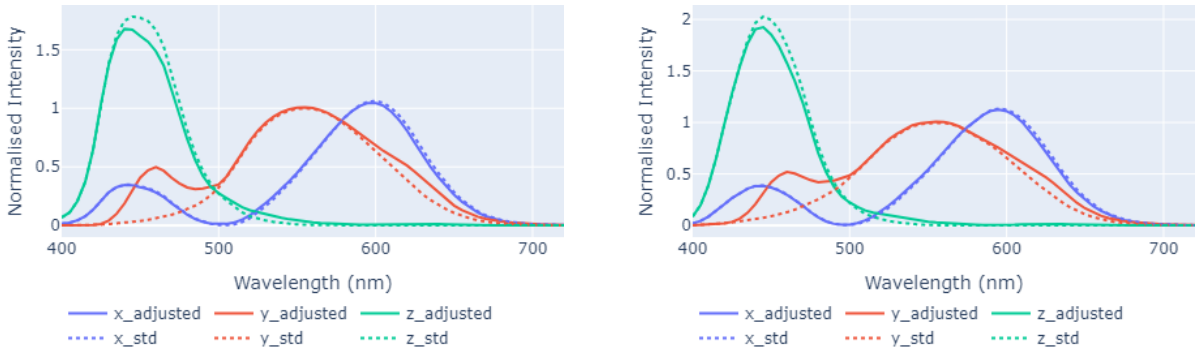


Fig. 3.26 *Mean Squares* regulariser resulted in a smooth set of CMFs but showed more fluctuation around the CMF prior compared to the *sum of squares* regulariser.

3.5.5 Effect of the CMF Prior

The sections below summarises the effects of the CMF prior on the results obtained. As outlined in section 3.5.3, we explored the following 4 types of priors; CIE 2° and 10° std. observers, a uniform prior, and a Gaussian triple prior.



(a) 2° CMF

(b) 10° CMF

Fig. 3.27 Optimizing with *sum of squares* regularizer show that both 2° and 10° standard observer CMF were adjusted to explain the observed colour confusions better.

CIE 2° and 10° Standard Observer Priors

Figures 3.27a and 3.27b show the results for *player A* for both 2° and 10° std. observer priors with sum of squares regularisation using a weight of 0.05 on the regularizer term. In both cases, the resulting CMFs were close to the priors with the key deviation being the secondary peak on the Y component of the CMF.

Uniform Priors

For the uniform prior, all three CMFs start off with a fixed value of 1 across the range of wavelengths. It was hoped that during the optimisation process these will deviate from their starting positions and settle to values that better explain the observed colour confusions. However, as shown in fig. 3.20, the process of calculating the tristimulus values for both chosen and target colours uses the same CMFs. It was soon found that this causes the gradients of the cost function to be quite small. Consequently, the optimisation resulted in CMFs which were also uniform. We tried adjusting the values of the three components to match the peak intensities of the 2° standard observer CMFs which produced a similarly unrealistic result. In hindsight, this result is not surprising but was worth presenting to provide a more complete picture.

Gaussian Priors

For the Gaussian prior, we approximated the 2° standard observer CMFs with three Gaussians (see fig. 3.28). Figure 3.29 shows the solution obtained with the sum of squares regularizer with a weight of 0.05 (identical setup to figs. 3.27a and 3.27b). The results show a striking similarity to the results obtained with the CMF priors with the same regularisation setting.

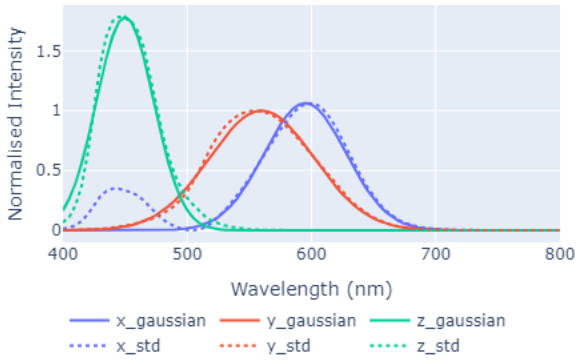


Fig. 3.28 The 2° standard observer CMFs were approximated with three Gaussians.

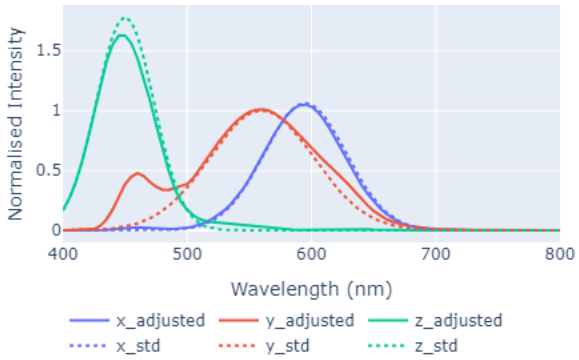


Fig. 3.29 The CMFs learnt with the gaussian approximation were similar to those obtained when using the 2° standard observer prior.

3.6 CMF Inference with a Psychologically-Relevant Observer Prior

The process summarised in the previous section has no physiological basis. Consequently, there is no guarantee that the results obtained will have any physical meaning. However, what it does show is that, an optimisation process can be used to infer CMF like shapes from the colour mismatches in the *Specimen dataset*.

Colour models like CIE-XYZ and CIE-LAB are based on in-person colour matching experiments and are defined using empirical CMFs. As a result, any optimisation process used to find optimised CMFs will have to operate on the wavelength domain (i.e., explicitly find what the CMFs have to be at each wavelength). This makes it difficult to constrain the problem and would lead to unrealistic results as shown in section 3.5.

Although there exists other vision models that aim to model the HVS more accurately these models do not fit the problem formulation. Many of them aim to model a specific component of the HVS (e.g., contrast sensitivity [91, 92], masking effects [93, 94], spatial frequency response [95, 96], etc.) and are particularly aimed at describing luminance related effects. The problem at hand requires a physiologically relevant model of human *colour* vision.

In 2016, Asano et al. presented a model that is an extension of the CIE-2006 Physiological Observer model (CIEPO06) [22]. This model, defined in eq. (3.10), uses the following 10 parameters to specify an individual's LMS cone spectral sensitivities.

$$f(a, v, d_{lens}, d_{mac}, d_L, d_M, d_S, s_L, s_M, s_S) \tag{3.10}$$

- a : Age of the observer (as defined in CIEPO06) [years]
- v : Visual angle (similar to field size in CIEPO06) [degrees]
- d_{lens} : Deviation from average lens pigment density [%]
- d_{mac} : Deviation from average peak optical density of macula pigment [%]

- d_L, d_M, d_S : Deviations from the average peak optical densities of the L-, M- and S-cone photopigments [%]
- s_L, s_M, s_S : Deviations from the average λ_{max} shifts of the L-, M- and S-cone photopigments [nm]

In this model, the age, a ,and visual angle, v, are defined the same way as the CIEPO06 model. Internally, the function defined in eq. (3.10) uses the eqs. (3.11) to (3.22), to calculate the LMS cone spectral sensitivities from the parameters described above. By using this model, we are not only reducing the dimensionality of the optimisation problem, but are also explicitly incorporating physiological properties of the human vision system into the optimisation process.

Equation (3.11) models the effect of the variability in the lens pigment, where $D_{ocul,avg}(\lambda)$ is the average spectral optical density for an observer of age a.

$$D_{ocul}(\lambda) = D_{ocul,avg}(\lambda, a)\left[1 + \frac{d_{lens}}{100}\right]. \tag{3.11}$$

The average spectral optical density, $D_{ocul,avg}(\lambda)$, is defined by Equation (3.12), based on work by Pokorny et al., where $D_{ocul,1}$ and $D_{ocul,2}$ are the age dependent and age independent components of average spectral optical density [97].

$$D_{ocul,avg}(\lambda, a) = D_{ocul,1}(\lambda)[1 + 0.002(a - 32)] + D_{ocul,2}(\lambda) \tag{3.12}$$

Equation (3.13) models the variability of the peak optical density in the macula pigment, where D_{mac}^{max} and $D_{mac}^{relative}(\lambda)$ are the peak and relative spectral optical densities of the macula pigment as defined in eq. (3.14) and CIEPO2006 respectively [98, 51].

$$D_{mac}(\lambda) = D_{mac}^{max} D_{mac}^{relative}(\lambda) \tag{3.13}$$

$$D_{mac}^{max} = 0.485e^{-v/6.132}\left[1 + \frac{d_{mac}}{100}\right] \tag{3.14}$$

Equation (3.15) models the shift in the location of peak spectral sensitivity of the photopigment of given type ($j \in \{L, M, S\}$), and $A_j(\lambda - s_j)$ is the average low optical density spectral absorbance for each cone type, as defined in CIEPO2006 [51].

$$A_{shift,j}(\lambda) = A_j(\lambda - s_j) \tag{3.15}$$

The eqs. (3.16) to (3.19) model the effects of variations in the peak optical density of the photopigments, where $D_{photopig,j}^{max}$ is peak optical density of a given cone type, j, as defined in eqs. (3.17) to (3.19) [99], and v is the visual angle in degrees.

$$\alpha_j(\lambda) = 1 - 10^{-D_{photopig,j}^{max} A_{shift,j}(\lambda)} \tag{3.16}$$

$$D^{max}_{photopig,L} = 0.38 + 0.54e^{-v/1.333}\left[1 + \frac{d_L}{100}\right] \tag{3.17}$$

$$D^{max}_{photopig,M} = 0.38 + 0.54e^{-v/1.333}\left[1 + \frac{d_M}{100}\right] \tag{3.18}$$

$$D^{max}_{photopig,S} = 0.30 + 0.48e^{-v/1.333}\left[1 + \frac{d_S}{100}\right] \tag{3.19}$$

And finally, eqs. (3.20) to (3.22) defines the L, M, S cone fundamentals as a function of the models defined by the equations above.

$$l(\lambda) = \lambda\,\alpha_l(\lambda)10^{-Dmac(\Lambda)-D_{ocul}(\lambda)} \tag{3.20}$$

$$m(\lambda) = \lambda\,\alpha_m(\lambda)10^{-Dmac(\Lambda)-D_{ocul}(\lambda)} \tag{3.21}$$

$$s(\lambda) = \lambda\,\alpha_s(\lambda)10^{-Dmac(\Lambda)-D_{ocul}(\lambda)} \tag{3.22}$$

We can calculate the 10° individual observer CMFs, from eqs. (3.20) to (3.22), using the linear transformation in eqs. (3.23) to (3.25) [22, 100].

$$\bar{x}_{10^\circ}(\lambda) = 1.93986443 \cdot l(\lambda) - 1.34664359 \cdot m(\lambda) + 0.43044935 \cdot s(\lambda) \tag{3.23}$$

$$\bar{y}_{10^\circ}(\lambda) = 0.69283932 \cdot l(\lambda) + 0.34967567 \cdot m(\lambda) \tag{3.24}$$

$$\bar{z}_{10^\circ}(\lambda) = 2.14687945 \cdot s(\lambda) \tag{3.25}$$

3.6.1 Methodology

The choice of players and device was done in the same way as before (see section 3.5.3). However, in addition to *player A*, we also evaluated the next 4 players by the number of mismatches to verify our method. Although the presented results will be for *player A*, in the discussion these auxiliary results will also be used.

In the original formulation, the optimiser was trying to find the CMFs that minimised the difference between the chosen colours and the target colours for the given player, in the *CIELAB* colour space. In the framework with the physiologically relevant model, the optimiser is trying to find estimates for the physiological parameters outlined in section 3.6 (except a and v) which are then used to calculate the CMFs. These CMFs are then used as before to find the *CIELAB* coordinates of the colour pairs. The optimiser uses the difference between these coordinates as the error signal in the optimisation. Figure 3.21 outlines this process graphically.

Initial Conditions

As there is no explicit indication of the player's age, the age parameter (a) was set to be 25. Similarly, the field size (v) was chosen to be 10°. These two values are chosen to be fixed values as they can be easily measured, which restricts the optimiser to only needing to find the other eight parameters. The initial values for those parameters were all set to 0.1. A value of 0 indicates that the physiological

parameters show no deviation from the standard observer. Setting the value to 0.1 instead of 0 was observed to help with the numerical stability of the optimisation process.

Optimiser Bounds

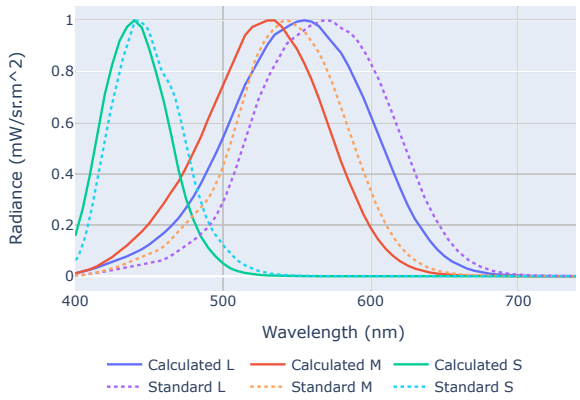
In their work, Asano et al. calculated the standard deviation of each of the eight parameters by analysing reports of physiological measurement data from the literature [22] (see table 3.4). These values were used as a baseline and we set the optimiser bounds for the allowed range of the parameter values to be 2 standard deviations on either side of zero.

d_{lens}	d_{mac}	d_L	d_M	d_S	s_L	s_M	s_S
19.1 %	37.2 %	17.9 %	17.9 %	14.7 %	4.0 nm	3.0 nm	2.5 nm

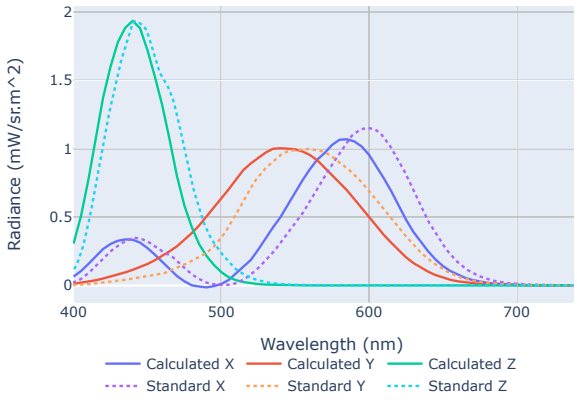
Table 3.4 Standard Deviation of the physiological parameters from Asano et al. [22].

In contrast to before, to mitigate gradient issues, the optimisation was carried out using the *SHGO* algorithm with 50 point *Sobol* sampling and 10 iterations for generating the simplicial complex [101, 102]. We chose a random subset of 100 mismatch pairs from the set of all player colour mismatches for the optimisation. We validated this sub-sampling process by carrying out the optimisation on five random subsets of the player's colour mismatches and comparing the calculated physiological parameters for parity.

3.6.2 Results



(a) The learned cone sensitivities and the standard cone sensitivities.



(b) The CMFs calculated from the learned cone sensitivities and the 10° standard observer CMFs.

Fig. 3.30 The optimisation process is able to learn spectral sensitivities that better explain the observed colour confusions.

Figure 3.30 shows the learned cone sensitivities and the corresponding CMFs for *player A*, and the learned values for the physiological parameters are given in the first row of table 3.5. Across the 5

optimisation restarts, the results were observed to be consistent. In comparison to the previous method (see section 3.5), owing to the embedded physiological relationships, the learned spectral sensitivities (and thus the CMFs) are much more realistic. This illustrates that the proposed method can potentially be used to extract hypothetical individual CMFs from online colour matching data.

User	d_{lens} (%)	d_{mac} (%)	d_L (%)	d_M (%)	d_S (%)	s_L (nm)	s_M (nm)	s_S (nm)
A	-38.20	-74.40	-35.80	-35.80	-29.40	-8.00	-6.00	-4.99
B	-19.55	-38.07	9.65	8.25	1.49	-7.59	3.96	-0.21
C	-28.05	-68.59	-7.27	0.56	-8.73	-6.88	2.72	-2.89
D	-28.05	-68.59	-7.27	0.56	-8.73	-6.88	2.72	-2.89
E	19.1	37.2	17.9	17.9	14.7	4.0	3.0	2.5

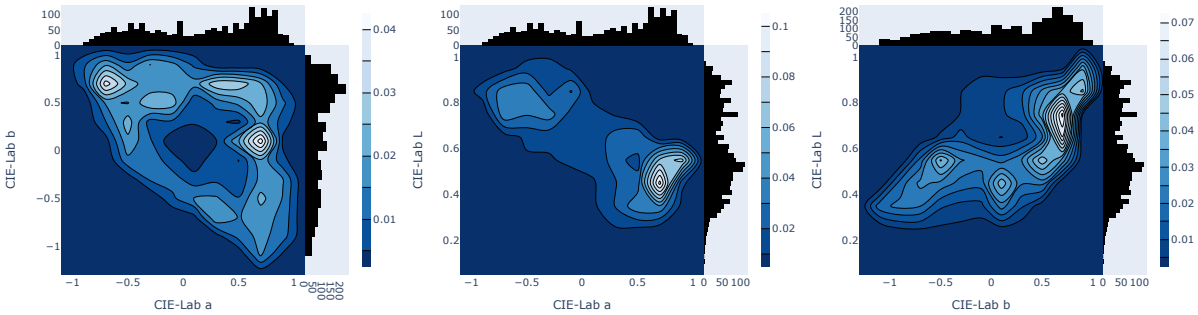
Table 3.5 Individual player's optical properties learned by the optimiser.

In the preliminary examination, it was observed that when the number of colour mismatches is large (>200), it led to numerical instabilities and to global optima that are independent of the colour mismatches. The 100-point resampling and the validation using subsets was chosen because of this. Although the results were better with this method, the optimiser sometimes still learned the same parameters between some players (see players C and D in table 3.5). This could be caused by mismatched colour pairs being very similar between the players or by the display model being too approximate, leading to erroneous minima. The mismatch distributions for players C and D (see figs. 3.31a and 3.31b) show substantial overlap between the locations of the most common colour mismatches, giving some credence to the first hypothesis. Further study beyond these preliminary results would be required for validating the proposed method.

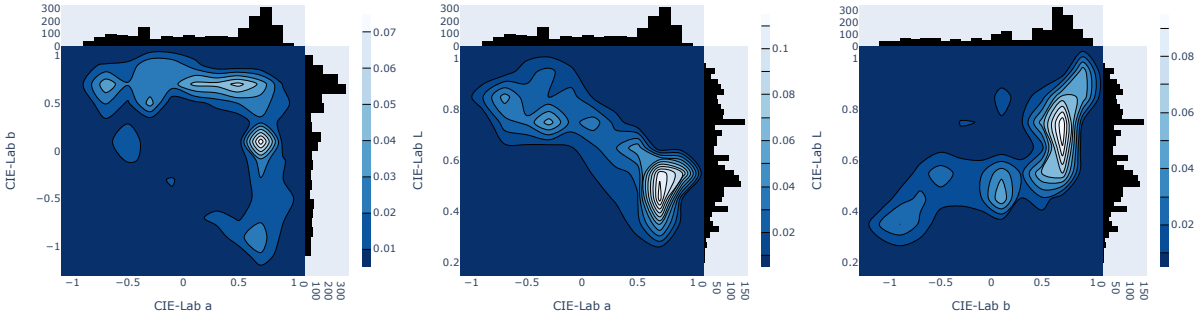
In addition to *SHGO*, we evaluated *Nelder-Mead*, *dual-annealing*, and *L-BFGS-B* methods to do the optimisation [103–106]. Due to the lack of explicit bounds, *Nelder-Mead* resulted in extreme and unrealistic results. *Dual-annealing* gave similar results to *SHGO*, but was slower to converge. *L-BFGS-B* failed to optimise at all due to a zero-valued Jacobian matrix error. This is similar to the issue encountered when trying to optimise in the *CIELAB* colour space in section 3.5.3. In that case, the *Adam* optimiser was unable to optimise due to 'broken gradients'. We believe that the conditional statement in the *CIE-XYZ* to *CIELAB* conversion might be preventing the optimiser from calculating cost function gradients (see eq. (2.11)).

3.7 Related work

To our knowledge this is the first work that uses measurements from mobile display systems to infer an individual observers' CMFs. It is also the first bit of work that uses such a large corpus of colour matching data. However, there are few articles in the literature, that use displays to conduct colour matching experiments. Sarkar et al. [107, 84] and Oicherman et at. [108, 109] used CRT and LCD/LED displays to evaluate metameric breakdown between the different display types and examine inter-observer variabilities in colour perception. Similar work was done by Fang and Kim [110], and Xie et al. [111]



(a) Density plots of player C's mismatches



(b) Density plots of player D's mismatches

Fig. 3.31 Density plots of the colour confusions by players C and D show significant overlap.

using LCD and OLED displays to model observer metamerism and to characterise its effects on display calibration. Pardo et al. [112] presents a computer-based method that uses a CRT display to simulate a Pickford-Nicolson anomaloscope to detect red-green colour vision deficiencies in observers. In their work they use a spectrometer to measure and model the primary emissions of the display. In all the above work, the results were obtained using in-person user studies and the authors had the opportunity to calibrate and model the displays accurately before the experiments.

The closest work to the work presented in this chapter would be the work done by Park et al. in 2016 [113]. In their work, they use a genetic algorithm to estimate CMFs of individual observers to calibrate a tiled LCD system. Their work represents the Y and Z CMFs using a single Gaussian distribution, and the X with a mixture of two Gaussian distributions. This is in contrast to our work which uses an extension of the physiologically accurate model presented by Asano et al. [22] to generate the CMFs from the cone fundamentals.

3.8 Conclusions and Future Work

This chapter presented work done on extracting individual observer CMFs from a large corpus of colour matching data from a mobile game about colour. It explored how to structure the optimisation process to embed physiological information into it such that the resultant CMFs have a physiological basis. The chapter also presented some statistics about the *Specimen dataset*, exploring how the players' colour

acuity is affected by different chromatic properties of the presented colour samples and is affected by the time of day.

The preliminary results in section 3.6 showed that potential values for a player's ocular properties can be calculated from the online colour matching data. However, as outlined in that section, the results of the optimisation process show numerical instabilities, and thus a more in-depth study into the implementation of the vision model will be needed to iron out the results. It would be useful if the models could be re-written in a way that the preserves the gradient information between the different mathematical operations, such that gradient-based methods can be used to learn the optimal parameters.

The method used to create the RGB-to-spectra model of the display introduces a substantial amount of epistemic uncertainty into the CMF inference process. In this work we used the optical emission at the white point of the display to extract the shape of the three primaries and made assumptions about the luminance of the display. We also assumed that their peak intensities are linearly mapped to the R,G,B values (i,e., R=255 implies that the red primary intensity was at its maximum). In reality none of these assumptions might hold. Furthermore, variations in the calibration settings of the individual displays would lead to further deviations between the behaviour of the model used in this work and the actual display dynamics. Ultimately, because the validity of this method hinges on how accurately the display model represents the actual display, priority should be given to measuring the performance of the real *iPhone X* device. A process similar to the one outlined in appendix A could be used for this.

Regarding the future use of this method, there are two main avenues: use of online colour matching for display calibration and extracting player CMFs from the *Specimen dataset*. For the first use case, the proposed next step would be to extract a better display model and carry out end-to-end testing of the method with a user study. An avenue for optimisation would be to figure out the optimal set of colour samples required to maximise the information gain about the player CMFs with as few colour mismatches as possible. As outlined before, the player age and the field size are hardcoded into the optimisation process as there is no information provided in the *Specimen dataset* about these parameters. However, during a user study, it would be possible to measure these values quite easily, and allow more accurate CMFs to be extracted.

For the second use case, the key bit of information that the *Specimen dataset* provides about the display conditions is in the *CIELAB* coordinates. It might be possible, through another optimisation process to find the mapping between *sRGB* values and the display primary radiance that gives the observed *CIELAB* coordinates. This would allow a more accurate value for the display spectrum to be obtained, leading to better optimisation results. By coordinating with the developers of Specimen, it might be possible to get access to the method used to calculate the *CIELAB* coordinates and be able to calculate the exact display spectrum for each colour mismatch. Furthermore, the anonymisation of the dataset has removed the player age. With permission from the Specimen developers and the players, it might also be possible to get, and use, the exact ages of the players to achieve more accurate results.

'Of course it is happening inside your head, Harry, but why on earth should that mean that it is not real?'.

— Albus Dumbledore.
Harry Potter and the Deathly Hallows by J.K. Rowling [116, p. 723].

Chapter 4

Metamerism and Metameric Breakdown

4.1 Introduction

As outlined in chapter 2, human vision acts like a many-to-one function, where quantifiably different spectra can be perceived as the same colour. This effect is referred to as metamerism, and the different spectra that are perceptually similar are called metameres. In this chapter we explore how to leverage metamerism to achieve energy savings in emissive pixel displays (e.g., OLED and QD-LED displays).

Our analysis begins with a simple proposition: spectra composed of different frequency components take different amounts of energy to generate. For *spectral generators* (e.g., lights and displays) based on solid-state electronics (LEDs, OLEDs, and QD-LEDs) there are two reasons for this. Fundamentally, the amount of energy in a photon, E, is proportional to its frequency, ν, and $E = h\nu$ [37, p. 59]. This means, for a spectrum, composed of photons of different frequencies, the total, minimum, energy required to generate that spectrum, is the sum of the energies of the photons. In most optoelectronic devices however, the actual amount of energy required to generate a photon of a given wavelength is higher due to differences in material chemistry, extraction losses, and other frequency-dependent in-efficiencies [114, p. 283][115]. This is further exacerbated by the non-uniform spectral sensitivity of human vision (see fig. 2.5). This leads to 'blue' and 'red' colours requiring a higher radiance to generate the same luminance as 'green' colours. As the radiance is correlated with current (and energy, for a given frequency/voltage), this means that 'blue' and 'red' colours will require more energy to generate than 'green' colours (i.e., lower lumens per watt).

As outlined in section 2.7.8, it is also true that different spectra can be perceived as the same 'colour'. Considering the previous exposition about the variation of energy as a function of wavelength, this chapter aims to explore the following questions;

- It is possible to leverage metamerism to increase the energy efficiency of spectral generators?
- If so, how much energy saving can be achieved hypothetically by leveraging metamerism?
- What hardware considerations are required to achieve this?

4.1.1 Contributions

In this chapter we present the following contributions:

- Show that metamerism can be leveraged to lead up to an order of magnitude less power usage.
- Exploration of the properties of the landscape of metamers using a brute force search.
- Present a mathematical framework for describing multi-primary displays (MPDs).
- Derive a metamerism index and an energy-saving index that captures the metameric freedom in a display.
- Evaluate how to use a MPD architecture to minimise metameric breakdown when using narrow primaries.

4.2 Metamere Search on Hypothetical Three-primary Displays

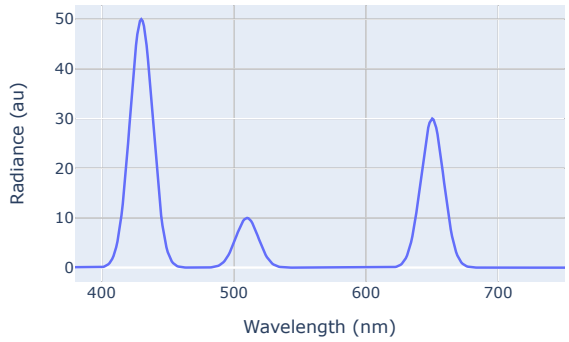


Fig. 4.1 The metamere search experiment was conducted on the set of all possible three-primary spectral shapes with fixed FWHM and Gaussian shaped primary emissions.

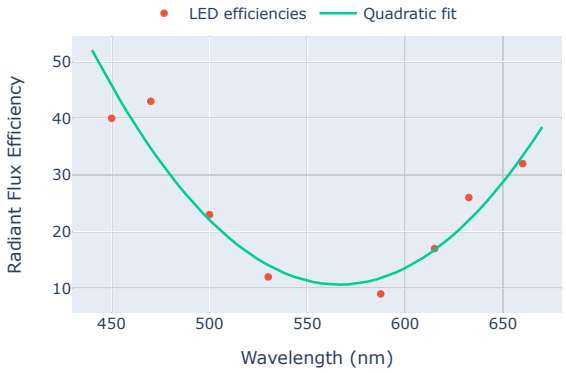


Fig. 4.2 The simple LED energy approximation follows a quadratic shape (based on data from [20]).

The investigation begins by looking at the ideal case; a three-primary display with the ability to control both the emission intensity and wavelength of the individual primaries. With such a display, a wide variety of spectral shapes can be generated. Given that the different wavelengths consume different amounts of energy to generate, each spectrum will have an associated energy cost. We can use this idealised setup to explore the following questions:

1. How many different ways are there of generating a particular colour sensation? (i.e., Given the chosen quantisation, how many metameres are there?)
2. For each metamere, what is the estimated energy consumption?

Based on these results, the potential energy saving that can be achieved can be calculated and the space of metameres can be explored for patterns and insights to devise further optimisations.

4.2.1 Experiment Setup

Consider a hypothetical three-primary display with Gaussian-shaped primary emission spectra, whose centre frequencies and the intensities can be altered. The state of such a display can be defined by six parameters; for each of the three primaries, two values denoting its intensity and the emission frequency. Consider a target colour sensation that we are trying to generate, defined as chromaticity coordinates in the CIE xyY coordinate space. The aim is to control the aforementioned six parameters till the *colour* of the spectrum matches the target colour (as calculated using the CIE 2° std. observer CMFs). To avoid redundancy, the visible spectrum is split into three non-overlapping, 'short', 'medium', and 'long' wavelength regions and each primary is assigned to one of them ('short' = [400, 516] nm, 'medium' = [516, 632] nm, and 'long' = [632, 748] nm).

Using this setup, spectra were sequentially generated by incrementing the wavelength of the 'short', 'medium', and 'long' wavelength emissions in sequence, creating an exhaustive set of spectra, uniformly distributed throughout the space of all possible three primary spectra with Gaussian shaped primaries and fixed full widths at half maximum (FWHMs). The discretisation of this space was controlled by the choice of wavelength and intensity quantisation values, which specified the size of the wavelength and intensity increments. The range of the primary intensities (in terms of Radiance) were chosen from the range [0,1] $W/Srm2$.

For each spectrum in the set of spectra, its CIE xyY coordinates were calculated and any spectra whose chromaticity values were further than a threshold value from the chromaticity of the target colour were discarded. It is important to note that, the thresholds used at this stage to identify metameres do not reflect discriminability of the colour as accurately as it would when using a uniform colour space like *CIELAB*. Discriminability of colours in the CIE-xyY space can be described using *MacAdam ellipses*, however, the shapes of these are not uniform in their eccentricities and orientation throughout the CIE-xyY colour space [43]. As this would have added extra complexity to this investigation, we decided that for this initial investigation, it was sufficient to use the xyY chromaticity coordinates. Carrying out this *brute-force* search in a perceptually uniform colour space, like *CIELAB*, would have been computationally intractable due to the increased complexity in calculating the colour coordinates.

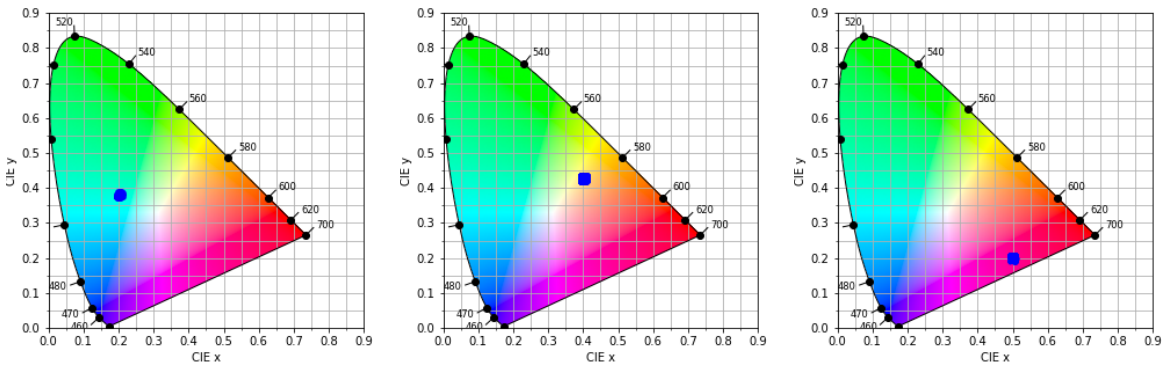
Once the metameric spectra were found, the amount of energy required to produce each of them were calculated using a simple power model derived from information about commercial LEDs (see fig. 4.2) [20]. Given the time it takes to do carry out the grid search, for this initial investigation, we chose three arbitrary colours as the target colours. Namely a *Turquoise* colour, a *Yellow* colour, and a *Pink* colour with the [x,y] chromaticity coordinates of [0.2031, 0.3806], [0.4, 0.45], and [0.5, 0.2] respectively. For the *Turquoise* colour samples we evaluated both the CIE 2° and 10° std. observer CMFs. For the *Yellow* colour sample we evaluated the CIE 2° std. observer CMFs and for the *Pink* sample we evaluated the CIE 10° std. observer CMF.

4.2.2 Simulation Results

Table 4.1 outlines the results of the brute-force search. Figure 4.3 illustrates the chosen colours and the distribution of the metameres on CIE-1931 xy chromaticity diagrams.

Turquoise	Search Parameters	Observer Intensity Quantisation Wavelength Quantisation Metamere Threshold	2° std. observer 0.2 10 nm +/- 0.01	10° std. observer 0.2 10 nm +/- 0.01
	Results	Number of Test Spectra Metameres found	388,800 379	388,800 242
Yellow	Search Parameters	Observer Intensity Quantisation Wavelength Quantisation Metamere Threshold	2° std. observer 0.1 5 nm +/- 0.01	
	Results	Number of Test Spectra Metameres found	19,677,504 3420	
Pink	Search Parameters	Observer Intensity Quantisation Wavelength Quantisation Metamere Threshold		10° std. observer 0.2 10 nm +/- 0.01
	Results	Number of Test Spectra Metameres found		388,800 673

Table 4.1 Search Parameters and results for Brute Force search for metameres.



(a) *Turquoise* (b) *Yellow* (c) *Pink*

Fig. 4.3 The chromaticities of the metameres closely matched the corresponding target colours as can be seen on the CIE-1931 xy chromaticity diagram.

Turquoise Colour Sample with 2° Std. Observer

For the *Turquoise* colour sample, from 388,800 test spectra, we found 379 metameres (0.0845%) with the 2° std. observer with an absolute threshold of +/- 0.01 in the [x,y] plane. Under similar conditions with the 10° std. observer, 242 metameres were found. The figures only show the results for the 2° std. observer. The results show that there is a *five-fold* difference between the maximum and minimum energies required to produce the spectra in this set of metameres (see fig. 4.5a). Figure 4.5b shows

the histogram of the energies, and it indicates that the energies form a multimodal distribution. The density plots of the 'short' and 'medium' wavelength emissions' intensity (*sIntensity* and *mIntensity*) show a similar shape (see fig. 4.14). The scatter plot of the 'long' wavelength intensity (*lIntensity*) and the energy also shows a similar bifurcation. This indicates that the space of metamere energies is not covered uniformly and that there is an opportunity for optimising the energy.

Figure 4.6 shows the effect of the variation in the primary intensities on the spectral generation energy. It shows that the energy is highly correlated with the 'medium' wavelength emission intensity (*mIntensity*), with the 'short' wavelength intensity (*sIntensity*) showing a smaller correlation and the 'long' wavelength emission intensity (*lIntensity*) showing a smaller correlation still.

Figure 4.14 shows that the density of the 'long' wavelength intensities (*lIntensity*) and mean wavelengths (*lMean*) are more uniformly distributed among the set of metameres. The plot of 'medium' wavelength emission wavelength (*mMean*) against 'long' wavelength emission wavelength (*lMean*) is in line with this observation, showing that large variations in *lMean* could be made without significantly affecting the output colour. The mean spectrum and the variance plot adds further credence to this hypothesis showing a spread out, low mean, moderate variance region in the 'long' wavelength range (650 nm – 820 nm) (see fig. 4.4). The oscillations in the variance are due to the coarse wavelength quantisation used in the grid search. One other point worth noting is the peak around 550 nm. The narrow peak is suggestive of the presence of a fundamental wavelength. The density plot for the *mMean* dimension in fig. 4.14 is further indicative of this, showing an anomalous aggregation of metameres around 550 nm.

We also observe that the distribution of the tristimulus values of the metameres are not uniform within the metamere threshold range as was initially expected. The density plots (plots along the diagonal on fig. 4.13) show that there are modes in the space of metamere tristimulus values. This means that, given the chosen space of metameres, some tristimulus values are more likely than others. Furthermore, as outlined before, there are modes in the energy where these tristimulus values preferentially occupy. We believe that this is partly due to the restriction imposed on the space of spectra by the wavelength and intensity quantisation. Looking at the results for the 'yellow' colour sample which used a finer quantisation (see the next section), we also hypothesise that this apparent modality is also a property of the space of metameres of this colour.

We also repeated the metamere search to find metameres with all three tristimulus values that are +/- 0.01 from the target (instead of just the x,y chromaticity coordinates as was done before). However, this yielded no metameres. We believe that this is a consequence of the limited expressive capacity of the hypothesis space. More specifically, we believe that the space from which spectra are drawn is too restrictive to be able to produce a spectrum that is symmetric in all three dimensions.

Yellow Colour Sample with 2° Std. Observer

The quantisation was reduced for the *Yellow* sample search; double the resolution in both intensity and wavelength axes. From 19.67 million test spectra, 3420 metameres were found (0.0173%). The results show that there is a ten-fold difference in energy between the maximum and minimum energies required

to produce the spectra in this set of metameres (see fig. 4.9a). The behaviour of the mean spectrum is similar to that of the *Turquoise* colour sample, showing a narrow peak around 550 nm and three modes (see fig. 4.7). The 'long' wavelength mean (*lMean*) peak has a higher mean and lower spread while the variance in the spectra in this range is much higher than that for the *Turquoise* sample. There is a high correlation between the metamere energy and the 'medium' wavelength emission intensity (*mIntensity*) while the correlations between energy and the 'short' and 'long' wavelengths (*sIntensity* and *lIntensity*) are less apparent (see fig. 4.8). The energy distribution shows an order of magnitude difference between the smallest and largest energies while the multimodality in the histogram is less distinct.

As expected, the energy is also highly correlated with the Y tristimulus value (see fig. 4.15). The scatter matrix for the tristimulus values shows a more uniform distribution of metameres, but the arrangement of x and y values along the Y dimension still shows preferential modes (see fig. 4.15). Similar to the *Turquoise* sample, the mass of the 'medium' wavelength emission wavelength (*mMean*) distribution is confined to a small region near 580 nm (see fig. 4.16).

Pink Colour Sample with 10° Std. Observer

The third target colour sample we investigated was the *Pink* colour. For this test, we used the same quantisation as the *Turquoise* sample to allow for a fair comparison but used the 10° std. observer CMFs. With this configuration, 673 metameres were found from 388,800 test spectra (0.173%). The results show that there is a seven-fold difference in energy between the maximum and minimum energies required to produce the spectra in this set of metameres (see fig. 4.12a). The mean spectrum for the metameres looks surprisingly similar to a mix between the *Turquoise* and *Yellow* mean spectra, with a broad 'long' wavelength emission, and narrow 'short' and 'medium' wavelength emissions (see fig. 4.10).

There are two key differences between the metamere distribution of this test compared to the two previous ones (see fig. 4.18). Firstly the density plot of the 'medium' wavelength emission mean (*mMean*) is multimodal with low kurtosis. Second, the energy density is more unimodal than the other two tests with a positive skew (i.e. bias towards the low energies). There seems to be an internal structure in the metameres that cause the mean spectra to change predictably. However, further investigation will be needed to devise a way to describe these variations and to predict the changes.

With all the colour samples, a substantial amount of energy saving could be observed by using metameres. However, this exploration only considered the metameres with matching chromaticities ('colour') and without considering luminance due to the limited flexibility of the spectral shapes (only 3-primaries with fixed FWHM). Although simply reducing the intensity of a metamere would clearly lead to a decrease in energy, the results give some evidence that energy saving could still be obtained while keeping the luminance the same. In figs. 4.13, 4.15 and 4.17 (*Energy versus Y* plots), vertical spread of metameres at each Y coordinate indicates that, although the luminance of those metameres were the same, the energy cost of producing them still varies. Although the energy saving is not as substantial (only up to approximately 2.5x energy difference), the results obtained illustrate the potential for achieving power saving while keeping the chromaticity and the luminance of the metamere the same.

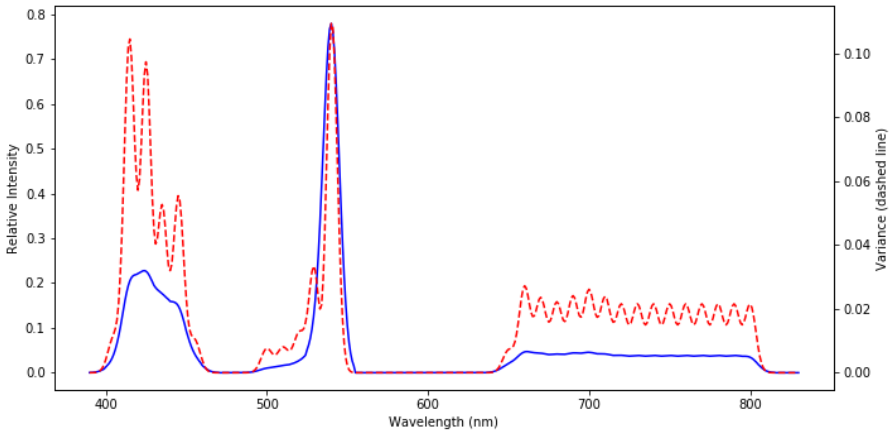
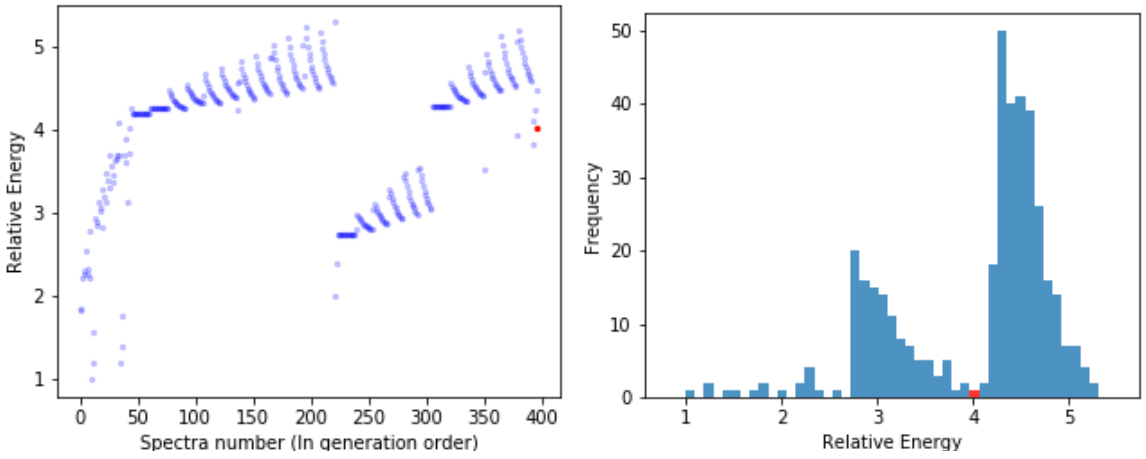


Fig. 4.4 Mean spectrum (blue) and variance (red) plot for the *Turquoise* colour sample show that some frequencies are more fundamental to generating a colour.



(a) Relative power required.

(b) Power histogram.

Fig. 4.5 Energy plots for the *Turquoise* colour sample show that there is a five-fold difference between the maximum and the minimum energies required to generate this colour (red = power of the initial target spectrum).

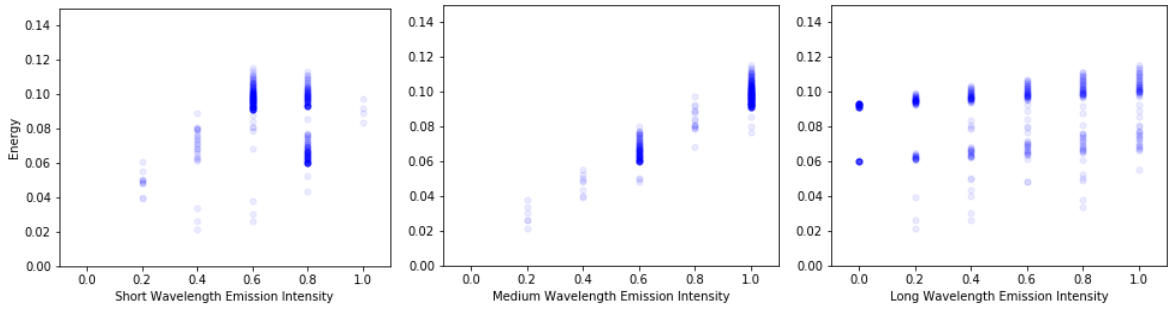


Fig. 4.6 Metamere energy is correlated most strongly with the medium primary intensity for the *Turquoise* colour metameres. Energy is in arbitrary units.

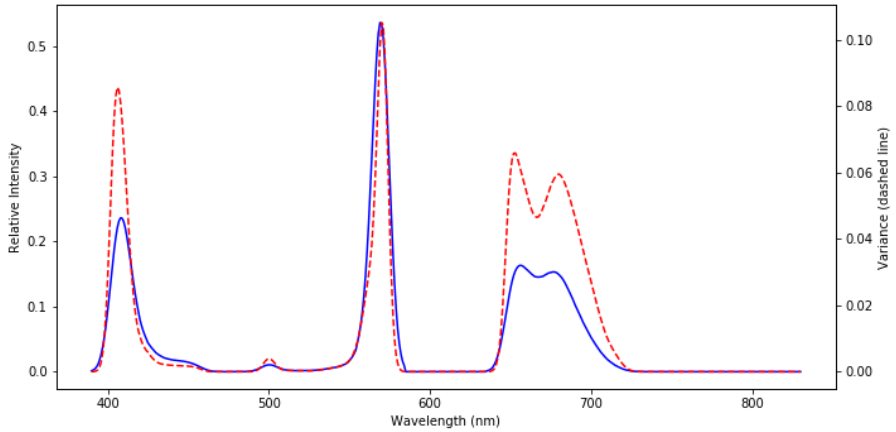


Fig. 4.7 Mean spectrum (blue) and variance (red) plots for the *Yellow* colour sample show that there are two 'fundamental' frequencies used in generating this colour.

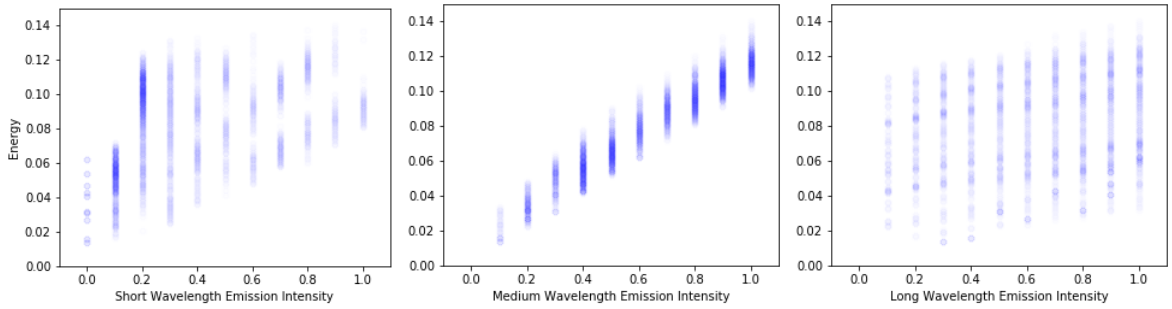
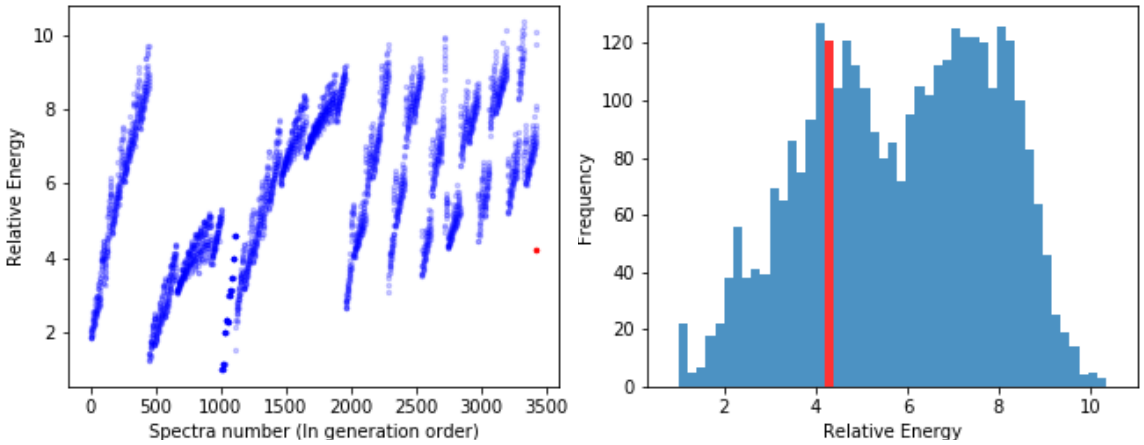


Fig. 4.8 Metamere energy is correlated most strongly with the medium primary intensity for the *Yellow* colour metameres. Energy is in arbitrary units.



(a) Relative power required.

(b) Power histogram.

Fig. 4.9 Energy plots for the *Yellow* colour sample show that there is a ten-fold difference between the maximum and the minimum energies required to generate this colour (Red = power of the initial target spectrum).

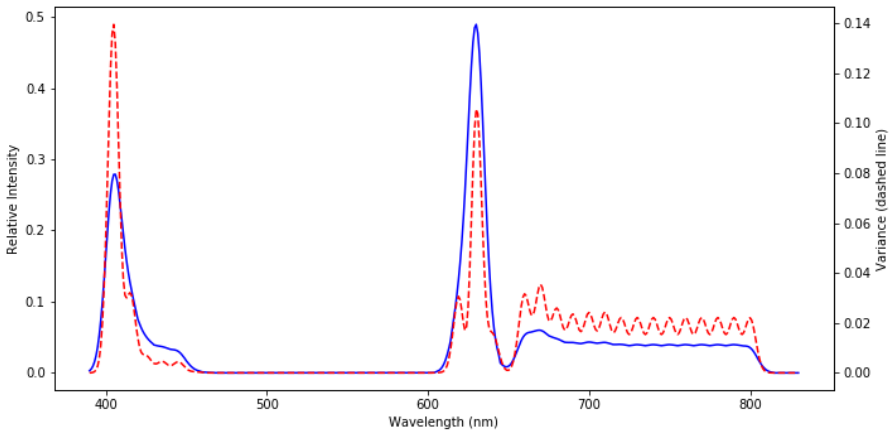


Fig. 4.10 Mean spectrum (blue) and variance (red) plots for the *Pink* colour sample show that there are two 'fundamental' frequencies used in generating this colour and the overall shape looks like a mix of the mean spectra for the *Turquoise* and *Yellow* colours.

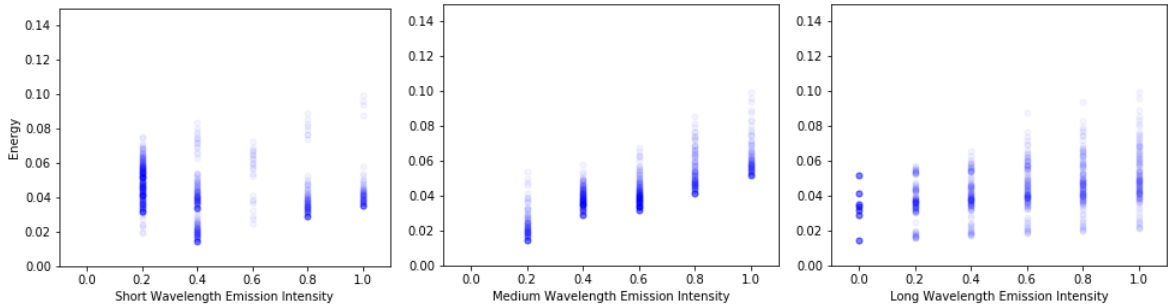
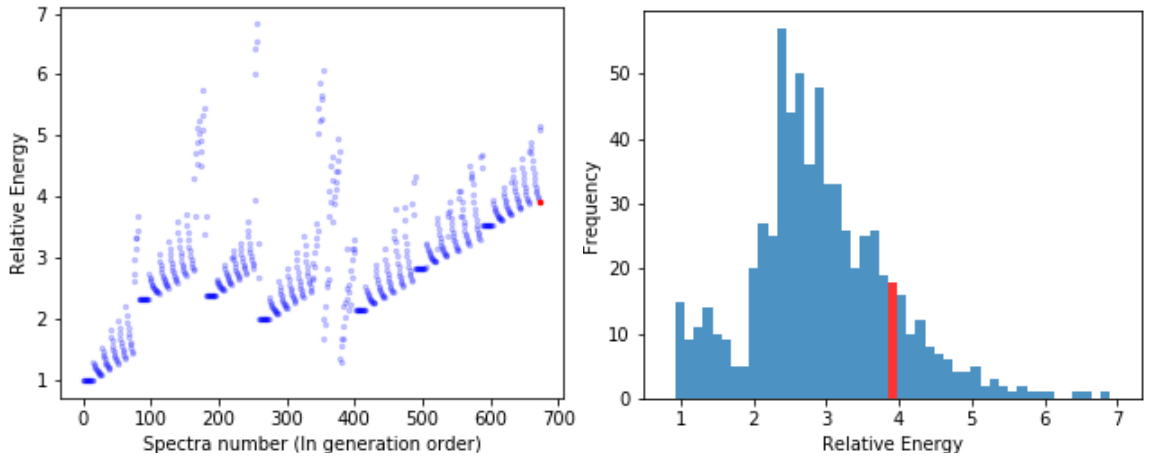


Fig. 4.11 Metamere energy is correlated most strongly with the medium primary intensity for the *Pink* colour metameres. Energy is in arbitrary units.



(a) Relative power required

(b) Power histogram.

Fig. 4.12 Energy plots for the *Pink* colour sample show that there is a seven-fold difference between the maximum and the minimum energies required to generate this colour (red = power of the initial target spectrum).

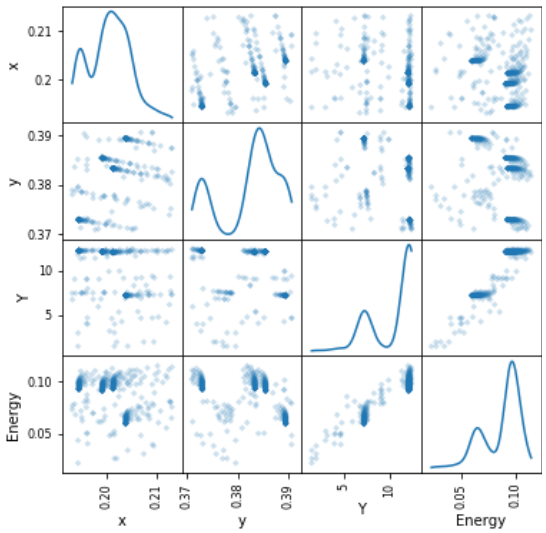


Fig. 4.13 Scatter matrix of metamere tristimulus values and energy for the *Turquoise* colour sample shows that the tristimulus values nor the spectral energies are uniformly distributed.

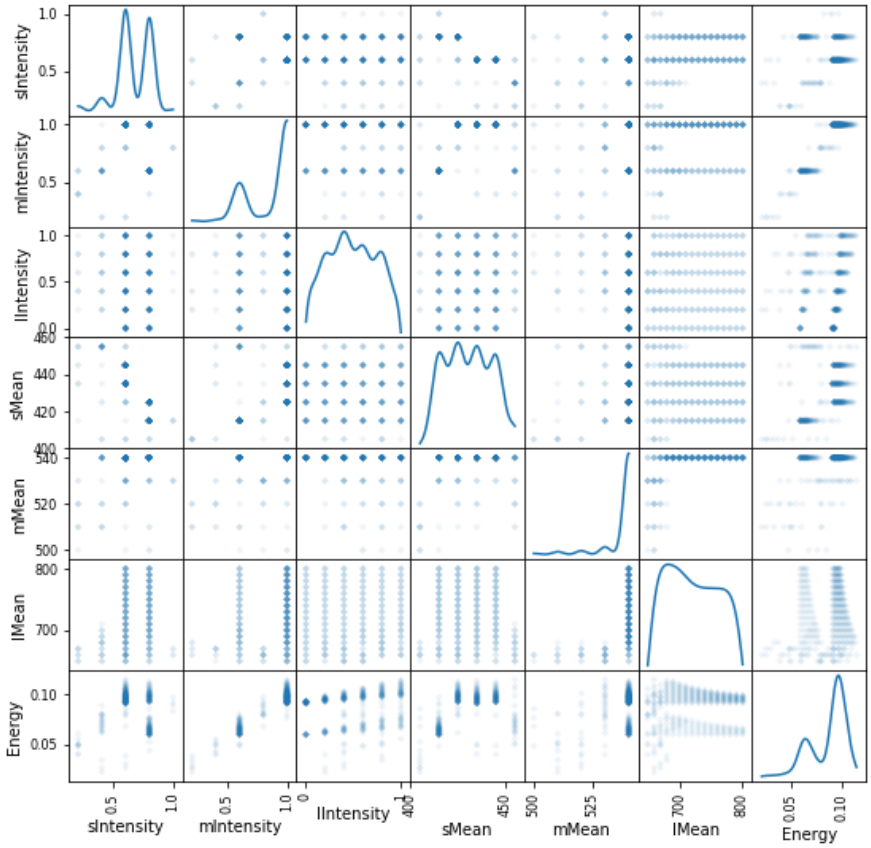


Fig. 4.14 Scatter matrix of the spectral properties of the *Turquoise* metameres show that there are modes in this space where metameres preferentially occupy.

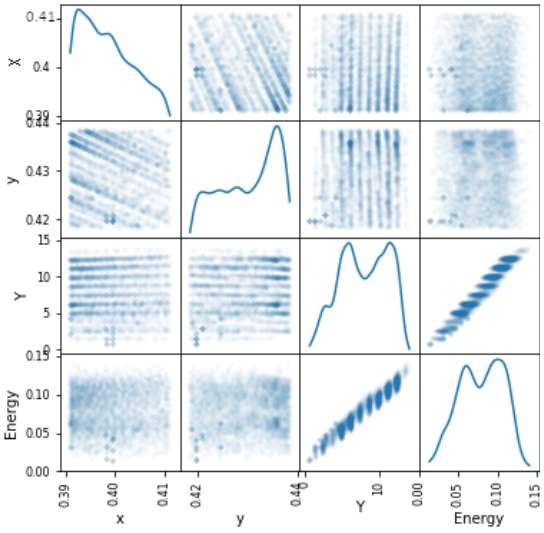


Fig. 4.15 Scatter matrix of metamere tristimulus values and energy for the *Yellow* colour sample shows that by increasing the quantisation resolution, the tristimulus values can be made to be more uniformly distributed.

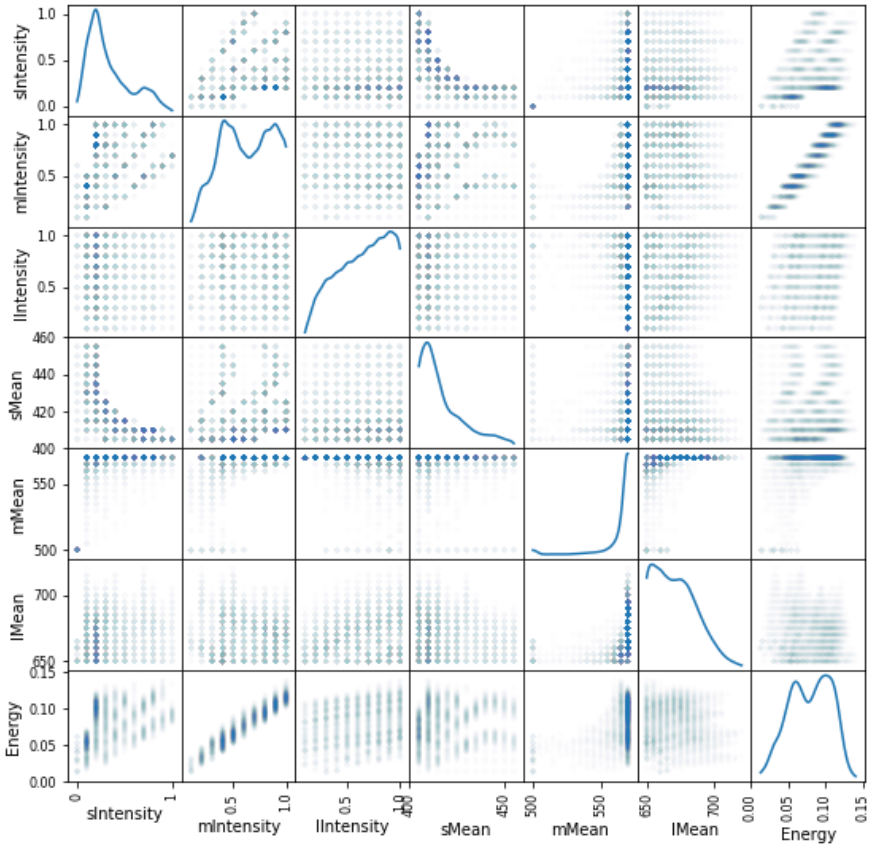


Fig. 4.16 Scatter matrix of the spectral properties of the *Yellow* metameres even with a finer quantisation, some properties (e.g., *mIntensity* and *mMean*) show a non-uniform metamere distribution.

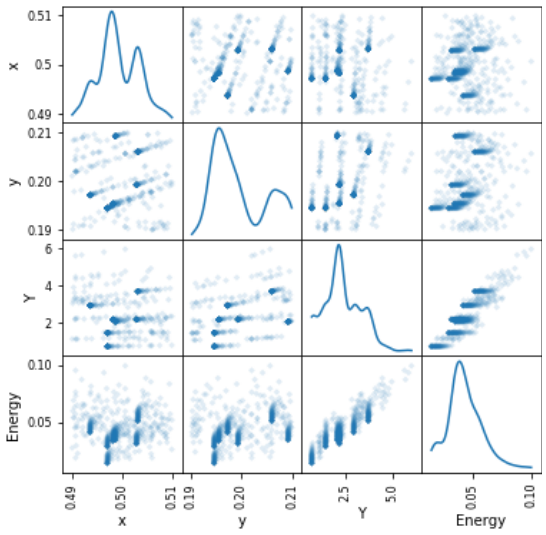


Fig. 4.17 Scatter matrix of metamere tristimulus values and energy for the *Pink* colour sample shows that the tristimulus values nor the spectral energies are uniformly distributed.

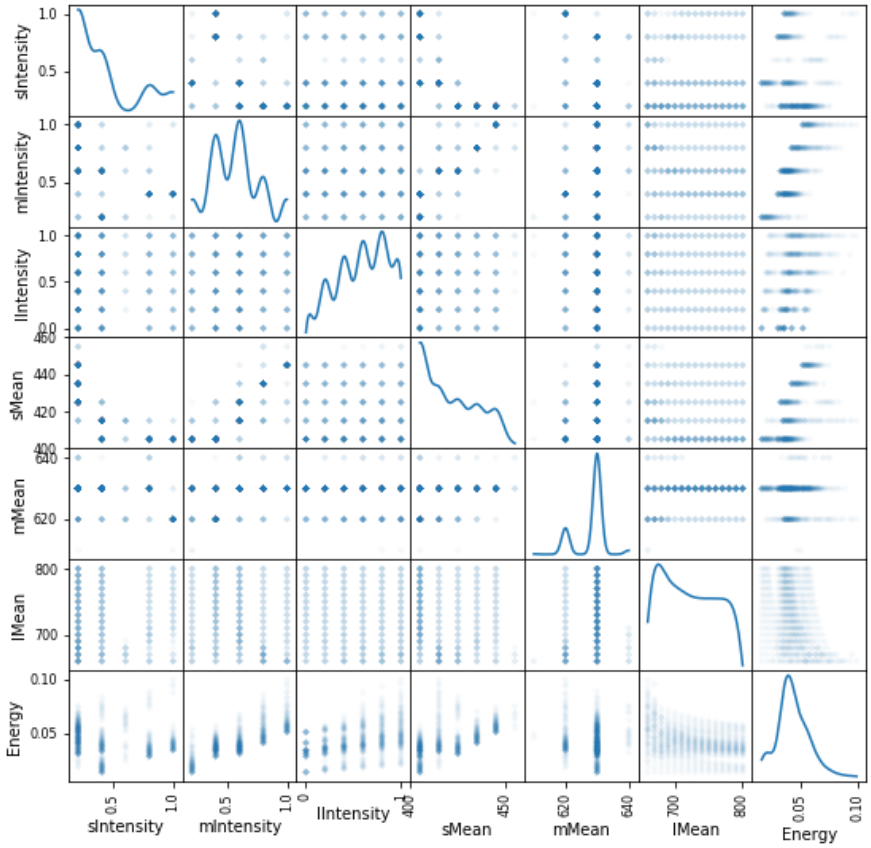


Fig. 4.18 Scatter matrix of the spectral properties of the *Pink* metameres are quite similar to that of the *Turquoise* sample indicating that the observed structure is biased by the coarseness of the quantisation.

4.3 A Mathematical Framework for Multi-Primary Displays

As outlined in section 4.2, by carefully choosing the primary frequencies, it is possible to use metamerism to generate colour sensations more energy efficiently. However, the method outlined before relies on the ability to change the wavelengths of the display's primary emissions. With the current state of the art in solid-state display technologies, it is not possible to alter the wavelength of the emission spectrum once the display has been fabricated. Unlike other display technologies, QD based displays might be uniquely positioned to allow for in-situ changes in primary wavelengths as their optical properties are tied to their physical structure. The initial idea was to investigate whether the optoelectronic properties of QDs can be modulated through piezoelectricity or QD polarisation. However, due to the COVID-19 pandemic, this line of work had to be suspended due to lab access restrictions.

Instead, we propose an alternative method that could be applied to a display, irrespective of the technology it is based on. We propose that displays with more than three primaries can be used to replicate the results observed in section 4.2. Even though the industry standard is to use three primaries, there is no fundamental limit to the number of primaries a display can have. This allows us to hypothesise potential displays with an arbitrary number of primaries. We will refer to these as *multi-primary displays (MPDs)* when referring to the idea of having more than three primaries, or as *n-primary displays* when the number of primaries needs to be highlighted.

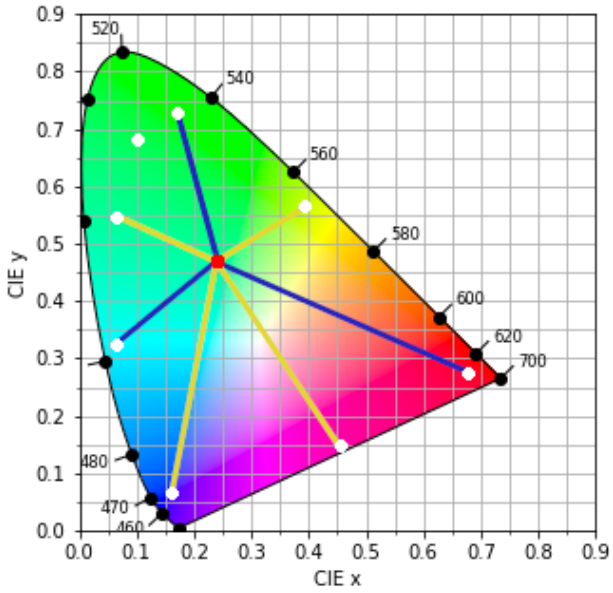


Fig. 4.19 Two Possible ways (blue and yellow paths) of achieving a colour c (red dot) in an 8-primary display. White dots are the chromaticities of the display's primary emissions

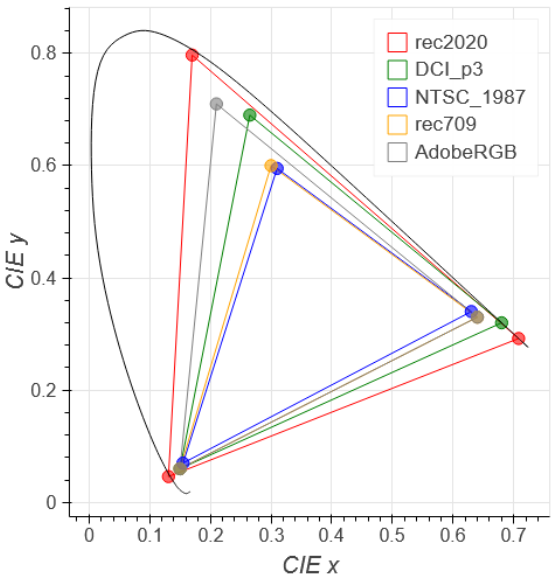


Fig. 4.20 Common display colour gamut standards plotted on the CIE-1931 chromaticity diagram.

4.3.1 Mathematical Framework

Consider an n-primary display, P_n, characterised by the set $\mathbf{P}_n$ containing its primaries, defined as,

$$\mathbf{P}_n = \{p_0, p_1, ..., p_{n-1} \mid p_i = [X_i, Y_i, Z_i], \forall i < n\} \tag{4.1}$$

The primaries, p_i are defined in terms of their coordinates on the CIE-1931 XYZ colour space, as is standard. The power set, $\wp(\mathbf{P}_n)$, represents the set of all the subsets of $\mathbf{P}_n$. In other words, $\wp(\mathbf{P}_n)$ contains all the possible ways the primaries can be combined together. For example, $S \in \wp(\mathbf{P}_n)$ such that $S = \{p_0, p_1, p_2\}$ is the set containing only the first three primaries. Although the empty set $\phi \in \wp(\mathbf{P}_n)$ and ϕ cannot produce any colour sensations, for notational simplicity it will be assumed that $\wp(\mathbf{P}_n)$ has 2^n elements each with a potential for generating useful colour sensations.

Given a particular colour, $C = [X_C, Y_C, Z_C]$, each $S \in \wp(\mathbf{P}_n)$ has a different way (i.e., different primary intensities) of achieving that colour. In other words, there are (potentially) 2^n possible ways of producing a given colour sensation. Figure 4.19 illustrates two of the possible ways of achieving the colour C in a hypothetical 8-primary display.

This can be represented by a set, $\mathbf{V}^C$, of 2^n n-dimensional vectors, each quantifying the intensities of the primaries used.

$$\mathbf{V}^C = \{V_{S_i}^C \mid \forall S_i \in \wp(\mathbf{P}_n), V_{S_i}^C \in \mathbb{R}^n\} \tag{4.2}$$

A function, f, can then be defined on $\mathbf{V}^C$ such that it takes in each n-dimensional vector and calculates a score. This f can be any arbitrary function based on the operational goals of the display. For example, the score calculated by f could be the energy required to produce the colour C using each of the 2^n possible ways, it could be the colour rendering index (CRI) for each combination of primaries or something much more complex like the effect on the wear level for each of the primaries. The score can then be used to pick the optimal way of displaying the colour C, given the operational objective dictated by f.

4.3.2 Measure of Metameric Capacity and Energy Saving Potential

Consider a colour space $\mathbf{k}$, described by the CIE-1931 XYZ colour coordinates of the primary vertices, such that $\mathbf{k} = \{p_0, p_1, p_2, \mid p_i = [X_i, X_i, Z_i]\}$. Let $G(\mathbf{x})$ be the colour gamut; the set of colours enclosed by the polygon defined by the set of primary vertices $\mathbf{x}$ (as described by CIE-1931 XYZ colour coordinates).

Thus, for each colour C in the colour gamut $G(\mathbf{k})$, the primary combinations that can achieve that colour C is given by,

$$\mathbf{h}_C = \{S \mid C \in G(\mathbf{S}), \forall S \in \wp(\mathbf{P}_n)\} \tag{4.3}$$

Consequently, the number of ways a particular colour sensation can be achieved is given by the number of primary combinations $S \in \mathbf{h}_C$ (i.e., the cardinality, $|\mathbf{h}_C|$). The total number of ways each colour in the gamut can be achieved is then given by,

$$\alpha = \{|\mathbf{h}_C| \mid \forall C \in G(\mathbf{k})\} \tag{4.4}$$

Let $H : G(\mathbf{k}) \rightarrow \alpha$, be a function that, given a colour, returns the number of ways the colour can be achieved with the available primary combinations. Similarly, let $E : \wp(\mathbf{P}_n) \times G(\mathbf{k}) \rightarrow \mathbb{R}$, be a function that, given a colour and a choice of primary combinations, returns the energy required to generate that colour.

Metamerism Index

Based on the above definitions, a metamerism index, M_I, can be defined, that describes the metameric freedom of a particular display (as described by the set of primaries $\mathbf{P}_n$), with respect to a chosen colour space $\mathbf{k}$. Let M_I be a measure of the average number of ways a given colour in the colour gamut can be generated and is defined by eq. (4.5).

$$M_I = \frac{\sum_{\forall C \in G(\mathbf{k})} H(C)}{|G(\mathbf{k})|} \tag{4.5}$$

Metameric energy-saving index

Similarly, the energy-saving index, M_E, can be defined as the total, minimum, power required to generate all the colours in the colour gamut.

$$M_E = \frac{\sum_{\forall C \in G(\mathbf{k})} \min_{S \in h_c} E(S,C)}{|G(\mathbf{k})|} \tag{4.6}$$

For both indices, the normalisation term $|G(\mathbf{k})|$ is used to take into account variations due to differences in the quantisation of the colour space. This can be changed to $|G(\mathbf{k}) \cap G(\mathbf{P_n})|$ if the effects of incomplete coverage of the colour gamut by $\mathbf{P}_n$ needs to be ignored.

4.3.3 Advantages of Multi-Primary Displays

The optimisation presented in section 4.3.1 can be done at run-time from the software level, and the MPD display controller can switch between different operational objectives by changing f on-the-fly. Further algorithms can be devised that consider other properties of the content on the screen such as the pixel change in colour between different screen refresh cycles to further optimise the colour transitions. System designers will even have the ability to develop new objective functions after the system has been launched to the public and sending the new optimisation objectives to the users' devices as a software/firmware update. This optimisation can be done either in the embedded layer (where the TFT drivers reside) or at the system level (alongside software display drivers). Although the implementation

details of such optimisation algorithms are beyond the scope of this thesis, we believe that an offline optimisation step followed by a simple lookup table based run-time optimisation setup could easily be implemented for such a display. In chapter 5, we explore how to use the limits of human colour acuity and the colour dependent properties of the display pixel primaries to optimise the content being displayed on with the objective of reducing the display power dissipation. With the use of a suitable abstraction layer, such an MPD can also be made to be backwards compatible with devices designed to work with existing three-primary displays.

Finally, redundancy introduced by having multiple primaries allows these displays to tolerate sub-pixel failures elegantly and avoid 'dead pixel' issues faced by traditional displays, allowing display manufacturers the opportunity to 'bin' displays in a similar way to how integrated circuits get 'binned'.

The main advantages of smart MPDs can be summarised as follows;

- Wider colour gamut.
- Increased energy efficiency.
- Optimisable for arbitrary operational objectives.
- Feature upgrades after device fabrication.
- Content-aware optimisations at runtime.
- Resilience to sub-pixel failures.

4.4 Metameric Breakdown in QD-LED Displays

Typically the CIE standard observers are used to quantify the perceived colour of a spectrum. However, the narrow bandwidths of the QD emissions make QD-LED based devices more susceptible to variations between observers, compared to devices using primaries with wider emission bandwidths, as outlined in section 3.5 [84]. The effect of two observers disagreeing on the perceived colour of a spectrum is termed *observer metameric breakdown*. In the sections that follow, we show, using Styles and Burch's original 10° individual colour matching functions the effect of emission bandwidths and primary wavelengths on *observer metameric breakdown* [43, 22, 47].

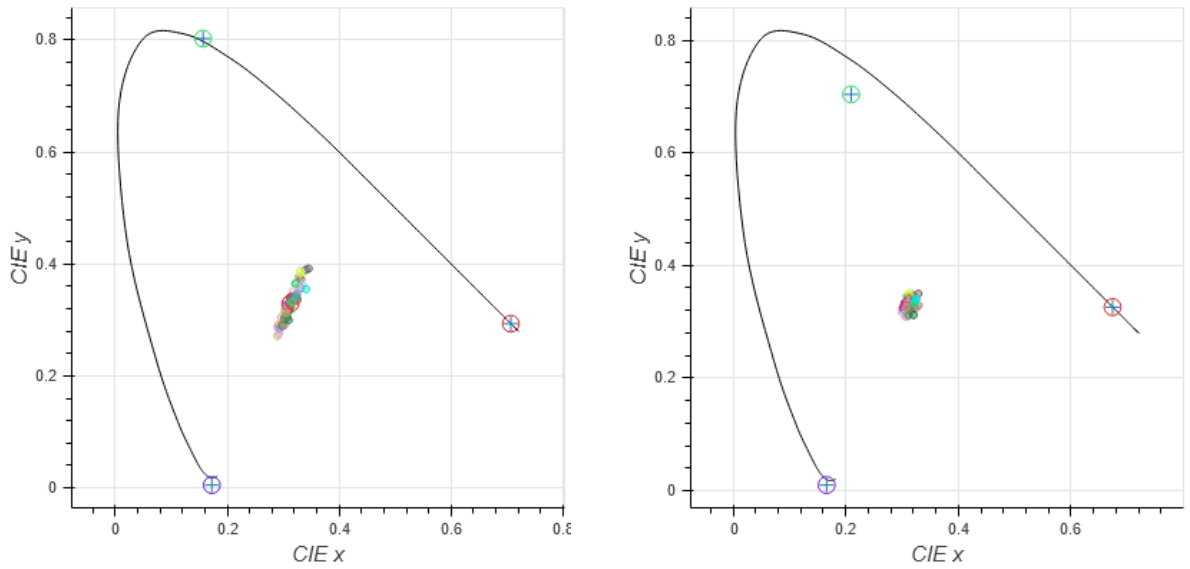
4.4.1 Metameric Breakdown Sensitivity to FWHM

For this exploration, a set of three-primary spectra were generated in a similar way to section 4.2, but with the emission wavelengths fixed. Using a simple *L-BFGS* optimiser, the primary intensities required to match the chromaticity of CIE-D65 spectrum was calculated. The process was repeated for the same wavelength combination, but with different FWHMs. Figure 4.21 shows the results of this process for a spectrum with primaries at 410nm, 530nm, and 630nm. We used the Styles and Burch 10° individual colour matching functions as the set of observers to quantify the observer variability. Each coloured dot in the centre of the plot indicates the tristimulus values of a particular observer when viewing the calculated spectrum, plotted on the CIE-xy chromaticity diagram. Circles with the plus sign are the locations of the primaries according to the 10° std. observer. As can be seen, for primaries

with a bandwidth of 10 nm (fig. 4.21a), the spread of the tristimulus values is quite large, indicating disagreement in the perceived colour, while at a 50nm (fig. 4.21b), the variation is much lower showing minimal observer metameric breakdown. Figure 4.22 shows the spectra with 10 nm, 20 nm, and 50 nm FWHMs along with a cropped in plot of the CIE-1931 chromaticity diagram showing the variation in the observers.

The initial hypothesis was that we could mitigate this by using multiple primaries to stabilise the colour. Although we did observe that using multiple primaries does indeed minimise the spread of the individual observer tristimulus values, we believe this is due to an artefact of the way the tristimulus values are calculated from the emission spectra and not due to a stabilising effect introduced by the additional primaries.

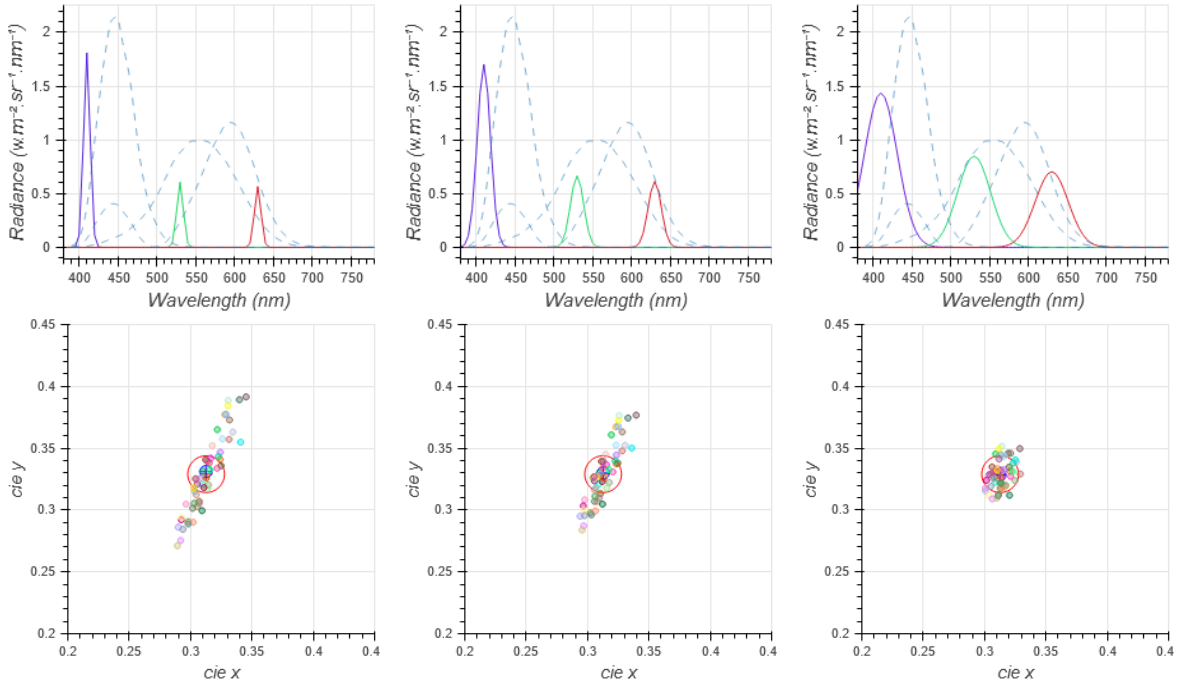
Figure 4.23 shows the same information as fig. 4.22, but for a device with the primaries arranged at 430 nm, 530 nm, and 600 nm. As can be seen, the metameric breakdown in this case is minimal. Figure 4.24 shows how the standard deviation of the individual observer tristimulus values vary with FWHM. In the mitigated case (fig. 4.24b), the deviation is more or less constant, showing that even extremely narrow QD primaries would not lead to substantial observer metameric breakdown. Similarly, fig. 4.25 shows the effect of the centre wavelength of the primaries on the individual observer std. deviation. As can be seen, at narrow FWHMs, the observer deviation is extremely sensitive to the primary wavelength (see fig. 4.25a), which reduces with increasing FWHM (see fig. 4.25d). Thus, by carefully picking the primary wavelengths to avoid the locations where the variation in the observers are maximal, observer metameric breakdown can be minimised.



(a) 10 nm FWHM.

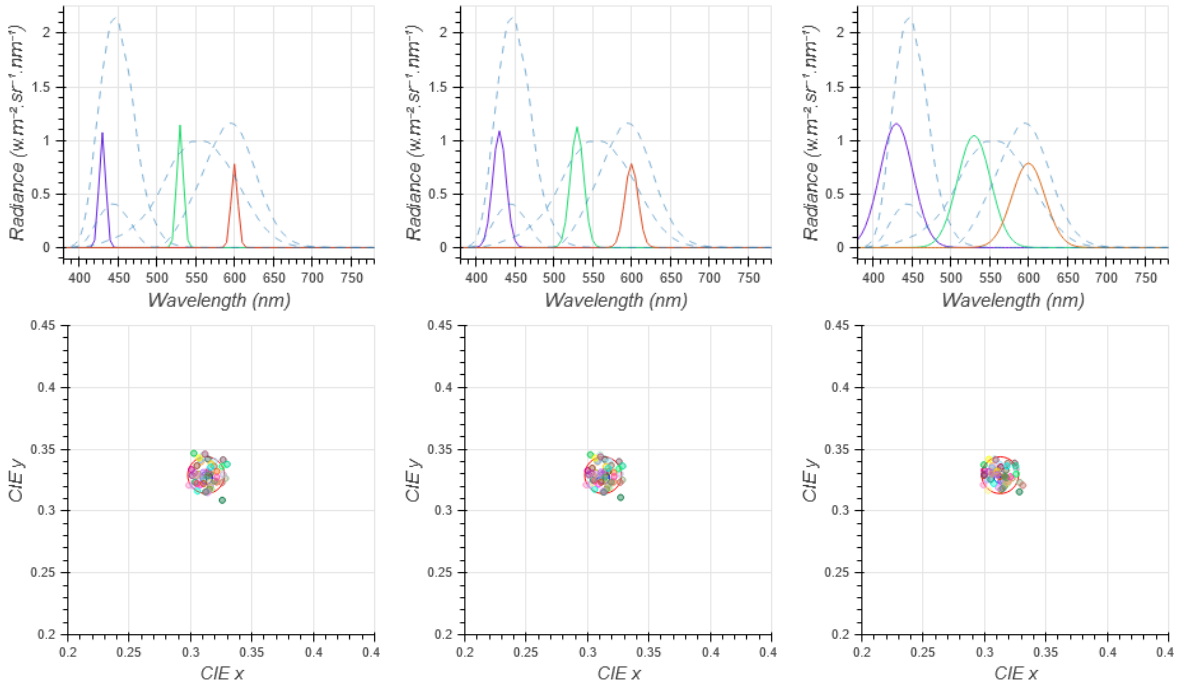
(b) 50 nm FWHM.

Fig. 4.21 At narrow FWHMs a substantial amount of observer metameric breakdown can be observed (left), which can be reduced by increasing the FWHM (right).



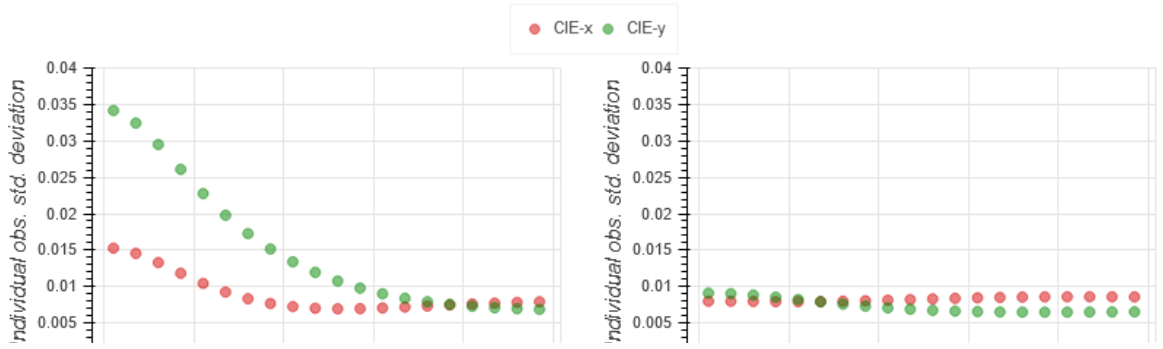
(a) 10 nm FWHM. (b) 20 nm FWHM. (c) 50 nm FWHM.

Fig. 4.22 In the extreme case, narrow wavelengths cause substantial deviation in the individual observers' tristimulus values, implying a high degree of observer metameric breakdown.



(a) 10 nm FWHM. (b) 20 nm FWHM. (c) 50 nm FWHM.

Fig. 4.23 The metameric breakdown can be mitigated be carefully picking the location of the primary wavelengths.



(a) Extreme Case. (b) Mitigated Case.

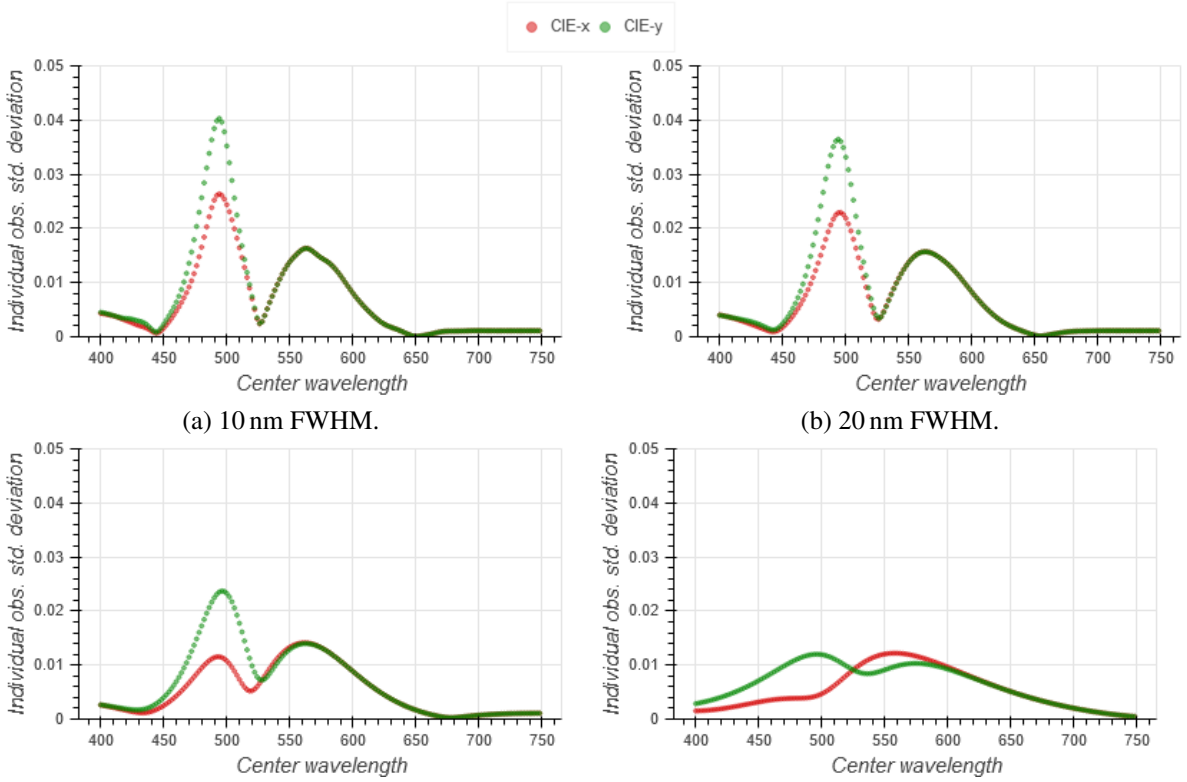
Fig. 4.24 In the extreme case, a substantial degree of metameric breakdown can be seen even up to FWHM=50. When primaries are carefully chosen, this deviation is minimal across the range of FWHMs tested.

4.5 Conclusions and Future Work

This chapter highlighted how metamerism could be leveraged to achieve the same colour sensation while consuming up to 10 times less power. It presented a mathematical description of an MPD and based on it, derived two indices for measuring the metameric capacity of a display. Finally, the chapter explored how narrow primary bandwidths can lead to metameric breakdown and how to mitigate it by evaluating the sensitivity of the Styles & Burch individual observers to the FWHM and the wavelengths of the primary emissions.

As highlighted in this chapter, the metamere search focussed exclusively on chromaticity match. Consequently, the effect of luminance on the power consumption was not factored out (as was seen). The next step would be to factor out the effect of luminance by using a more flexible device architecture (e.g., more than three primaries) with a range of target colours. Using an optimisation process would be more ideal in this case as it would be intractable to do a grid search in such a high dimensional space.

The MPD architecture presented in this chapter was not evaluated for its metameric capacity. Thus, conducting a metamere search on hypothetical, fixed primary, QD displays would be required to evaluate the efficacy of using such an architecture for energy saving. Work along this avenue was carried out as a part of the work presented in this chapter. However, the preliminary results were inconclusive at the time of writing and hence was left out. The metameric indices that are presented should also be evaluated with respect to these displays to assess their usefulness in a real-world application.



(a) 10 nm FWHM.

(b) 20 nm FWHM.

(c) 50 nm FWHM.

(d) 100 nm FWHM.

Fig. 4.25 When using primaries with narrow FWHMs, the individual observer deviation is extremely sensitive to the primary wavelength. The sensitivity decreases with increasing primary FWHM.

Inspiration Is for Amateurs. The Rest of Us Just Show Up and Get to Work.

— Chuck Close.
Inside the Painter's Studio [122, p. 42].

Chapter 5

Approximate Colour Transformations

5.1 Introduction

This chapter focusses on exploring how to use the dynamic power consumption of emissive pixel displays (e.g., OLED/QD-LED displays) to minimise the display power consumption through a class of image transforms called *approximate colour transforms*. In particular, it is concerned with *Crayon*; an approximate colour transform presented by Stanley-Marbell et al. in 2016 [21]. This chapter initially presents an outline of how *Crayon* works and how the authors evaluated its performance. We then present the design of an updated user study aimed at better quantifying the effect of *Crayon* on the perceptual quality of images. We detail how the user study was conducted, the results we obtained, and how these results were used to derive a method for calculating optimal transformation parameters from image features.

5.1.1 Operating Principles of Approximate Colour Transforms

The aim of an *approximate colour transform* is to modulate the colours on an image with respect to some target display, such that the resulting image consumes less energy when displayed on this display. This is analogous to how typical image compression algorithms aim to transform images such that they consume less storage space when stored. Image compression compresses images in the domain of disk space, *approximate colour transforms* 'compress' images in the energy space. In both cases, it is important to ensure that this 'compression' process does not severely degrade the perceptual quality of the image (e.g., by introducing noise or compression artefacts). Typically, such transforms comprise a model of the display energy, a model of human vision, an objective function, and an optimiser [117–121].

5.1.2 Contributions

This chapter presents the following contributions:

- Confirmation of previous work in literature through a user study with more than 60 participants.
- A parameter normalisation scheme for *Crayon*.

- A content-aware parameter selection method for the *Crayon* colour transform that uses simple image features to efficiently identify an optimal parameter-lower-bound required to achieve given image quality.

5.2 What is *Crayon*?

In 2016, Stanley-Marbell et al. presented *Crayon*, a system of image transformations that aims to achieve display power savings by leveraging perceptual limitations of the human vision system and the dynamic power consumption of OLED displays [21]. The original *Crayon* system has 2 image transforms; a colour transform and a shape transform. These transforms aim to minimise the power consumption of an image when displayed on an emissive pixel display by modulating the colours and the shapes in the image, respectively. The work presented in this chapter focusses exclusively on the *Crayon* colour transform.

Crayon was designed to work as an image pre-processor ('offline' optimisation) or as a display library add-on to intercept display drawing calls to carry out these optimisations while the display device is in use ('online' optimisation). Compared to similar work, the efficiency of *Crayon* is primarily due to the use of a low order polynomial to approximate the display power model [117, 119, 120]. Using this, a closed-form solution to the optimisation process was derived by the authors. When used as an online optimiser, *Crayon* is able to access both shape and colour of the content by intercepting low-level calls from the drawing library application programming interface (API) and optimise the content prior to rasterisation.

It was shown in the original work that, at an extended graphics array (XGA) resolution, using just the colour transform gave a net positive power saving if the transformed image was displayed on the screen for more than 2 seconds [21]. For the case where both colour and shape transforms were used, the display time for break-even power was shown to be as small as 1ms.

5.3 How does the Crayon Colour Transform work?

The *Crayon* transforms use a simple, sum of quadratics model of the display power as its basis. The power dissipation of each sub-pixel is modelled as a quadratic function. This allows the problem of minimising display power to be posed as a minimisation for each pixel independently, assuming that the power of a pixel is decoupled from its neighbouring pixels. A constraint is added using a Lagrange multiplier to control the magnitude of the pixel colour modulation. This system does not incorporate properties of the HVS explicitly, but rather uses results from a user study to evaluate its perceptual performance and quantify the effect of the transform parameters on the perceptual quality of the transformed images. The section below describes the *Crayon* colour transform mathematically.

5.3.1 Mathematical Formulation of Crayon

Let the sub-pixel power for the colour channel c be:

$$p_c(s) = \frac{1}{2}\alpha_c(s)^2 + \beta_c s + \gamma_c \tag{5.1}$$

where $s \in [0,1]$ is the normalised intensity of the colour channel. In the case of an image in *sRGB* colour space, $c \in \{R,G,B\}$. See appendix A for details on the procedure for measuring the model parameters α_c, β_c, and γ_c. Quadratic functions were shown to provide a good fit for the observed power measurements while being simple enough to be used in real-time optimisation [21].

Let x be an N-pixel, *sRGB* input image with colour channels r, g, and b. The total power required to display this image on the display is given by:

$$P(x) = \sum_{c\in\{r,g,b\}} \sum_{i=1}^{N} \frac{1}{2}\alpha_c(x_c[i])^2 + \beta_c x_c[i] + \gamma_c \tag{5.2}$$

where $x_c[i]$ is the image intensity of channel c at pixel i, and α_c, β_c, and γ_c are the power model parameters for the channel c.

The goal is to find an image y that approximates x but dissipates less display power, while ensuring that the perceptual difference between x and y is small. This can be formulated as the following minimisation problem:

$$\min_y \; P(y) \text{ s.t } \phi(y-x) < \varepsilon \tag{5.3}$$

where ϕ is a convex function that measures the distance between the images, and ε quantifies the acceptable approximation (i.e., the image similarity measured by ϕ).

The chosen function ϕ should simplify the optimisation while providing a meaningful comparison for images. In their original work, Stanely-Marbell et al. used Euclidean (ℓ_2) and Squared Euclidean (ℓ_2^2) distances.

With $\alpha_c > 0$, $P(y)$ is a convex function and the minimisation problem is convex and has a unique minimiser. Moreover, there exists a Lagrange multiplier Λ, whose value depends on ε, that defines an unconstrained problem with the same minimiser:

$$\min_y P(y) + \Lambda\phi(y-x) \tag{5.4}$$

This unconstrained formulation can then be used to calculate stationary solutions for ℓ_2 and ℓ_2^2 distances.

5.3.2 Least-Squares (ℓ_2^2) Solution

For the case where ϕ is the ℓ_2^2 distance, the image distance is given by, $\phi(y-x) = \frac{1}{2}\|y-x\|_2^2$. Consequently, the minimisation is given by,

$$\min_{y} \sum_{c\in\{r,g,b\}} \sum_{i=1}^{N} = \frac{1}{2}\alpha_c(y_c[i])^2 + \beta_c y_c[i] + \gamma_c + \frac{\Lambda}{2}\|y_c[i] - x_c[i]\|_2^2 \tag{5.5}$$

The problem above is decoupled for each pixel and image channel. As a result, the global minimiser can be obtained by independently minimising

$$\min_{y_c[i]} \frac{1}{2}\alpha_c(y_c[i])^2 + \beta_c y_c[i] + \gamma_c + \frac{\Lambda}{2}\|y_c[i] - x_c[i]\|_2^2 \tag{5.6}$$

for each pixel i and colour channel c. Calculating the derivative with respect to $y_c[i]$ and finding the stationary point leads to the closed form solution for optimum value of the colour channel c of the i^{th} pixel:

$$y_c[i] = \frac{\Lambda x_c[i] - \beta_c}{\Lambda + \alpha_c} \tag{5.7}$$

5.3.3 Euclidean-distance (ℓ_2) Solution

For the case where ϕ is the ℓ_2 distance, the image distance is given by, $\phi(y-x) = \|y-x\|_2$. β_c and γ_c are set to zero to simplify the power model. The minimisation then becomes,

$$\min_{y} \sum_{c\in\{r,g,b\}} \sum_{i=1}^{N} = \frac{1}{2}\alpha_c(y_c[i])^2 + \gamma_c + \Lambda\|y_c[i] - x_c[i]\|_2 \tag{5.8}$$

Unlike with the ℓ_2^2 distance, the square root term in the ℓ_2 distance couples the colour channels together, thus the optimisation will need to be carried out with respect to all the colour channels together.

Let $\vec{y}_i = (y^r[i], y^g[i], y^b[i])$ and $\vec{x}_i = (x^r[i], x^g[i], x^b[i])$ be column vectors, where $\vec{y}[i], \vec{x}[i] \in \mathbb{R}^3$. Furthermore, let D_α be a diagonal matrix with the second order power model coefficients α_r, α_g, and α_b, on the diagonal. This would reformulate the optimisation for the i^{th} pixel to be

$$\min_{\vec{y}_i} \frac{1}{2}\vec{y}_i{}^T D_\alpha \vec{y}_i + \Lambda\|\vec{y}_i - \vec{x}_i\|_2 \tag{5.9}$$

Note that, expanding $\vec{y}_i{}^T D_\alpha \vec{y}_i$ gives $\alpha_r y^r[i]^2 + \alpha_g y^g[i]^2 + \alpha_b y^b[i]^2$, which is the result from the summation in eq. (5.8).

Consider a change of variables to $\vec{z}_i = \vec{y}_i - \vec{x}_i$:

$$\min_{\vec{z}_i} \frac{1}{2}(\vec{z}_i + \vec{x}_i)^T D_\alpha (\vec{z}_i + \vec{x}_i) + \Lambda\|\vec{z}_i\|_2 \tag{5.10}$$

The second term in the above function depends only on the magnitude of the vector $\vec{z}_i$. The first term is minimised when $\vec{z}_i$ is opposite to $\vec{x}_i$ (i.e., equal magnitude and antiparallel). Considering the more general anti-parallel case, the vector $\vec{z}_i$ that minimises the function satisfies $\vec{z}_i = -\mu\vec{x}_i$, for some $\mu > 0$. This simplifies the problem in $\mathbb{R}^3$ to an optimisation in $\mathbb{R}$.

Substituting $\vec{z_i} = -\mu\vec{x_i}$ gives:

$$\min_{\mu} \frac{1}{2}(-\mu\vec{x_i} + \vec{x_i})^T D_\alpha(-\mu\vec{x_i} + \vec{x_i}) + \Lambda\|-\mu\vec{x_i}\|_2 \tag{5.11}$$

Which simplifies to:

$$\min_{\mu} \frac{1}{2}(1-\mu)^2 \vec{x_i}^T D_\alpha \vec{x_i} + \mu\Lambda\|\vec{x_i}\|_2 \tag{5.12}$$

The problem is again differentiable and we can solve it by differentiating with respect to μ and equating to zero. Doing so gives us a closed-form solution for the optimum value of the i^{th} pixel in the transformed image:

$$\vec{y_i} = (\mu+1)\vec{x_i} \qquad \text{with } \mu = max\left(1 - \frac{\Lambda\|\vec{x_i}\|_2}{\vec{x_i}^T D_\alpha \vec{x_i}}, 0\right) \tag{5.13}$$

5.3.4 Image Colour Spaces

Most modern displays and high definition multimedia content use the *rec.709* colour standard [54]. This means the content designed for it natively represents colours in terms of the *RGB* primaries. In the original *Crayon* work, the authors also explore colour transforms in the *CIELAB* colour space, which was designed to reflect the way humans perceive colour differences [123].

5.3.5 Multi-Primary Displays

The *Crayon* transforms can also be adapted to work with MPDs quite simply. As an example, for an *RGBY* display, the power model can simply be extended to have another colour axis

$$p_c(s) = \frac{1}{2}\alpha_c s^2 + \beta_c s + \gamma_c \text{ with } c \in \{R,G,B,Y\} \tag{5.14}$$

In general, for an n-primary display, the power model could be expressed as;

$$p_c(s) = \frac{1}{2}\alpha_c s^2 + \beta_c s + \gamma_c \text{ with } c \in \{\sigma_1, \sigma_2, \ldots, \sigma_n\} \tag{5.15}$$

where $\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$ represents the primaries of the display. This new power model can then be used to derive the *Crayon* transforms outlined in section 5.3.1. However, since most images are in the *sRGB* colour space, a function for transforming from the image space to the display space would be required. Since the manufacturer would need to provide these functions for the nominal operation of an MPD to begin with, this does not place a substantial extra burden on the manufacturer. However, this transformation will need to be simple and analytic to allow for closed-form solutions to the *Crayon* transformations to be obtained.

5.3.6 Controlling the Transform

The *Crayon* transform applies the eq. (5.7) or eq. (5.12) at each pixel individually to calculate the optimised image. Thus the *Crayon* transform can be vectorised and be made to run efficiently in parallel on the the target device's graphics processing unit (GPU).

There are three things that dictate the output of the *Crayon* transform, namely:

1. The colour space in which the optimisation is carried out
2. The distance metric that specifies image similarity
3. The trade-off parameter, Λ (Lambda), that specifies the strength of the approximation

Results from the original study showed that, typically, the ℓ_2^2 distance produces a solution with many small colour approximations while the ℓ_2 distance produces a solution with a fewer but larger colour approximations [21]. Smaller the Λ, stronger the transform will be and larger the power saving, but more significant the image quality degradation. The authors showed that these transforms can be implemented efficiently to have a low computational cost. The ℓ_2^2 colour transform (eq. (5.7)) for example requires only four arithmetic operations per colour channel. They show that these transforms produce images of acceptable quality to a significant proportion of participants in a large user study while significantly reducing the display power dissipation.

5.4 The Original *Crayon* User Study



Fig. 5.1 The original *Crayon* study used three bitmap images and two vector images as the baseline [21].

As outlined in section 5.3, the *Crayon* formulation does not rely on an HVS model to carry out its colour optimisation. Thus, there is no explicit guarantee that the images transformed using *Crayon* will have the required perceptual quality. To assess the perceptual performance of the transformed images as a function of Λ, a user study was carried out by Stanley-Marbell et al. as a part of the original *Crayon* work.

The original *Crayon* user study comprised two parts; one for the colour transform and one for the shape transform. The colour transform component of the study used three bitmap images and two vector images as the input images (see fig. 5.1) [124]. These images were transformed using the ℓ_2 and ℓ_2^2 distance metrics in the *sRGB* and *CIELAB* colour spaces (four possible distance metric and colour space configurations). For each transform configuration, a range of Λ values were chosen, with the maximum

Λ producing transformed images 'indistinguishable from the original' and the minimum Λ producing 'almost unacceptably approximated' images. The study also evaluated two 'darkness transforms' which simply reduced the image brightness in *sRGB* or *CIELAB* colour spaces with 40 uniformly-sampled brightness multiplier values in the range [0,1]. In total, the study used 1240 colour-transformed image pairs; 160 for each original image with 40 for each transform configuration per image with an additional 80 'darkness-transformed' images for each original image.

The study was carried out on the Amazon mechanical turk (AMT) crowd-sourcing platform with participants from the AMT public workforce [125]. The images were presented to the participants in sets of 10, randomly chosen, matched, image pairs, with one of the pairs being a control image pair where one of the two images was swapped with a completely black image.

The participants were asked to rank the images on a 4 point ordinal scale with the following adjectival scores [126]:

- score 3: *Identical*
- score 2: *Minor Difference*
- score 1: *Significant Difference*
- score 0: *Completely Different*

For each pair, the participants were also asked to describe their rationale for the chosen score. Following accepted practice [127], the following criteria were used to verify the responses:

- Verbal rationale provided for all image-pairs
- Score provided for all image-pairs
- Control image pair was correctly scored
- All 10 image pairs were not given the same score (to check that the scorer is not arbitrarily picking the same score)

Responses from participants that failed to match any one of the criteria were rejected.

The study was run till 20,000 image-pair evaluations were obtained. The authors note that the final tally showed that there were more than 400 participants in the study.

This study showed that the *Crayon* system was able to achieve significant power savings while preserving perceived image quality. It also showed that the score is correlated with typical image quality metrics but it alone was not enough to predict the user score. Furthermore, it showed that the perceptual score was correlated with Λ. In comparing *Crayon* to the darkness transforms (i.e., simply making the images darker to reduce the display power consumption), the authors observed that *Crayon* was able to achieve a better user score while reducing the power consumption by the same amount. The authors also observed that, for a given Λ, the perceived quality of different images was *qualitatively* quite different. This means that simply averaging Λ across a large set of images would be a poor way to decide the transform parameter required to achieve a given quality. Finally, the Λ values suggested in the original *Crayon* work are not normalised and would not translate to different display models.

As outlined before, when *Crayon* is being used, a key requirement is to find what the optimal distance metric, colour space, and Λ combination to use to get a required user score and power saving. The aim of the contemporary studies were to formalise the notion of picking these parameters such that a system can dynamically pick colour transform parameters optimised for the image being transformed, or even pre-emptively calculate if its worth transforming the image, based on the properties of the image.

In the sections that follow, we present the design and the results of contemporary user studies aimed at learning how to efficiently pick the *Crayon* transform parameters (in particular Λ).

Overall, the user studies aim to answer the following questions:

1. What properties of the images make them better or worse suited to be transformed with *Crayon*?
2. Would the transformed images be an 'acceptable' replacement for the original?
3. Is the 'acceptability' correlated with the controllable transformation parameters?
4. How can we pick these parameters based on the image to be transformed?
5. How can we normalise the Λ values such that they can be used across different displays?

5.5 Contemporary *Crayon* user studies

As outlined before, we require a display power model for the *Crayon* transforms. In the original study, the authors used an OLED hardware development platform as the test device and measured the power consumption of the built-in display. We instead chose to explore how *Crayon* could be used in a real device. We chose a *Huawei P30 Pro* device as our target platform and measured the variation of the display power over the range of *sRGB* values the device could display; [0 to 255] for each colour channel.

For this, we built an *Android* application to sweep the range of *sRGB* pixel values and measure the display power variation. This was done in a *Binder MK56* thermal chamber with temperature regulated at 25°C to remove effects of ambient temperature variations on the display power consumption (see fig. 5.20). Appendix A provides details of this process along with details of some adjacent explorations we carried out into the power consumption of the OLED displays.

To ensure parity with the original study, we chose to use both ℓ_2 and ℓ_2^2 distance metrics. However, we investigated three colour spaces instead of two; *sRGB*, *CIELAB*, and *CIEUVW*. This gave us six distance metric and colour space configurations.

As with the previous study, this study was also done on the AMT platform with the AMT workers. In the time since the previous study, AMT had been integrated with Amazon *SageMaker*, which operates as a platform for human-in-the-loop machine learning model training [128].

Instead of the three bitmap images used in the original study, we used 80 images from *Unsplash*, which are provided under a CC0-like license (referred to as the Unsplash license) [129].

In the sections that follow, we outline the design details and the results of the studies we carried out into answering the questions posed in section 5.4.

5.6 User Study 1: Internal User Study

We chose to carry out an initial, exploratory study with the members of the *Physical Computation Lab* at the University of Cambridge. This small scale study would allow us to cheaply evaluate the efficacy and the user-friendliness of the study setup.

5.6.1 Experimental Setup

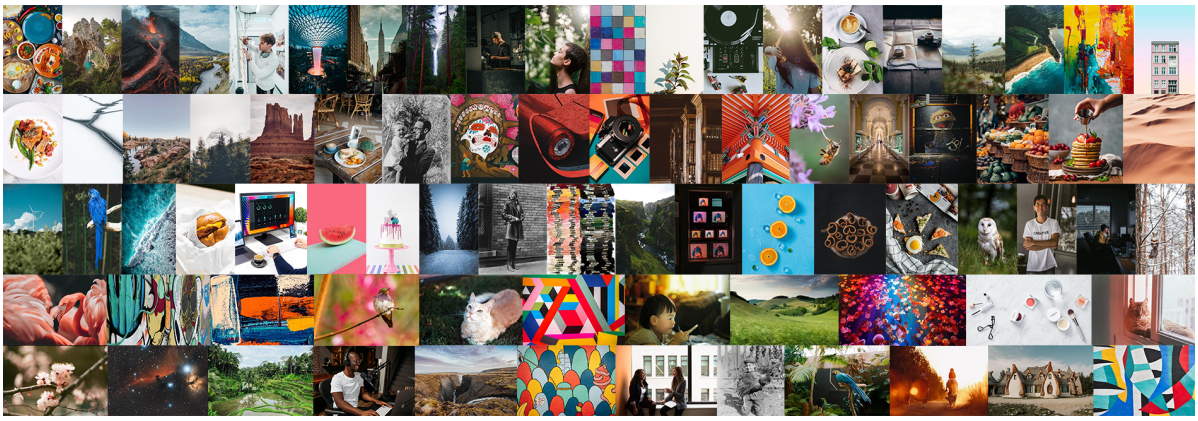


Fig. 5.2 A collage of the images used in the internal study showing the diversity in the images' colour, source, and style.

For the first iteration of the user study, we chose 80 images from *Unsplash*. Figure 5.2 shows a collage of the images used in this study. We ensured that this set of images represented a diverse range of natural and artificial images containing a variety of subjects, and had distinctly different colour palettes. Using these 80 images, we generated the transformed images by randomly sampling an image from the original image set and then picking a random colour space, distance metric, and Λ combination. By repeating this process, we generated a set of 200 transformed images, which we matched with their original images to create a set of 200 image pairs to be presented to the participants.

The participants of the study were to be presented with a set of 11 image pairs from the aforementioned image set. One of the 11 pairs were randomly chosen to be a control pair where the transformed image was replaced with a completely black image.

The participants were asked to answer the question *'Image on the right is an acceptable replacement for the image on the left'* where the image on the right was the transformed image. The answers were required to be given on the following 5-point Likert scale [130, 131]:

- Score 1: Strongly disagree
- Score 2: Disagree
- Score 3: Neither agree nor disagree
- Score 4: Agree
- Score 5: Strongly agree

In addition to the image-scoring, each test also included a version of the *Farnsworth-Munsell 100 Hue colour vision test* (FM100). The aim was to quantify the test participant's colour discriminability. However, we will not discuss the results of this in this chapter as it was meant to be used as an auxiliary measurement and because it was discarded in the subsequent user studies.

5.6.2 User Study Results

As outlined before, we used 200 unique image pairs, transformed using one of two distance metrics and in one of three colour spaces at different Λ values. For each pair, we obtained scores from three different participants, which gave us scores for 600 image pair comparisons. Having multiple scores for each image pair allowed us to average the user scores to get a more accurate estimate of the score.

5.6.3 Score Correlation with Lambda (Λ)

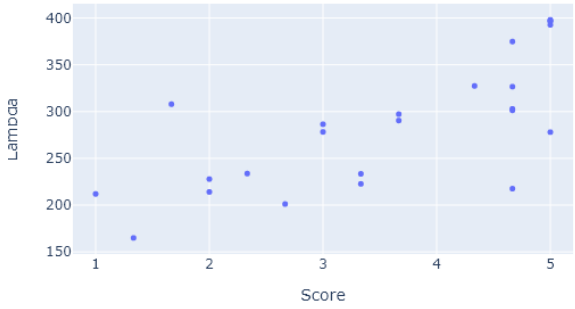
Figure 5.3 shows what Λ is required to achieve a given score for the different distance metrics and colour spaces used in the study. The images transformed using the ℓ_2 distance show the clearest correlation between mean opinion score (MOS) and transform parameter, Λ. This trend is most obvious in figs. 5.3a, 5.3c and 5.3e. In the results for images transformed using the ℓ_2^2 distance, the trend is less obvious. Crucially, the results show that, throughout the Λ range used in the user study, the majority of the user responses were clustered towards the higher scores (4 and 5). On one hand, this implies minimum of the Λ range was not sufficient to create image artefacts substantial enough to make the transformed images unusable, except in few cases (i.e., the transformed images were *acceptable* substitutes for the originals). On the other hand, this highlights the need for more intermediary Λ values such that the user score trends are more clearly visible.

5.6.4 Image Power Saving Correlation with Score

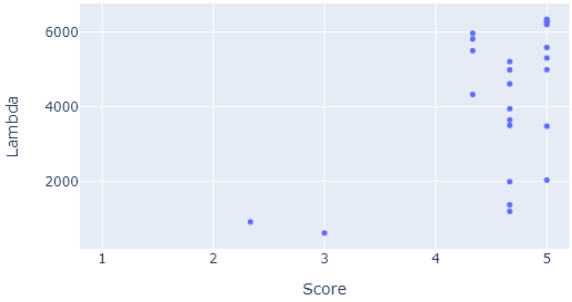
For each image pair, we calculated the power consumption of the image on the display using the previously mentioned display power model. We then calculated the relative power of the transformed image in each pair, with respect to the original image.

The intention was to evaluate how well the power saving introduced by the *Crayon* transform is correlated with user score. All things being equal, there should be a correlation similar to the Λ against user score since image power is correlated with Λ. The expectation was that, by normalising the images pairwise, we would be able to factor out variation of the power due to differences in the structural and chromatic properties of the images, and see a clearer trend.

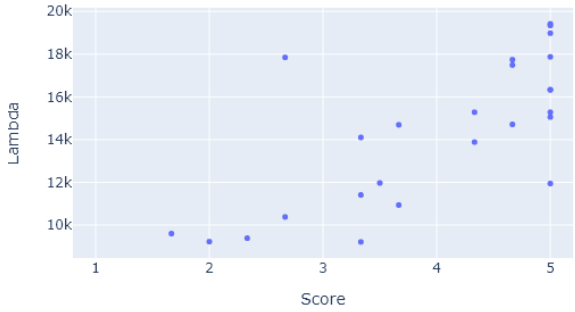
Figure 5.4 show the results of this normalisation. A clear positive correlation was observed for all 6 combinations. However the number of samples available was too small to meaningfully quantify the correlation, especially when using the ℓ_2^2 distance. Qualitatively, *sRGB* and *CIEUVW* colour spaces showed the clearest correlation, with the images using the ℓ_2 distance showing a clearer correlation compared to those transformed with the ℓ_2^2 distance. In most cases, particularly when using ℓ_2 distance, the variance of the power was higher at lower scores.



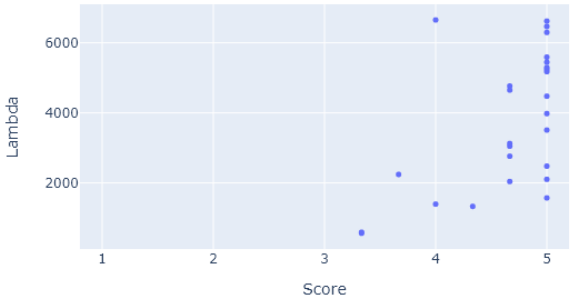
(a) ℓ_2 distance in the *sRGB* colour space.



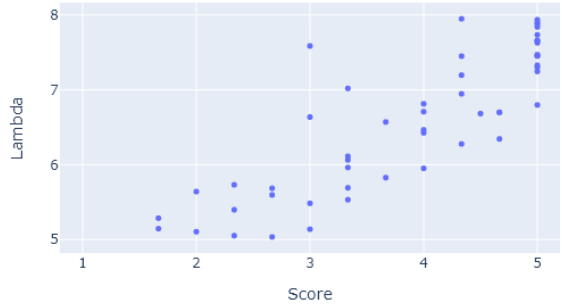
(b) ℓ_2^2 distance in the *sRGB* colour space.



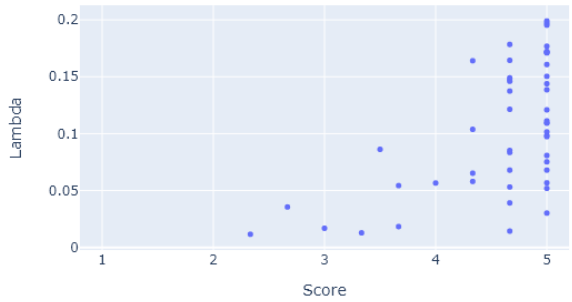
(c) ℓ_2 distance in the *CIELAB* colour space.



(d) ℓ_2^2 distance in the *CIELAB* colour space.



(e) ℓ_2 distance in the *CIEUVW* colour space.

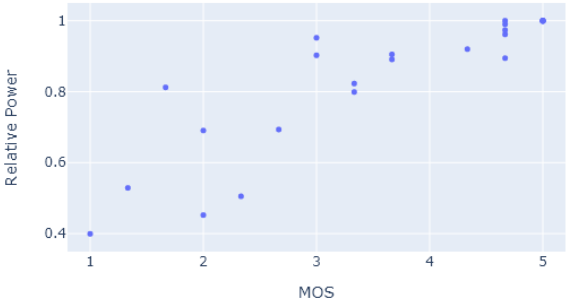


(f) ℓ_2^2 distance in the *CIEUVW* colour space.

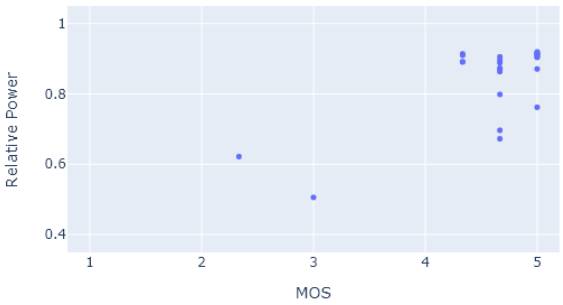
Fig. 5.3 The Λ value required to achieve a given score depends on the distance metric and colour space.

Furthermore, in the images using the ℓ_2^2 distance, the clustering of the data points towards the top right provide further indication that the choice of Λ used in the study was insufficient. We will use this observation later in devising a way to normalise the Λ values.

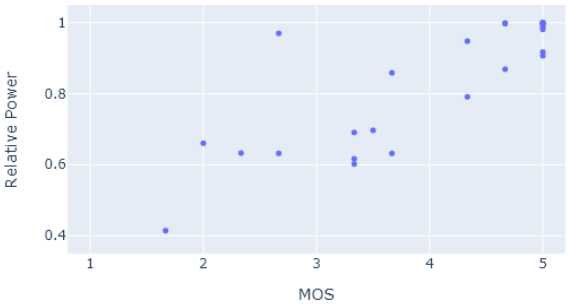
An image with low power and high score implies that the image can tolerate a larger colour approximation without the artefacts adversely affecting the perceptual quality of the image. This is important since a key part of this investigation was to investigate what images perform well under the *Crayon* colour transform, as outlined in section 5.4.



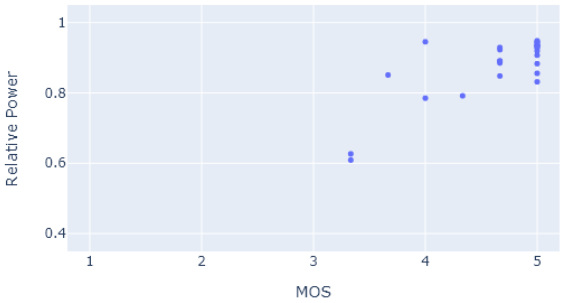
(a) ℓ_2 distance in the *sRGB* colour space.



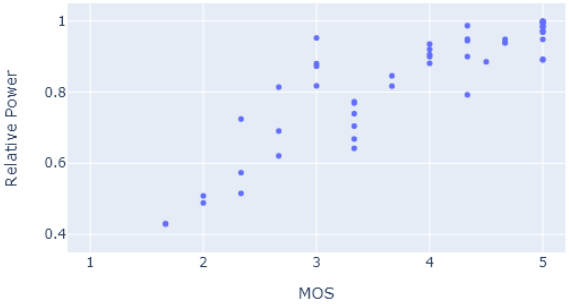
(b) ℓ_2^2 distance in the *sRGB* colour space.



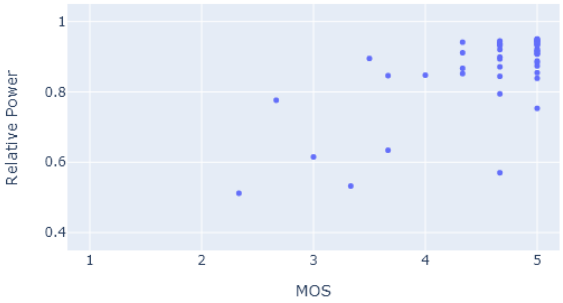
(c) ℓ_2 distance in the *CIELAB* colour space.



(d) ℓ_2^2 distance in the *CIELAB* colour space.



(e) ℓ_2 distance in the *CIEUVW* colour space.

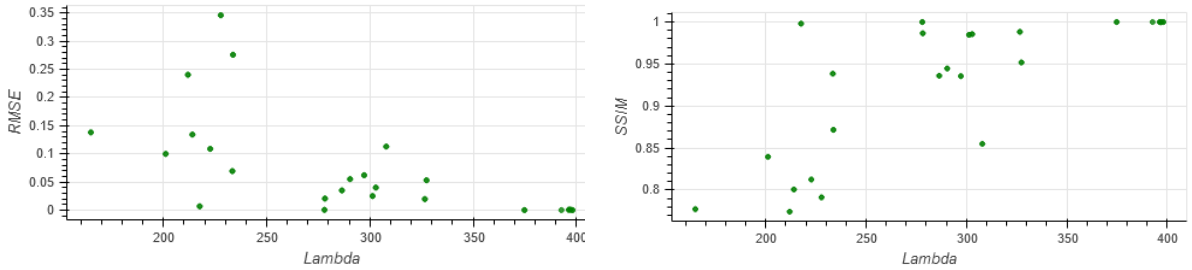


(f) ℓ_2^2 distance in the *CIEUVW* colour space.

Fig. 5.4 The relative power saving is most strongly correlated with user score in the images transformed using the ℓ_2 distance.

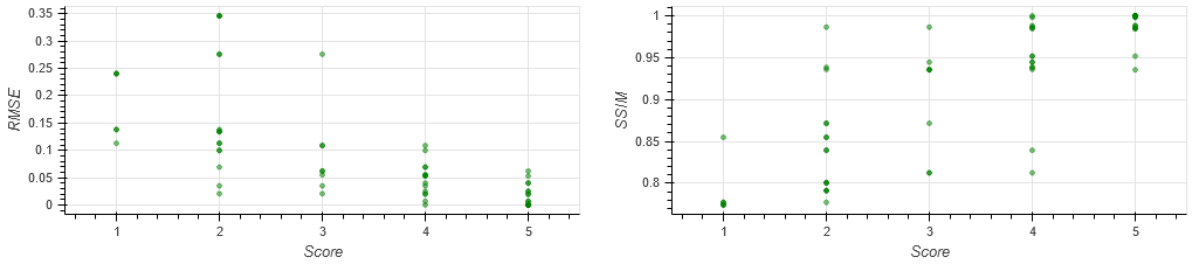
5.6.5 Correlation with Simple 'Image Quality Metrics'

In the original study, the authors compared the user scores with simple image quality metrics (IQMs); namely mean squared error (MSE), root mean squared error (RMSE), and structural similarity index measure (SSIM) (although the former two have been shown to not correlate with perceptual quality [132]). The aim, other than to replicate the results of the original study, was to find a metric that correlates with user score.



(a) RMSE correlation with Lambda (Λ). (b) SSIM correlation with Lambda (Λ).

Fig. 5.5 The correlation of RMSE and SSIM metrics with Lambda (Λ) is heteroscedastic for images transformed using the ℓ_2 distance in *sRGB* colour space.

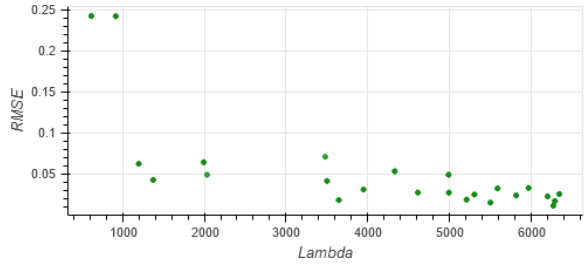


(a) RMSE correlation with user score (b) SSIM correlation with user score

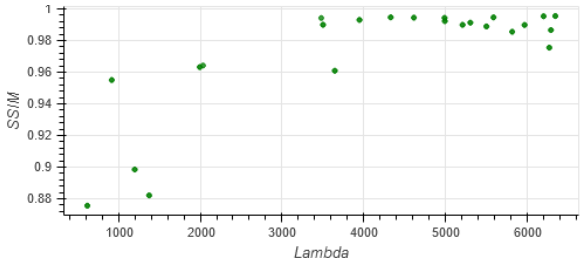
Fig. 5.6 The correlation of RMSE and SSIM metrics with score for images transformed using the ℓ_2 distance in *sRGB* colour space is similar the variation of the IQMs with Lambda (Λ), but discretised.

As was shown in the original *Crayon* work, there was correlation between the simple IQMs and Λ, but the actual variation appears to be heteroscedastic, making it difficult to use a simple model to calculate optimal Λ based on some target IQM value (see fig. 5.5). The variation of IQMs with user score shows a similar, but discretised trend, because the scores take discrete integer values from 1 to 5 (see fig. 5.6)

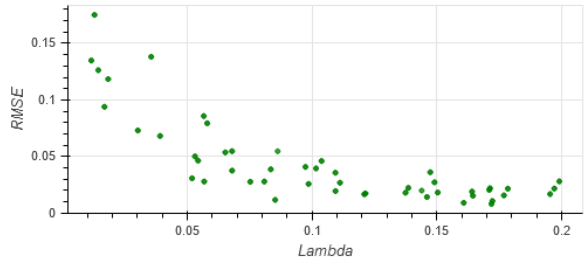
In general, images transformed with the ℓ_2^2 loss show more homoscedastic behaviour when looking at the variation with respect to Λ (see fig. 5.7). This is especially prominent in the results for images transformed in the *CIEUVW* colour space (see figs. 5.7c and 5.7d).



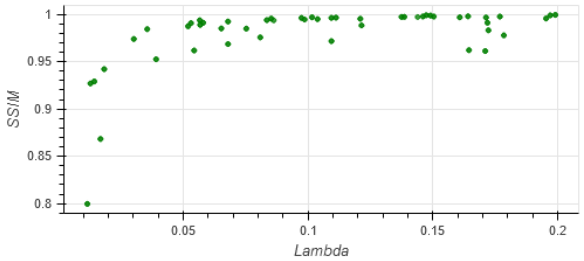
(a) RMSE correlation with Lambda (Λ) in the *sRGB* space.



(b) SSIM correlation with Lambda (Λ) in the *sRGB* space.



(c) RMSE correlation with Lambda (Λ) in the *CIEUVW* space.



(d) SSIM correlation with Lambda (Λ) in the *CIEUVW* space.

Fig. 5.7 The correlation of RMSE and SSIM for images transformed using the ℓ_2^2 distance in the *sRGB* and *CIEUVW* colour spaces show more homoscedastic variation.

5.6.6 Subjective Assessment of the 'Best' and 'Worst' Images

As seen in section 5.6.3, there is a range of MOS observed for a given value of Λ. As MOS averages scores for a given image, this variation is caused by the variation between the properties of the different images. As such, the next step was to investigate which images performed best and worst under the *Crayon* transform. As outlined in section 5.6.4, images that have a low relative power and a high score are images that respond well to the *Crayon* transform. These images are worth analysing to find out why they lead to large power savings without a significant loss in quality. In particular, what features should an image have to allow it to preserve high perceptual quality when transformed. On the contrary, images with high relative power and low score are images that showed significant degradation in image quality without a substantial energy saving. These images are also worth exploring to evaluate what image properties lead to poor performance. This information can be used to decide whether it is worth carrying out the *Crayon* transform on a given image in the first place.

If a heuristic could be found that can identify these images that transform well and the ones that perform poorly, it could be used to triage images such that the *Crayon* system only needs to transform those images that perform well. This would ensure that the overall image quality is maintained at a high standard, while optimising the use of compute resources.

To extract images that have both high MOS and low relative power, the ratio between MOS and relative power, was used as a simple heuristic for sorting images. Figures 5.10 and 5.11 show a collage of the 'best' 9 images and 'worst' 9 images respectively sorted by the aforementioned heuristic, where

'power ratio' is the ratio between the power of the transformed image and the power of the original image.

The first observation that can be made is that, most of the 'best' images were transformed using the ℓ_2^2 distance while most of the 'worst' images used the ℓ_2 distance. This is in line with the expectation that the ℓ_2 distance would favour fewer but larger colour changes compared to ℓ_2^2. It should be expected that this would lead the more apparent artefacts.

The second observation is that, images that are overall darker, could be transformed without many artefacts or significant loss of quality. Furthermore, texture-less white regions lead to more apparent artefacting, especially with the ℓ_2 distance. Transforming images in the *sRGB* colour space led to a shift in the colour temperature of the image, making them look warmer, especially with a smaller Λ (i.e., more approximation).

In addition to visually examining the properties of the images, we also extracted the colour palettes for each image using a median-cut algorithm to cluster the image pixels [133]. Figure 5.8 shows the 15 most common colours in each image. Visually it can be seen that the colour palette of the 'best' images have a marginal bias towards darker colours, but the difference is not stark.



(a) Best images (b) Worst images

Fig. 5.8 Colour swatches for the best and worst images show that the 'best' transformed images show a bias towards darker colours.



Fig. 5.9 Some of users' scores are inconsistent with the observed perceptual quality of images.

It was also apparent that the scoring was quite inconsistent. For example, the image pairs on the top row of fig. 5.10 have been given the same average perception score. Although in our personal opinion, image 2 has far less obtrusive artefacts compared to the others. A similar situation is highlighted in the image pairs shown in fig. 5.9. It is clear that, in making these claims, we are relying on our own biases; having 'normal' colour vision, and making the assessments on a 4K-high dynamic range (HDR) certified display that covers 98% of the *sRGB* colour space and is factory colour calibrated. This would

be beyond the resources available for a typical participant of the study. Given their conditions, the scores they recorded might actually correlate with the perceptions (assuming that they weren't being malicious in recording their scores). However, this brings to the fore the need to sanitise the data more carefully, and to design the user studies to minimise these variations.

In particular, how does one identify if a given observer is reliable? And perhaps just as importantly, can one identify observers that are intentionally sabotaging the test by providing misleading responses. In section 5.6.8 we explore a potential avenue for factoring out observer biases.

The control image was in part supposed to address this by identifying participants that are randomly choosing the options. But when the test is being carried out, the control image pair is prominently out of place and it could be that a 'saboteurs' can easily circumvent the trap by noticing that the control image pair is obviously different to the other images. Although it can be assumed that the members of our research group that took part in the study were not being malicious, they are still vulnerable errors made due to loss of focus, especially after repeatedly completing multiple iterations of the FM100 test.

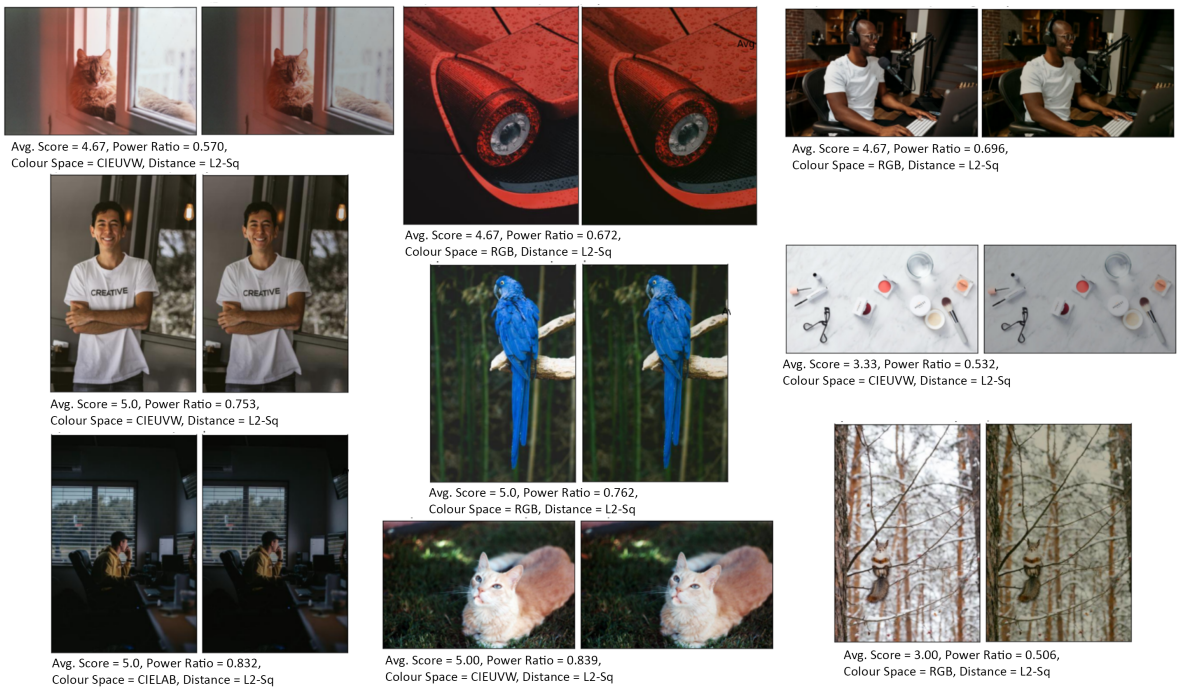


Fig. 5.10 The majority of the 'best' transformed images were transformed using the ℓ_2^2 distance and contain more darker colours.

Issues with the Score Scale

This was further explored with discussions with some of the participants of the study. It was apparent that there was significant variation between the interpretation of the question in answering whether '*Image on the right is an acceptable replacement for the image on the left'* on the 5 point scale. We

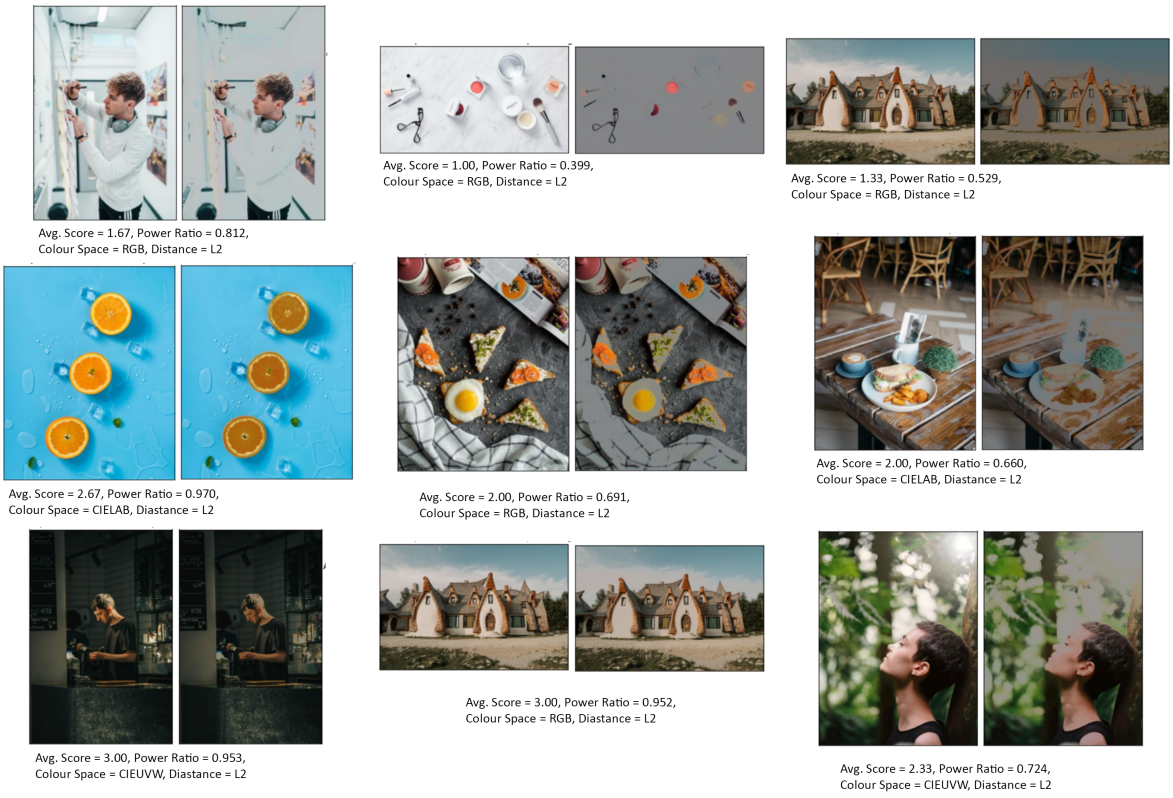
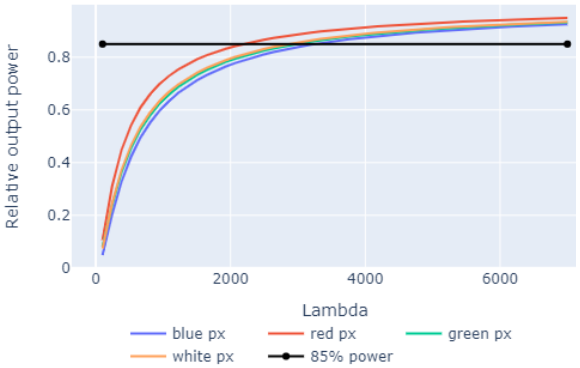


Fig. 5.11 All of the 'worst' transformed images were transformed using the ℓ_2 distance.

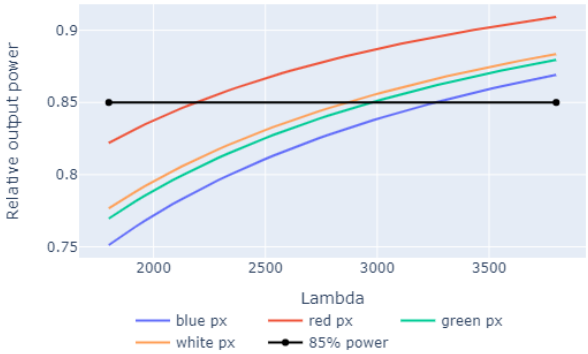
believe that this might be a cause for the apparent discrepancy between the variability of the reported scores.

5.6.7 Sensitivity of Power to Image content

In analysing image power ratio at different Λ values, it can be seen that, for a given Λ value, the relative image power is different for different images. This harkens back to the statement we made before about some images responding better to the *Crayon* transformation.

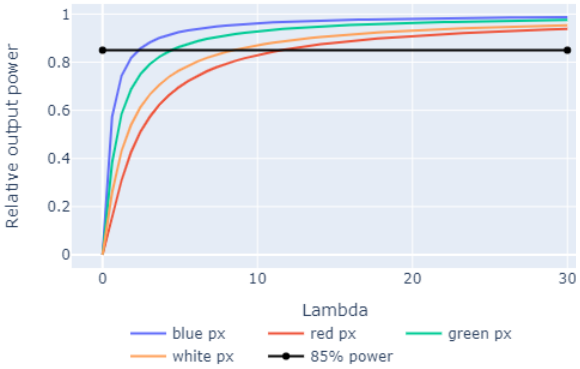


(a) Power sensitivity to Lambda (Λ) depends on the pixel colour but its variation is small with the *Huawei P30 Pro* model.

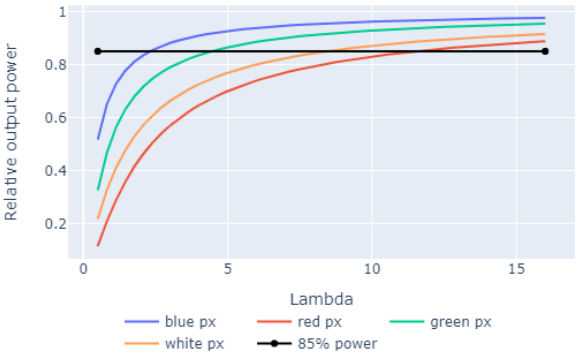


(b) Red pixels are least sensitive to Lambda (Λ).

Fig. 5.12 Variation of the single-pixel power over a range of Lambda (Λ) values with the *Huawei P30 Pro* model. The black horizontal line indicates 85% relative power.



(a) Power sensitivity to Lambda (Λ) depends on the pixel colour and varies significantly with the power model from the original *Crayon* work.



(b) Blue pixels are least sensitive to Lambda (Λ).

Fig. 5.13 Variation of the single-pixel power over a range of Lambda (Λ) values with the model used in the original *Crayon* work. The black horizontal line indicates 85% relative power.

To explore this more clearly, we considered a single-pixel image. As with the user study, we calculated the relative output power of this pixel for a range of Λ values, for four pixel states; White

(1,1,1) and each of R,G,B components at their maximum intensity (e.g max R = (1,0,0), etc.). Using the *Huawei P30 Pro* model, we observed that the 'red' pixel power was least sensitive to variations in Λ (see fig. 5.12a). What this means is that, for an image that has a large number of 'red' pixels, a smaller Λ is required to achieve the same power saving (see fig. 5.12b). However, because the same Λ is used for the entire image, this means that some other parts of the image would end up getting transformed more strongly. As observed before, the MSE and SSIM of the image is correlated with Λ and so is MOS, albeit to a lesser extent. As a consequence, it can be said that using a smaller Λ would lead to more perceptible artefacts in the transformed image.

We repeated the above test with the power model used in the original *Crayon* work. We observed that in this case, it was the 'blue' sub-pixel that was the least sensitive to Λ (see fig. 5.13a). Looking at the power curves for each sub-pixel, it can be seen that, in both cases, it was the sub-pixel with the 'shallowest' curve that was least sensitive to Λ (see figs. 5.14 and 5.15).

Thus, qualitatively, it can be said that the Λ sensitivity is related to the gradient of the power curve, which is a somewhat obvious finding in hindsight. In particular, images that use the display sub-pixel with the smallest gradient more, are less sensitive to Λ. Given that MOS is correlated with Λ, one could expect good quality transformations of images that use the aforementioned small-gradient sub-pixel the least. This will need to be verified with a larger user study with a less noisy set of MOS values.

In summary, it is clear that when transforming different images with a given Λ value, the power saving achieved is not the same. Furthermore, its clear that although MOS is correlated with Λ, the exact value is also dependent on the image content.

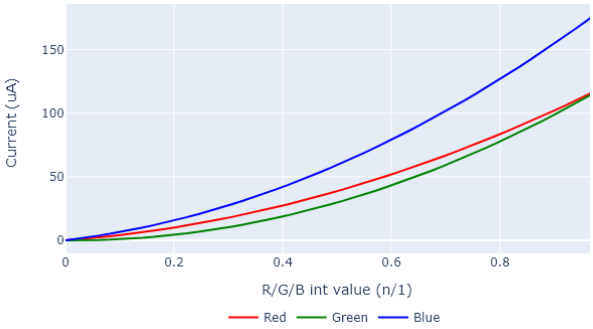


Fig. 5.14 *Huawei P30 Pro* power model.

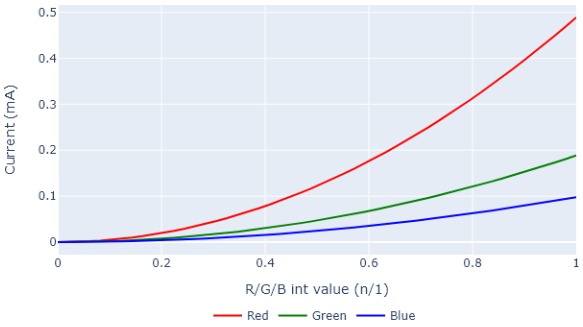
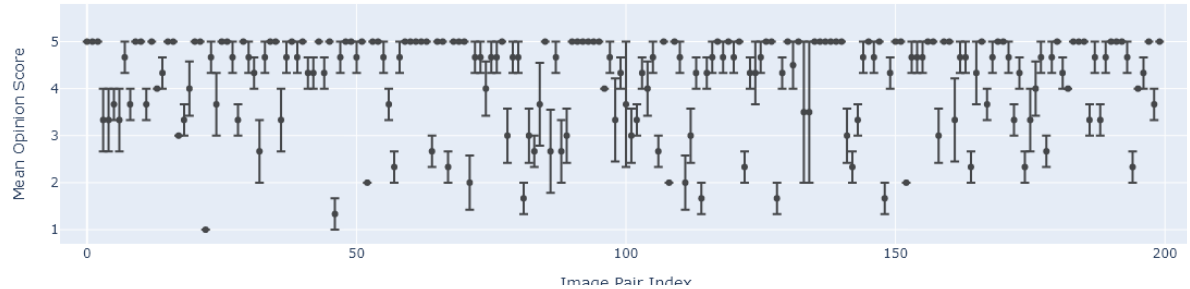


Fig. 5.15 Original *Crayon* power model.

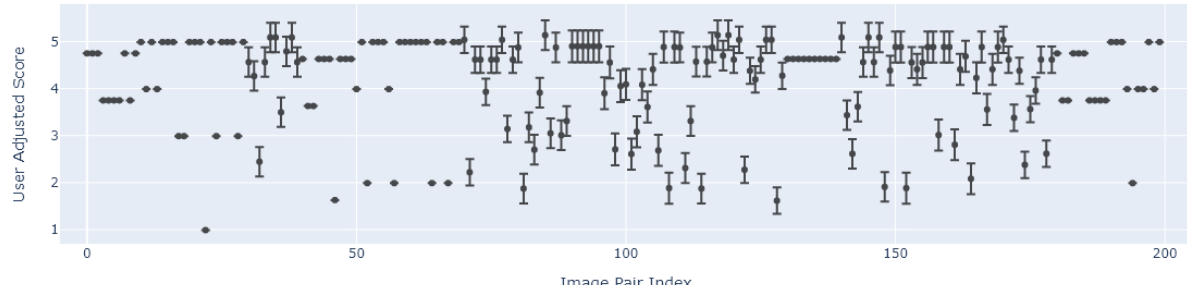
5.6.8 Factoring Out User Bias

As outlined in section 5.6.6, the variability between the observers substantially affects the quality of the user scores. To this end Li et al. has proposed a method that aims to factor out observer bias and inconsistency through a belief propagation and maximum likelihood estimation based method [134]. We used their publicly available implementation, *SUREAL*, to evaluate its efficacy in extracting less noisy opinion scores [135].

Figure 5.16a shows the results for each image pair after calculating MOS with the error bars highlighting one standard deviation from the mean (although with the small number of observations we have, the error bars have questionable utility). Figure 5.16b shows the results of the *SUREAL* algorithm. As can be seen the variance in the observations are substantially lower.



(a) Mean opinion scores (MOS) with error bars showing 1 standard deviation.



(b) User adjusted scores (from *SURREAL*) with error bars showing 1 standard deviation.

Fig. 5.16 The *SUREAL* method aims to factor out inconsistencies and biases imposed by the users and the images to achieve more stable estimates for the user score.

At this initial stage of the user studies, we took the results of from *SUREAL* at face value as evaluating the claims made in their paper is beyond the focus of this work. The results are presented in this section for completeness, repeatability, and to ensure that reader is aware of the post-processing steps we carried out on the data. We used this user adjusted score (UAS) in the following section as the target values, the machine learning models were to predict from Λ and the image properties.

5.6.9 Predicting Score

A key issue with the current formulation of the *Crayon* transform is the need to pick the trade-off parameter Λ. As shown in the sections prior (especially in section 5.6.6), there are visually appreciable properties that seem to predict what images are good at being transformed. Furthermore, there is an observable correlation between the MOS and the Λ value used in the transform.

This stage of the analysis is focussed on using machine learning to see if we can learn a model from the user study data that can predict the user score from the Λ value and other image properties and Λ. The hope was that this model would learn how different image properties affect the user score at a given Λ and allow *Crayon* to pick the Λ value based on the required score.

Given that this model will need to be used in near real-time when used in the 'online' mode (see section 5.2), it was paramount that the model itself be simple to ensure that the amount of energy used to compute the optimal transform parameters do not outweigh the energy saving due to the *Crayon* transform. As such, we looked at simple linear and low order polynomial models as well as support vector machines (SVMs) to model the Λ-score relationship. In addition to that, by using simple models, we would be able to examine model parameters and get an insight into the numerical relationships between the different image properties. However, we also explored the use of gaussian processes (GPs) to model the heteroskedasticity.

Predicting MOS from Lambda (Λ)

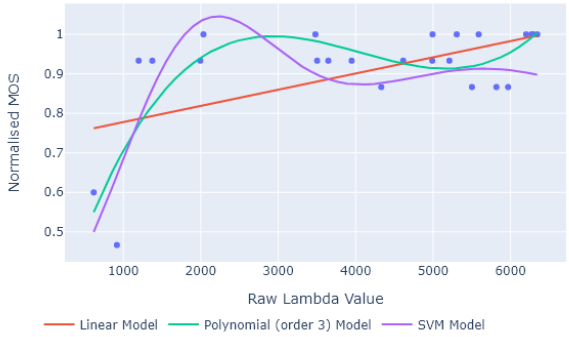
The simplest case is to predict the user score from just Λ. We trained a linear model, a cubic model, and an SVM regressor using the data for each colour space and distance metric combination with 10-fold cross-validation. Figure 5.17 shows the learned trend-lines when using MOS (fig. 5.17a) and UAS (fig. 5.17b) with the results from the images transformed in the *sRGB* colour space using the ℓ_2^2 distance. As can be seen, the line of best fit for the polynomial and SVM models are not monotonic increasing. Given how the *Crayon* transform works, this trend needs to be strictly increasing. Figures 5.17c and 5.17d shows the results of using GPs to model the trend of the data. In this case the same issue with the aforementioned simple models can be observed (i.e., with the predictions not being monotonic increasing with Λ).

Numerically, the models trained on the images transformed with the ℓ_2^2 distance in the *CIELAB* colour space had the lowest MSE across the training folds. Tables B.1 and B.2 in the appendix summarise the results of the models we trained.

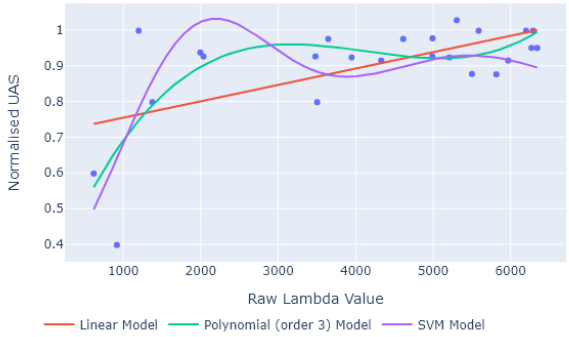
Predicting MOS from Normalised Lambda (Λ)

As a slight modification of the previous setup, we explored the effect of normalising the Λ value on the performance of the learned model. The normalisation was a linear map between the range of Λ used in the study and the interval [0,1], such that the maximum Λ would correspond to 1 and minimum Λ would correspond to 0. The normalisation helped the SVM model achieve a closer fit to the data when comparing to the model trained on raw Λ values. However, we believe that this is a consequence of using the same hyper-parameters for the SVM models (see fig. 5.18). Aside from that, the results were more or less identical to the models using raw Λ values with the key difference being in the values of the learned model parameters.

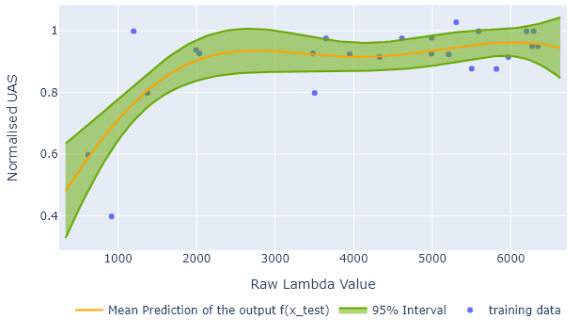
This was motivated by the expectation that normalising Λ across the different colour spaces and distance functions would allow the results from different colour spaces to be considered together giving the models more data to learn with. At this point we did not expect this to be that successful as the initial choice of the Λ range was quite arbitrary, and any normalisation process would be biased because of this. However, in section 5.7.1 we will explore how to normalise the Λ values to allow for this, and also allow for comparisons between different power models.



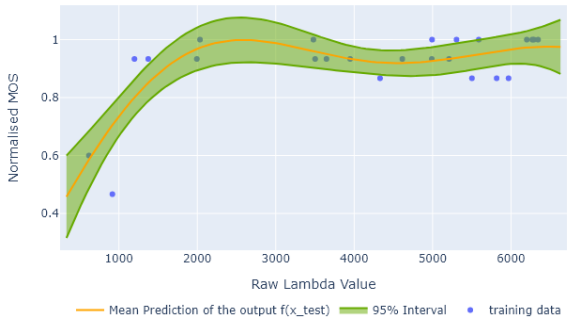
(a) Curve fits for predicting MOS from Lambda (Λ).



(b) Curve fits for predicting UAS from Lambda (Λ).



(c) GP model for predicting MOS from Lambda (Λ).



(d) GP model for predicting UAS from Lambda (Λ).

Fig. 5.17 Simply training models using Lambda (Λ) to predict score led to predictions that were close numerically, but were unrealistic (in the *sRGB* space using ℓ_2^2 distance).

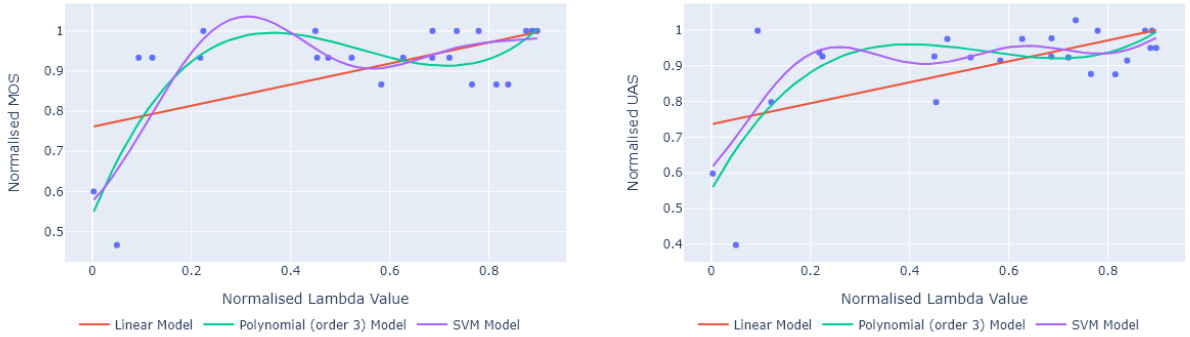
Predicting MOS from Normalised Λ and Image Colour Properties

As outlined before, qualitatively, some properties of images made them more tenable for effective transformation. We explicitly incorporated simple image colour properties like mean and variance of the *hue*, *saturation*, and *value* across the entire image. The results of these models were sub-par even compared to the previous models. This might be due to the limited expressiveness of the simple models explored. Tables B.4 and B.5 details the numerical results achieved from these models.

5.7 User Study 2: Redesigned Public-study

As detailed in the previous section, the first contemporary user study evaluating *Crayon* had the following issues:

1. Un-normalised Λ
2. Very subjective perceptual quality scale
3. User scores were too sparse and too non-uniformly distributed along the range of allowed Λ values
4. Redundant FM100 test increasing the likelihood of fatigue in participants
5. Low number of observations per image pair limiting the insight into observer variability



(a) Curve fits for predicting MOS from normalised Lambda (Λ).

(b) Curve fits for predicting MOS from normalised Lambda (Λ).

Fig. 5.18 Normalising Lambda (Λ) enables the SVM model to achieve better performance but still with a non-monotonic-increasing trend as required.

To mitigate the aforementioned issues and to increase the reliability of the results, the following changes were made when designing the next study:

1. Λ was normalised according to relative display power saving. See section 5.7.1 for details.
2. Re-designed the quality scale to fit to the ITU recommendations as explained in section 5.7.2.
3. Modified the experiment generation script to pick Λ at shorter intervals in the normalised range.
4. Removed the FM100 test from the experiment.
5. Increased the number of participants per experiment setup to 5.
6. A pair of identical images were added as an extra control pair.
7. The pool of base images was restricted from 80 to 14 to increase the number of scores obtained for each base image.
8. Increased the image pairs per experiment from 11 to 20.
9. The participants were asked not to spend more than 20 seconds on each image to emulate real-world observation time.
10. As per ITU recommendations, we used MOS, instead of the UAS.

5.7.1 Lambda (Λ) Normalisation

For each transform configuration, Λ range was chosen such that, given the display power model being used, the maximum Λ would lead to 5% power saving when transforming a pure white image. Similarly, the minimum Λ was chosen such that it would give a 60% power saving for the same configuration and power model. This range was then normalised to $[0, 1]$ with 0 and 1 corresponding to the minimum and maximum Λ values respectively. Table 5.1 summarises the maximum and minimum raw lambda values corresponding to each colour space, distance metric combination.

This was done for each transform configuration independently such that, normalised Λ values would be comparable across the transform configurations. Given that the normalisation process takes

into consideration the display power models being used, the Λ values would also be comparable (and transferable) between different power models (i.e., across different displays).

	sRGB [0,1]	CIELAB [0,1]	CIEUVW [0,255]
ℓ_2	[241,393]	[11350,17610]	[5.425,8.052]
ℓ_2^2	[420,9500]	[310,7200]	[0.00825,0.183]

Table 5.1 Normalised Lambda (Λ) values allow comparison between different display power models.

5.7.2 Adjusted Perceptual Quality Scales

The ITU-R BT.500 enumerates a list of guidelines for subjective quality assessment when carrying out experiments with human participants. In particular, it presents a set of scales recommended to be used in experiments where the participants are asked to compare two stimuli [30]. The BT.500 recommendation presents two *Double-Stimulus* scales; double-stimulus impairment scale (DSIS) and double-stimulus continuous quality-scale (DSCQS). As per the recommendation, DSIS is for evaluating the effect of artefacts caused by a 'system' on the quality of a particular stimulus on which it acts upon. For example, it can be used to quantify the positive effects of an error-correcting data encoding scheme applied to an image before being sent through a noisy channel. In this case, the channel and the encoding scheme in the aggregate will be considered to be the 'system'. The recommended DSIS with the following adjectival scores were incorporated into the new version of the user study.

- Score 5: Imperceptible
- Score 4: Perceptible, but not annoying
- Score 3: Slightly annoying
- Score 2: Annoying
- Score 1: Very annoying

The participants were requested to *'comment on the appearance of artefacts on the transformed images'* using this scale.

5.7.3 Experiment Setup

We chose 14 images from the original 80 to be used in the second study (see fig. 5.19). As before, we used the *Huawei P30 Pro* device as the target platform, reusing the power model used in the first user study. The 14 input images were transformed with each of the original 6 configurations using 20 Λ values, at 5% intervals in the normalised range (as explained in section 5.7.1). We then split the resulting 1,680 unique transformed images into batches of 20 images and paired them up with their originals. Two control image pairs were added to each batch at random locations, with one identical image pair and the other with one of the images set to a black image. These would allow us to validate the participant responses.

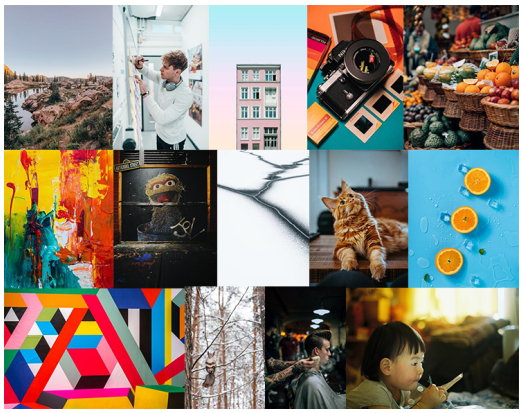


Fig. 5.19 Collage of images used in the second contemporary study to highlight the diversity of images.

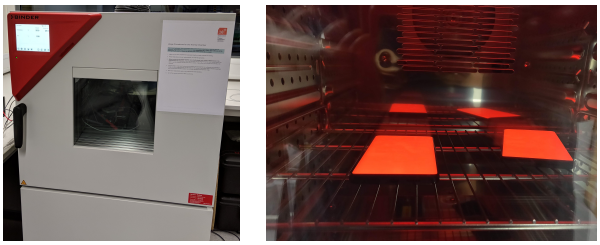


Fig. 5.20 We used a *Binder MK56* thermal chamber (left) to measure display parameters for the *Huawei P30 Pro* devices (inside the chamber, right) at 25°C.



(a) (b) (c) (d) (e) (f)

Fig. 5.21 Only the results from the above subset of images are used in figs. 5.23 and 5.27 to make the graphs of the study results more readable. The images are referenced by the author's last name; (a) Hojo, (b) Reis, (c) Owens, (d) Rajendharkumar, (e) Mu, (f) Johnson.

Each participant was presented with a batch of images and was asked to comment on the visual impairments on the transformed image of the pair using the 5-point DSIS outlined in section 5.7.2 [30]. To ensure that the image observation time is similar to a real-world use case, the participants were also asked not to spend more than 20 seconds on each image pair. We ran the study till we had responses from five different participants for each image batch.

5.7.4 User Study Results

The study resulted in 8,400 image pair comparisons from 62 unique participants. We discarded all the batches where either of the control image pairs were scored incorrectly. A batch was also rejected if the identical pair had a score less than 4 or if the pair with the black image had a score greater than 2. We then removed control images from the remaining batches. After this initial cull, 6,800 image pairs were left. Figure 5.24 shows the distribution of the raw scores as a function of the normalised Λ value for both ℓ_2^2 and ℓ_2 distances.



(a) Original image. Image power = 100.00%.
(b) $\Lambda = 1.0$. Image power = 95.86%.
(c) $\Lambda = 0.5$. Image power = 92.64%.
(d) $\Lambda = 0.01$. Image power = 54.19%.

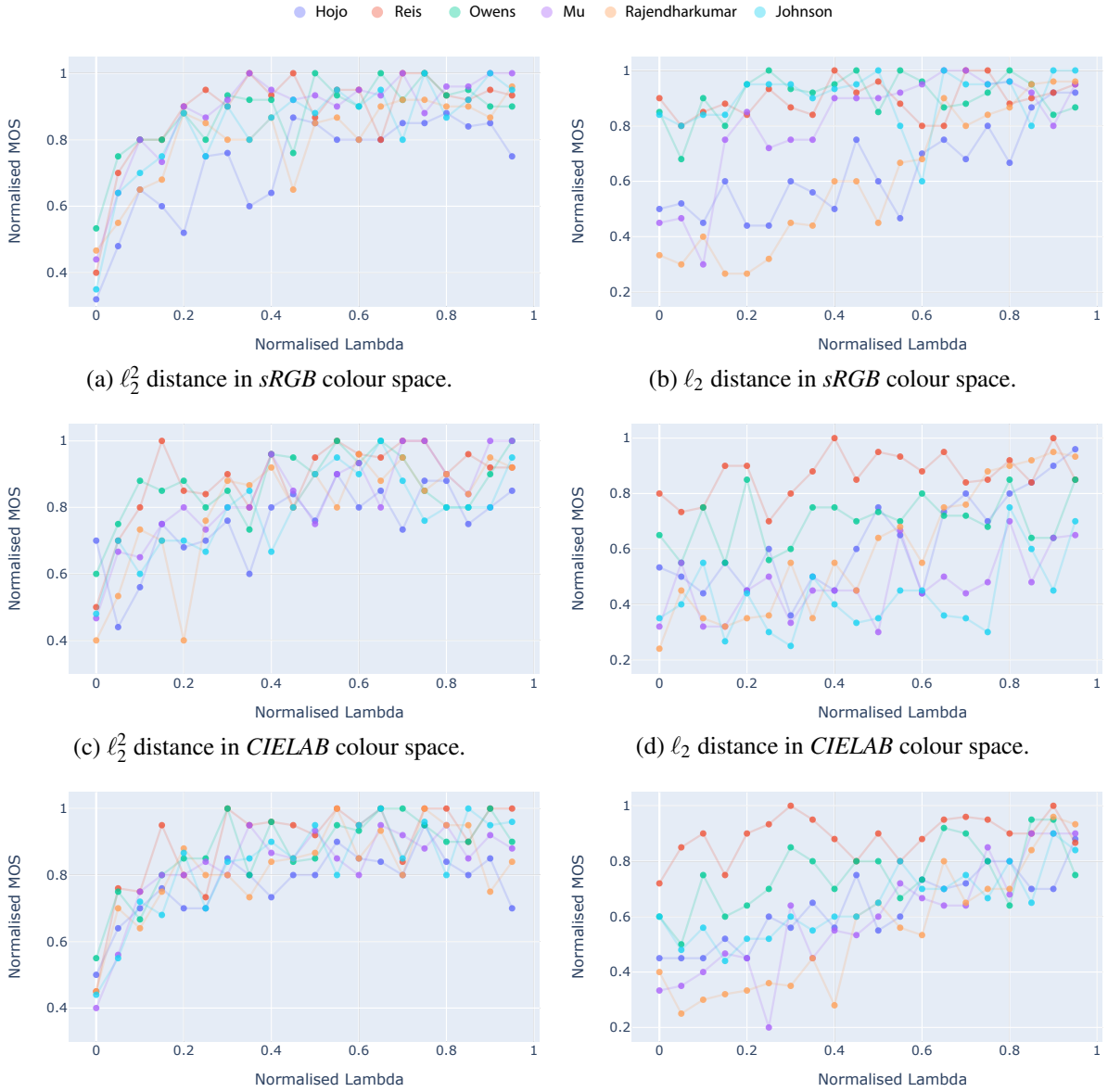
Fig. 5.22 The strength of the *Crayon* transform can be controlled by the parameter Lambda (Λ). Lower Λ achieves improved power savings at the expense of image quality. Images were transformed in the RGB colour space using the ℓ_2^2 distance.

We observed that the proportion of images with scores of 4 and 5 drop of more rapidly when using the ℓ_2 distance when compared to the results of the images transformed with the ℓ_2^2 distance. This implies that the ℓ_2^2 distance is better at preserving perceptual quality across a larger range of Λ. Since Λ is normalised using power savings, this means a higher power saving can be achieved with the ℓ_2^2 distance while preserving image quality. This matches the observations made in the first contemporary study. However, as the Λ values used in that study were not normalised, it is difficult to directly compare them with the results from this study. Notably, a majority of the participants reported that the artefacts were *'Imperceptible'* or *'Perceptible, but not annoying'* until a Λ value of 0.1 with the ℓ_2^2 distance (approximately 50% power saving). This is in-line with the results of the previous *Crayon* study. However, compared to the original work, this result is for a corpus of images five times bigger.

Figure 5.23 shows the normalised MOS against Λ for the images in fig. 5.21. The results show that when using the ℓ_2^2 distance, the change in MOS with the change in Λ follows a similar trend, across all the images. This is in contrast to the results when using the ℓ_2 distance.

5.8 Predicting Optimal Transform parameters

In comparison to the previous study, the smaller image-set and the increased number of user scores led to more stable MOS estimates for all the experiment configurations. Although this increased the accuracy of the models learned through methodologies outlined in section 5.6.9, the inherent variation observed between the images due to the effect of latent image properties would have placed an upper bound on the accuracy given the simple models being used and have limited their usefulness. Having a high numerical accuracy is does not necessarily imply that the predictions of a model are meaningful.



(a) ℓ_2^2 distance in *sRGB* colour space.

(b) ℓ_2 distance in *sRGB* colour space.

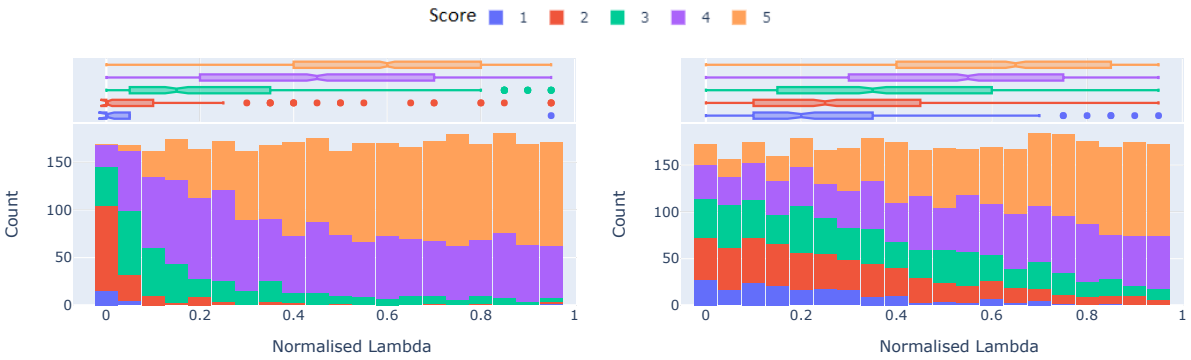
(c) ℓ_2^2 distance in *CIELAB* colour space.

(d) ℓ_2 distance in *CIELAB* colour space.

(e) ℓ_2^2 distance in *CIEUVW* colour space.

(f) ℓ_2 distance in *CIEUVW* colour space.

Fig. 5.23 MOS variation with Lambda (Λ) for different images follow a similar trend when using ℓ_2^2 distance (figs. 5.23a, 5.23c and 5.23e) in contrast to ℓ_2 distance (figs. 5.23b, 5.23d and 5.23f) in all the colour spaces (*sRGB* (figs. 5.23a and 5.23b), *CIELAB* (figs. 5.23c and 5.23d), and *CIEUVW* (figs. 5.23e and 5.23f)).



(a) ℓ_2^2 distance (b) ℓ_2 distance

Fig. 5.24 Image quality is preserved for a larger range of Lambda (Λ) when using the ℓ_2^2 distance (a) in contrast to ℓ_2 distance (b).

As enumerated before, the usefulness of the *Crayon* system is limited by the requirement that the transform parameter Λ needs to be chosen based on the required power saving and the perceptual quality. As the aim of *Crayon* is to be used in energy-constrained systems, it is paramount that this calculation can be done in an energy-efficient manner which places a limit on the tolerable complexity of the machine learning models.

Looking at the MOS variation with Λ in fig. 5.23a, it is evident that the trend-lines for all the images follow a similar shape. The following sections explore how to use this observation to efficiently calculate the optimal Λ value (through a lower bound) given a required user score.

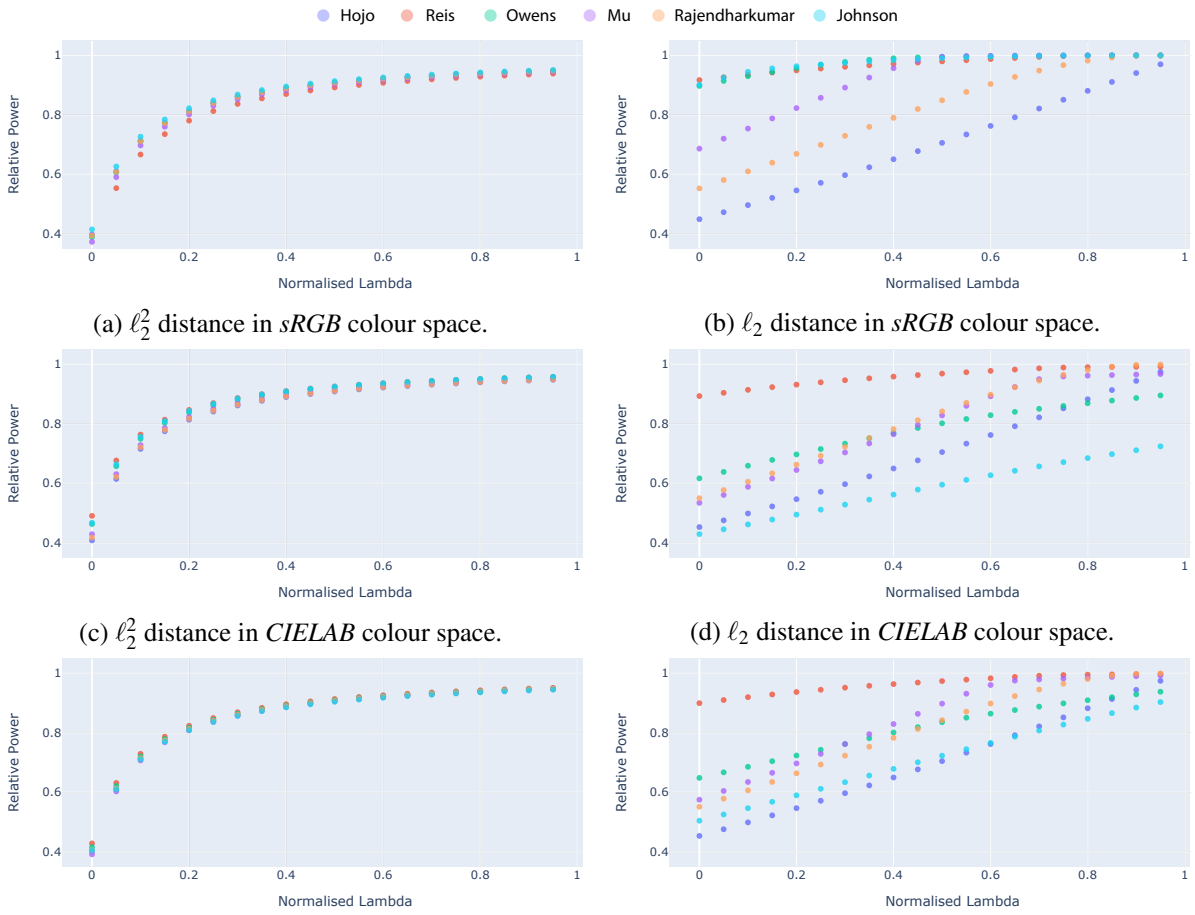
5.8.1 Lambda (Λ) Lower Bound Calculation

The upper bound for the curves shown in fig. 5.23a represent a Λ lower bound. This is the smallest value of Λ that has been observed to produce a given perceptual score. Qualitatively, one can see that the function of Λ against user scores is somewhat monotonic. The points on the lower bound can be extracted by algorithm 1.

Algorithm 1: Algorithm for finding the points on the lower bound.

Data: D = Pairs of Score and Λ, sorted by Λ
Result: LB= Array of data points on the lower bound

$L_{min}, S_{min} \leftarrow D[0]$;
$LB[0] \leftarrow ScoreLambdaPairs[0]$;
forall $i, (L,S) in ScoreLambdaPairs$ **do**
    if $L > L_{min}$ *and* $S > S_{min}$ **then**
        $LB \leftarrow (L,S)$;
        $L_{min} \leftarrow L$;
        $S_{min} \leftarrow S$;
return LB



(a) ℓ_2^2 distance in *sRGB* colour space.

(b) ℓ_2 distance in *sRGB* colour space.

(c) ℓ_2^2 distance in *CIELAB* colour space.

(d) ℓ_2 distance in *CIELAB* colour space.

(e) ℓ_2^2 distance in *CIEUVW* colour space.

(f) ℓ_2 distance in *CIEUVW* colour space.

Fig. 5.25 The relative display power variation with Lambda (Λ) for different images also follow a similar trend when using ℓ_2^2 distance (figs. 5.25a, 5.25c and 5.25e) in contrast to ℓ_2 distance (figs. 5.25b, 5.25d and 5.25f) in all the colour spaces (*sRGB* (figs. 5.25a and 5.25b), *CIELAB* (figs. 5.25c and 5.25d), and *CIEUVW* (figs. 5.25e and 5.25f)).

These data points were then fitted to the following exponential function by adjusting the rate parameter k,

$$\Lambda_{\mathrm{LB}}(s) = \frac{e^{ks} - 1}{1000} \tag{5.16}$$

This gives an analytic function describing the lower bound, which itself is described using a single parameter k. We carried out this process of finding the points on the lower bounds and fitting the exponential for all 14 images used in the study with each transform configuration. Figure 5.27 shows the results for the 6 images in fig. 5.21.

Thus, the shape of the Λ lower bound can be expressed by a single parameter k, which encapsulates the information provided by the individual user scores about the effect of Λ on the user score. Each image and colour space will have its own k value and thus the process of finding a model for user score, given Λ, has been reduced to a problem of finding k, given an image and the transform colour space.

Predicting the Lower Bound from Image Heuristics

Although the process of algorithmically finding the lower bound allows an educated judgement to be made about the choice of Λ given a required MOS, the lower bounds for the different images vary drastically (see fig. 5.27). Furthermore, the process of explicitly finding the exponential fit using known lower bound points only works for the images used in the study where the user responses are readily available. In practice, it is required to calculate this lower bound for a previously-unseen image.

Similar to what is outlined in section 5.6.9, the problem is now finding k given an image and the transform colour space. However in contrast to what was done in section 5.6.9, the problem is much simpler and the model space is restricted to the family of exponentials described by eq. (5.16) for varying values of k. This strong prior would ensure that the learned model would have the properties we expect from the user score (e.g., a monotonic increasing trend). As an added benefit, because the rate parameter k dictates the rate at which the exponential curve grows, the parameter is a good proxy for how well a given image performs under the *Crayon* transformation.

Given our aim of energy (and computational) efficiency, the following simple heuristics were evaluated as proxies for image properties to explore whether they correlate with k.

1. ***Luminance* Mean Threshold** *(Lum.Mean.Thr.)*: Split the image into 51x51 pixel blocks and count the number of blocks with mean *luminance* > 0.8.
2. ***Luminance* Std. Deviation Threshold** *(Lum.Std.Thr.)*: Split the image into 51x51 pixel blocks and count the number of blocks with standard deviation of the *luminance* > 0.02.
3. **Mean *Value*** *(Mean Val.)*: Mean *value* across the image, calculated in the HSV colour space.
4. **Mean *Saturation*** *(Mean Sat.)*: Mean *saturation* across the image, calculated in the HSV colour space.
5. **Std. Deviation of *Value*** *(Std.Val.)*: Standard deviation of the *luminance* across the image, calculated in the HSV colour space.
6. **Std. Deviation of *Saturation*** *(Std.Sat.)*: Standard deviation of the *saturation* across the image, calculated in the HSV colour space
7. **Std. Deviation of Hue** *(Std.Hue.)*: Standard deviation of the *Hue* across the image, calculated in the HSV colour space
8. **Mean *Luminance* (in *CIELAB*)** *(Mean Lum.Lab)*: Mean *luminance* across the image, calculated in the *CIELAB* colour space
9. **Ratio between the Mean *Value* and the Std. Deviation of the *Value* in HSV** *(Mean Std. Val. ratio)*: Mean *luminance* divided by the standard deviation of the *luminance*, calculated in the HSV colour space.
10. **Felzenszwalb Clusters** *(Flz.)*: Cluster the image using Felzenszwalb Clustering [136] with scale=500, σ=1, and count the number of clusters. Normalise the number of clusters by the image resolution and scale by 1000.

11. **Blocks with Low *Saturation* and High *Value*** *(L.Sat.H.Val.)*: Split the image into 13x13 pixel blocks and count the number of blocks with mean *saturation* < 0.3 and mean *value* > 0.7, calculated in the HSV colour space.
12. **Blocks with Low *Saturation* and High *Value* weighted by Std. deviation of value** *(Weighted L.Sat.H.Val.)*: as above but weighted by the standard deviation of the *value* across the image.

The exact values for the heuristic parameters (e.g., block size, threshold values, etc.) were found through manual experimentation. For each heuristic, the correlation between the heuristic's value and the lower bound shape parameter k was calculated. Figure 5.26 shows the Pearson correlation coefficients and the Spearman rank-correlation coefficients for each heuristic with k, along with the p-value being indicated by the error bars. Table 5.2 summarises these results numerically.

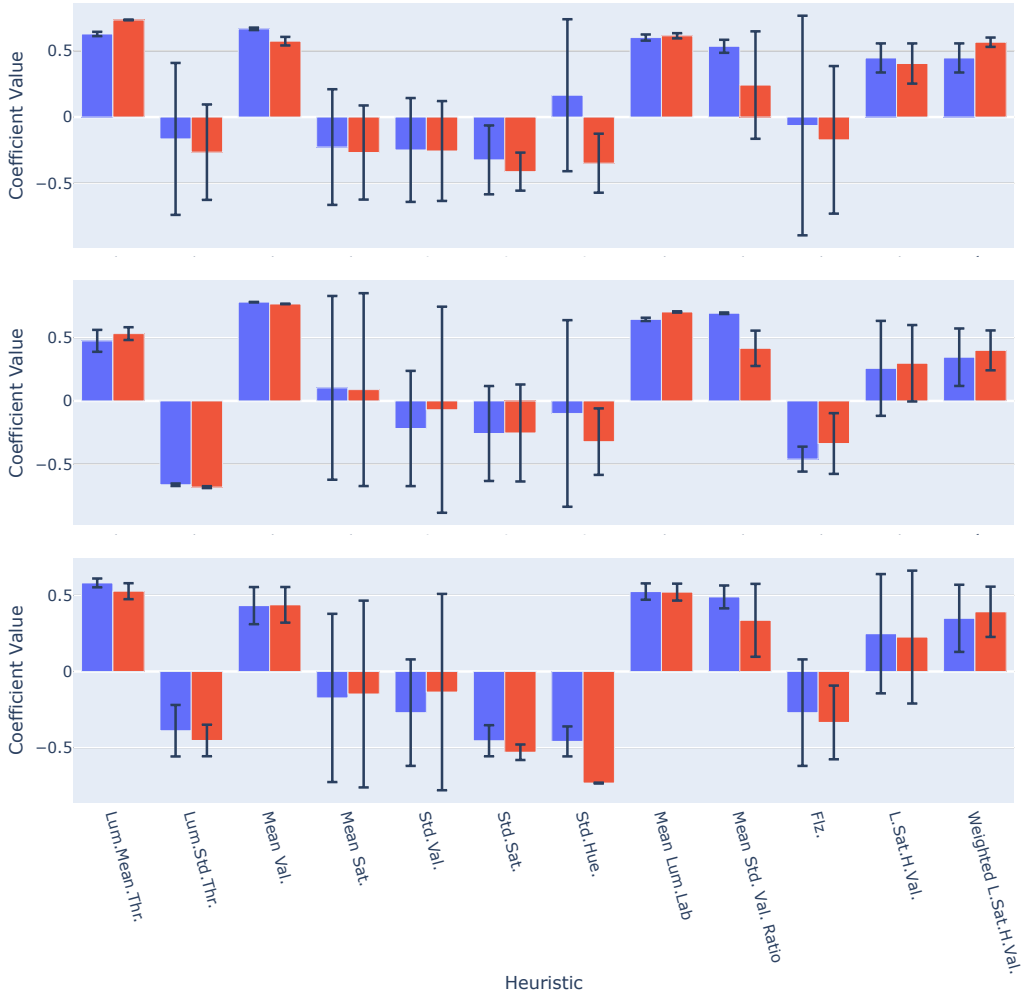


Fig. 5.26 Correlation coefficients for different heuristics (ℓ_2^2 distance). Top to bottom, *sRGB*, *CIELAB*, *CIEUVW*.

As can be seen in table 5.2, the mean *luminance* threshold (*Lum.Mean.Thr.*) shows the largest Pearson correlation with k (0.733 with p<0.01 in the *sRGB* colour space). This is followed by the mean luminance heuristic in the *CIELAB* space (*Mean Lum.Lab*) and the mean value heuristic in the HSV colour space (*Mean Val.*). However, this is as expected, as all three heuristics encode information about the brightness of the image. In addition to this, standard deviation of saturation and hue across the image were also somewhat correlated with k.

Based on this we chose to use the mean *value*, and the standard deviation of *hue*, *saturation*, and *value* as the image features to use in predicting k. At the interest of keeping the number of tunable parameters low, *mean value* was chosen instead of the *mean luminance threshold*, as it has a similar correlation to *mean luminance threshold* across the three colour spaces.

Heuristic	**Correlation**	*sRGB*	*CIELAB*	*CIEUVW*
Lum.Mean.Thr.	Pearson	0.733 (p=0.003)	0.531 (p=0.050)	0.528 (p=0.052)
Lum.Mean.Thr.	Spearman	0.627 (p=0.016)	0.474 (p=0.087)	0.582 (p=0.029)
Lum.Std.Thr.	Pearson	-0.265 (p=0.360)	-0.680 (p=0.007)	-0.453 (p=0.104)
Lum.Std.Thr.	Spearman	-0.165 (p=0.573)	-0.662 (p=0.010)	-0.389 (p=0.169)
Mean Val.	Pearson	0.573 (p=0.032)	0.766 (p=0.001)	0.438 (p=0.117)
Mean Val.	Spearman	0.666 (p=0.009)	0.780 (p=0.001)	0.433 (p=0.122)
Mean Sat.	Pearson	-0.267 (p=0.355)	0.089 (p=0.761)	-0.148 (p=0.614)
Mean Sat.	Spearman	-0.226 (p=0.436)	0.103 (p=0.725)	-0.174 (p=0.553)
Std.Val.	Pearson	-0.256 (p=0.377)	-0.069 (p=0.813)	-0.135 (p=0.645)
Std.Val.	Spearman	-0.248 (p=0.392)	-0.218 (p=0.455)	-0.270 (p=0.350)
Std.Sat.	Pearson	-0.412 (p=0.144)	-0.253 (p=0.383)	-0.531 (p=0.051)
Std.Sat.	Spearman	-0.323 (p=0.260)	-0.257 (p=0.375)	-0.455 (p=0.102)
Std.Hue.	Pearson	-0.348 (p=0.223)	-0.321 (p=0.262)	-0.733 (p=0.003)
Std.Hue.	Spearman	0.165 (p=0.573)	-0.099 (p=0.737)	-0.459 (p=0.098)
Mean Lum.Lab	Pearson	0.614 (p=0.020)	0.702 (p=0.005)	0.522 (p=0.056)
Mean Lum.Lab	Spearman	0.600 (p=0.023)	0.644 (p=0.013)	0.525 (p=0.054)
Mean Std. Val. Ratio	Pearson	0.242 (p=0.405)	0.415 (p=0.140)	0.336 (p=0.240)
Mean Std. Val. Ratio	Spearman	0.534 (p=0.049)	0.692 (p=0.006)	0.490 (p=0.075)
Flz.	Pearson	-0.172 (p=0.557)	-0.336 (p=0.240)	-0.335 (p=0.242)
Flz.	Spearman	-0.064 (p=0.829)	-0.459 (p=0.098)	-0.270 (p=0.350)
L.Sat.H.Val.	Pearson	0.404 (p=0.152)	0.297 (p=0.302)	0.226 (p=0.437)
L.Sat.H.Val.	Spearman	0.446 (p=0.110)	0.257 (p=0.375)	0.248 (p=0.392)
Weighted L.Sat.H.Val.	Pearson	0.565 (p=0.035)	0.399 (p=0.157)	0.392 (p=0.165)
Weighted L.Sat.H.Val.	Spearman	0.446 (p=0.110)	0.345 (p=0.227)	0.349 (p=0.221)

Table 5.2 Different heuristics are correlated to varying degrees with the parameter k.

Training Models to Predict the Rate Parameter; k

We experimented with using linear and cubic models as well as SVMs with radial basis function (RBF) kernels for predicting k from these features. Table 5.3 summarises the results for these models outlining

Colour Space	Model	MSE	Var.	% error
sRGB	Linear	0.7499	0.3619	38.325
	Cubic	0.5701	0.1925	29.134
	SVM	**0.2989**	**0.0782**	**15.277**
CIELAB	**Linear**	**0.1436**	**0.0047**	**7.114**
	Cubic	1.8258	4.4521	90.457
	SVM	0.1753	0.0107	8.682
CIEUVW	**Linear**	**0.1835**	**0.0135**	**9.955**
	Cubic	0.3691	0.0692	20.020
	SVM	0.2855	0.0168	15.488

Table 5.3 The model is able calculate the lower bound parameter k to a mean percentage MSE of 7%, with 5-fold cross-validation.

Colour Space	Model	MSE	% error
sRGB	SVM	0.0801	4.091
CIELAB	SVM	0.1753	8.686
CIEUVW	SVM	0.0295	1.601

Table 5.4 When using 13 of the 14 original images, the SVM model is able to further reduce the error to 1.6% showing that the accuracy could be increased with more data.

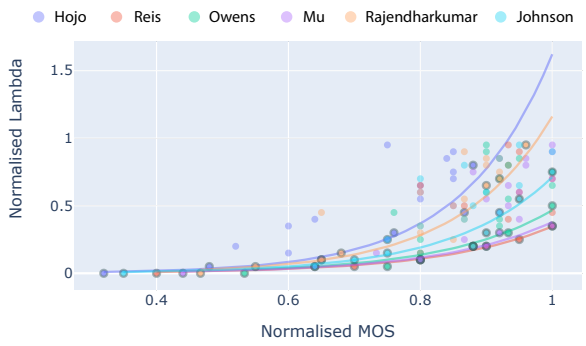
the average MSE and MSE variance achieved for each colour space, with 5-fold cross-validation. The MSE is defined as the difference between the predicted k value and k value calculated using the user study results. We used the observed range of k for each colour space to calculate the percentage error.

The results show, that this model is able to calculate k to a minimum mean error of 7% over 5 cross-validation folds. It can be seen that for both *CIELAB* and *CIEUVW* spaces, a simple linear model outperforms all other model architectures. We believe that this is due to the way those two colour spaces are defined. As outlined before, the *CIELAB* and *CIEUVW* aim to linearise the perceptual difference between colours. This allows, a Euclidean distance metric to be used on these spaces to measure the differences between colours, unlike the *sRGB* colour space.

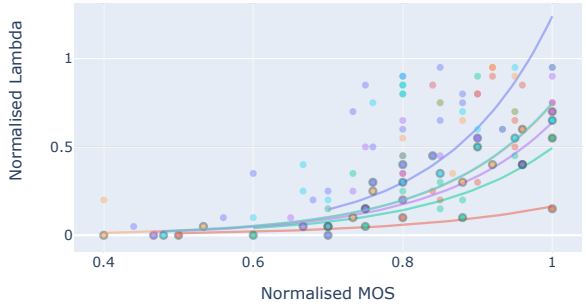
We trained a further set of models using the results from all the images except the image in fig. 5.21b. Table 5.4 shows the results for these models. Figure 5.28 shows the lower bound prediction for the image in fig. 5.21b alongside the calculated lower bound. These models show a decreased mean error of 1.6% MSE. This shows that our problem formulation is flexible enough to improve the model performance beyond what is shown here if a larger dataset containing user scores and corresponding images was provided. Additionally the results show that the model is able to avoid fitting to the outlier point at MOS = 1 in the *CIELAB* colour space (fig. 5.28b). This indicates that the learned model has generalised well and is not just 'remembering' the observations.

5.9 Applying the modified Crayon framework

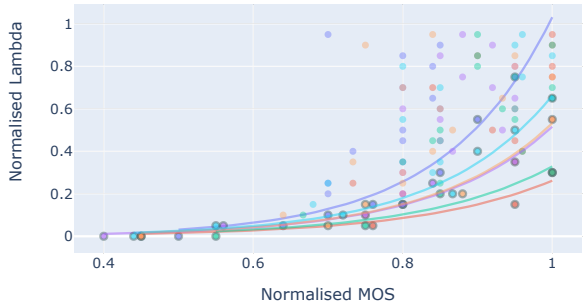
Given the findings presented in the preceding chapters, we can now outline how they can be used in conjunction with the *Crayon* transform. As before, we start with an image to be transformed. In the original framework, the transform parameter Λ is specified somewhat arbitrarily. However, in this work we showed how the transform parameter could be calculated based on the required user score (i.e., the required image quality). Thus, in the modified framework, the user-agent (human or software) specifies the minimum required image quality as a normalised MOS in the interval [0, 1]. Additionally, the user-agent can specify a lower bound for the power saving (e.g., "reduce the display power by at least



(a) ℓ_2^2 *distance in sRGB colour space.*

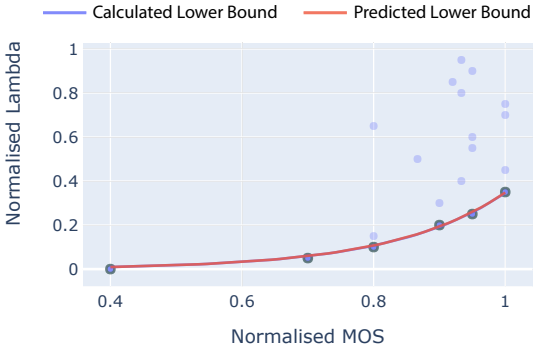


(b) ℓ_2^2 *distance in CIELAB colour space.*

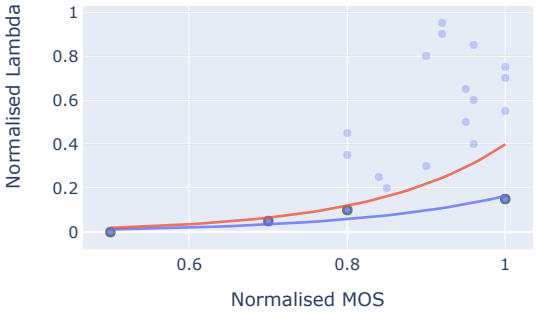


(c) ℓ_2^2 *distance in CIEUVW colour space.*

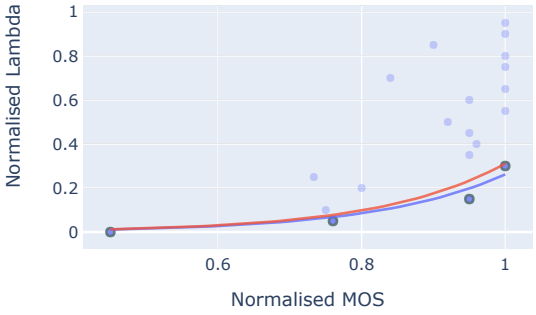
Fig. 5.27 The exponential model can fit the lower bound for Lambda (Λ) against MOS when using ℓ_2^2 distance for all three colour spaces; *sRGB* (fig. 5.27a), *CIELAB* (fig. 5.27b), and *CIEUVW* (fig. 5.27c).



(a) ℓ_2^2 *distance in sRGB colour space.*



(b) ℓ_2^2 *distance in CIELAB colour space.*



(c) ℓ_2^2 *distance in CIEUVW colour space.*

Fig. 5.28 Predicted lower bounds for the image in fig. 5.21b using the *leave-out* SVM model follow the trend of the data points better than the calculated fit for all three colour spaces; *sRGB* (fig. 5.28a), *CIELAB* (fig. 5.28b), and *CIEUVW* (fig. 5.28c).

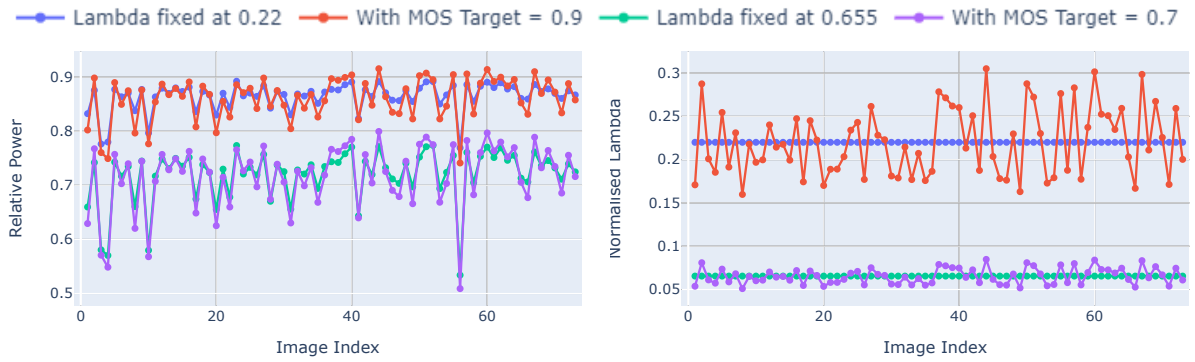
10% ($P_{max} = 0.9$)"). The user-agent can also specify a colour space, but by default we would opt to use the *RGB* colour space as CIE-LAB and CIE-UVW colour spaces require extra computation steps to convert the RGB pixel values to and from their respective colour coordinates. Unlike with the original *Crayon* framework, the modified framework can only be used with the ℓ_2^2 distance.

The first step in this framework is to calculate the values of the four image heuristics. These values are fed into the trained rate parameter model to calculate the parameter k. This gives us the shape of the lower bound of Λ as a function of the MOS. By using the specified minimum image quality as an input to this function, we can obtain the minimum Λ that can achieve this score (Λ_{min}). Since Λ is inversely correlated with display power, Λ_{min} also specifies the maximum display power saving.

In the case where there is no minimum power saving requirement, Λ_{min} will be used in the *Crayon* colour transform to obtain the transformed image. If a minimum power saving, P_{max}, is specified, the first step is to calculate the Λ value that corresponds to that power saving. Since this framework is based on a normalisation scheme that maps $\Lambda \in [0,1]$ to $P \in [0.40, 0.95]$, we can invert this map to get the inverse map $f^{-1} : [0.40, 0.95] \rightarrow [0,1]$, where $P = f(\Lambda) = 0.40 + 0.55\Lambda$. Using this we can find the value Λ_{Pmax} corresponding to P_{max}. Finally, the Λ to be used in the *Crayon* transform will be $min\,(\Lambda_{min}, \Lambda_{Pmax})$.

A further heuristic can be devised based on the sub-pixel power sensitivity differences shown in section 5.6.7. This would involve finding out the sub-pixel whose power is least sensitive to variations in its intensity. Given an image, the total intensities of the sub-pixels can be calculated simply without much computational overhead. The relative proportions of the different sub-pixel intensities can then be compared with the sub-pixel power sensitivities calculated before to decide whether its worth calculating the *Crayon* transform through a simple threshold operation, given the required power saving and image quality. However, the exact threshold would depend on the computational cost of calculating the optimal parameters and carrying out the *Crayon* transform, which we have not explored in this work.

Figure 5.29 shows the results of using the proposed method versus picking Λ arbitrarily by using the mean across the images, as was done with the original Crayon formulation, on the images used in the first user study (except those used to train the models, see fig. 5.21). The proposed method adaptively picks the best Λ for each image (red and purple traces in fig. 5.29b) instead of applying the same (average) Λ for all the images (blue and green traces in fig. 5.29b). This causes some images to be over transformed, leading to image artefacts or not transformed enough, leaving behind potential power saving. Even though the power saving difference is not as drastic, this is due to using the average of the predicted optimal scores. The fixed Λ was calculated by taking the average of the Λ values calculated by the proposed method. Without the method and the Λ normalisation, this choice would have had to be made arbitrarily, for each display model. The proposed method provides a way to intuitively control the transformation using an image quality target and/or a power target, which was not previously possible. The normalisation allows these calculated values to be applicable across multiple display models.



(a) The proposed method picks Lambda (Λ) by considering power saving and image quality target for each image.

(b) Image-specific Lambda (Λ) values lead to better image quality.

Fig. 5.29 Given a target image quality (MOS), the proposed method calculates the most optimal Lambda (Λ) for each image instead of globally choosing Λ which can be too high, causing increased image artefacting or be too conservative, leading to sub-optimal power saving. ('Image index' arbitrarily indexes the images used in the first study (fig. 5.2), without the images used to train the models (fig. 5.21).)

5.10 Spatial Dependence of Display Power

While working on *Crayon*, the authors had observed that there seemed to be a dependence between the spatial arrangement of the pixel values and the current draw that was observed. Their hypothesis was that the pixels that are on, are causing localised heating, which in turn were leading to differences in the resistivity across the panel and thus the panel power draw.

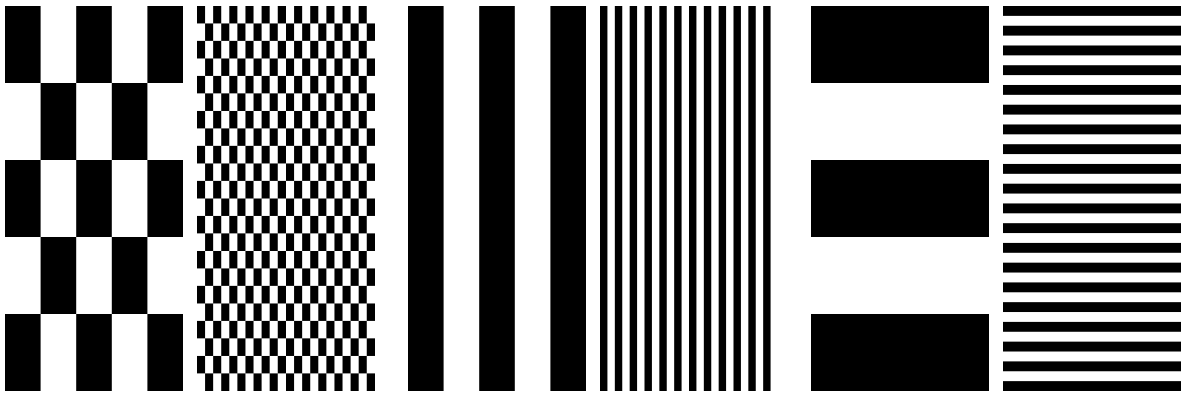
As will be detailed in appendix A, the current (and thus the power) of the *Huawei P30 Pro* display was not significantly affected by the overall panel temperature. Thus we decided to carry out further experiments to fully characterise the dependence of the display power with different spatial arrangements of pixels on the screen. If the local heating hypothesis is true, the power draw would change as the pattern on the screen changes.

5.10.1 Methodology

For simplicity, we explored the power consumption of black and white chequerboard patterns, and vertical and horizontal stripes. Figure 5.30 shows a sample of the patterns used in this study.

The test protocol was as follows. First, a set of images with an equal number of black and white pixels were created. The images were made to be an exact match with the display resolution and aspect ratio. The 'fineness' of the pattern is determined by the number of black and white segments on the horizontal direction for the chequerboard and vertical patterns, and the number of black and white segments along the vertical direction for the horizontal patterns. For example, the coarse patterns in fig. 5.30 (left pattern of each pair), has a 'fineness' of 5.

Starting from a coarse pattern where half the display was black and the other half was white, we generated increasingly fine patterns, at steps of one till a coarseness of 100, after-which the coarseness was increased in steps of 10 until 991, giving 190 images for the vertical patterns and 309 for the horizontal patterns. This was mainly done to reduce the number of images being tested with extremely fine patterns with a negligible difference. Without this step, there would have been 1079 and 2339 images for the vertical and horizontal patterns respectively.



(a) Checkerboard (b) Vertical (c) Horizontal

Fig. 5.30 Images of chequerboard patterns along with vertical and horizontal black and white stripes were used as test patterns in the spatial dependency tests.

We used the *Huawei P30 Pro* device as the target hardware and used the same software setup used to characterise the display power (see appendix A) to show the images and measure the power consumption of the display. The images were displayed periodically, interspersed with completely black images, which served as the points at which baseline power measurements were taken. If the power consumption is affected by the spatial arrangement of the black and white pixels and not just the quantity of them, the current draw should vary as the different images are shown on the screen.

5.10.2 Results

Figure 5.31 shows the results for the chequerboard patterns. It shows that, as the fineness of the pattern is increased, the power draw goes down. The results for the vertical patterns show a similar result (see fig. 5.33). The power draw with the horizontal patterns deviates from this trend, showing that the power draw is more consistent across the range of patterns.

The oscillations in these plots are due to the intermediate measurements of the baseline power draw. The reason for the peaking artefacts in the measurements is unclear. A potential hypothesis is that this could be due to intermittent ‘power-hungry’ background processes happening in the device. It could also just be ‘measurement noise’ or an artefact of the way the *Android* battery subsystem interacts with the device hardware that measures power.

There are several hypotheses for why the power changes as the ‘fineness’ of the patterns are increased.

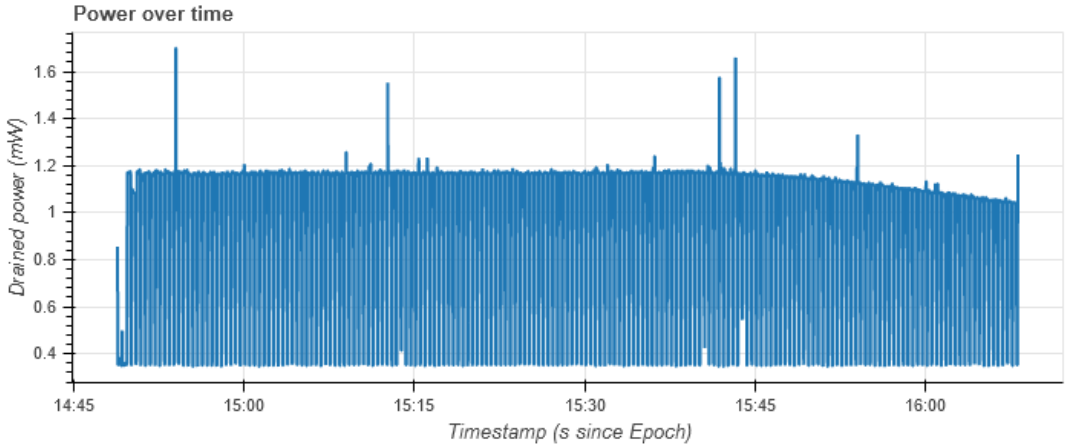


Fig. 5.31 Display power decreases as finer chequerboard patterns are displayed.

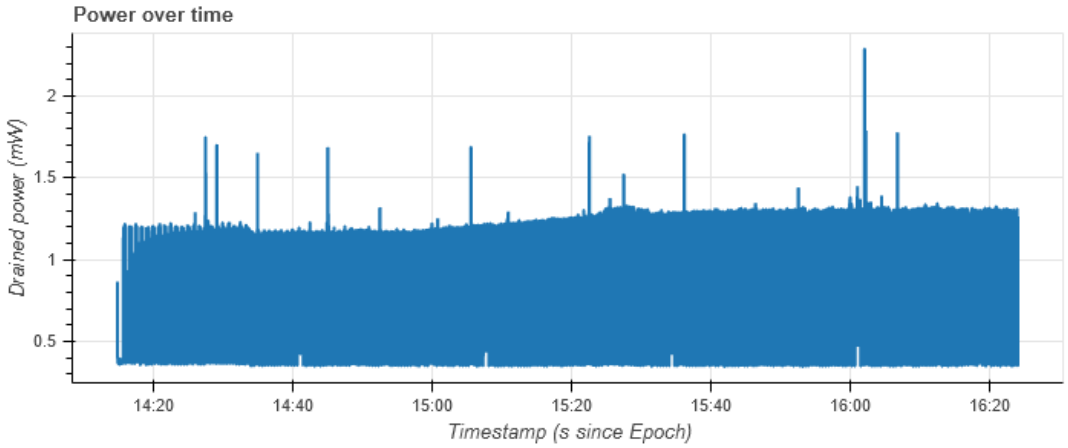


Fig. 5.32 Display power is mostly invariant to the fineness of the horizontal pattern.

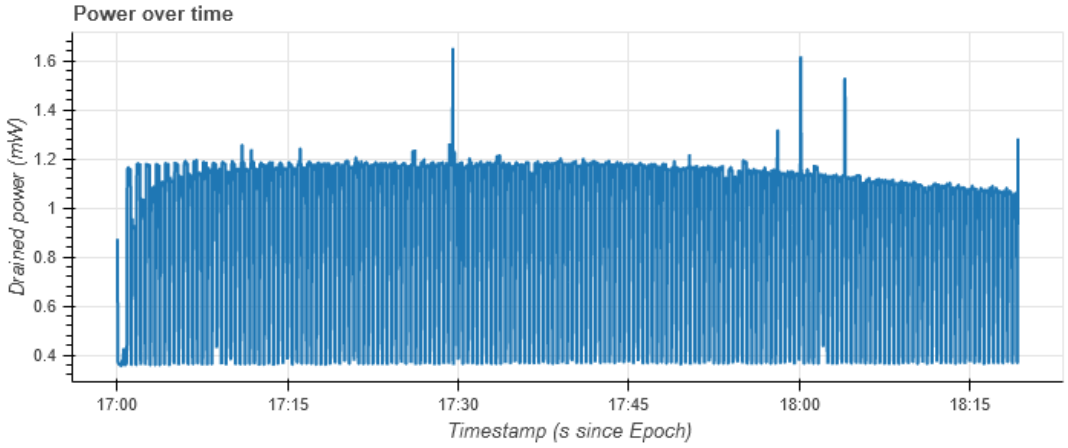


Fig. 5.33 Display power decreases as finer vertical patterns are displayed.

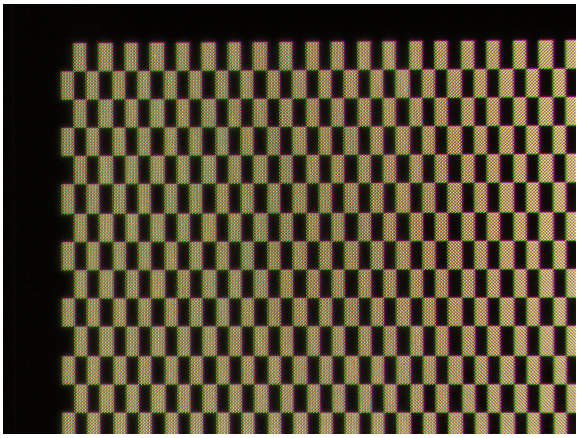
- It could be that when displaying the finest patterns, the display pixels do not align with the image pixels, leading to aliasing and, the actual proportion of black and white pixels being different to the calculated values.
- The *Huawei P30 Pro* has curved display edges on the left and right sides of the device, as well as rounded corners. This would lead to an error in our assumptions about the resolution of the display and thus exacerbate the issue of image pixel misalignment with the display pixels. This might explain why the power drop-off does not affect the horizontal patterns.
- It could also be that the device uses a *PenTile* matrix scheme for the sub-pixels. Since with such a scheme each 'pixel' does not have tri-colour sub-pixels, the assumptions made about the distribution of black and white pixels would be invalid. As *PenTile* matrix displays use proprietary algorithms to carry out sub-pixel rendering, our measurements would be subject to the nuances and the artefacts caused by them.

Pixel Alignment Artefacts

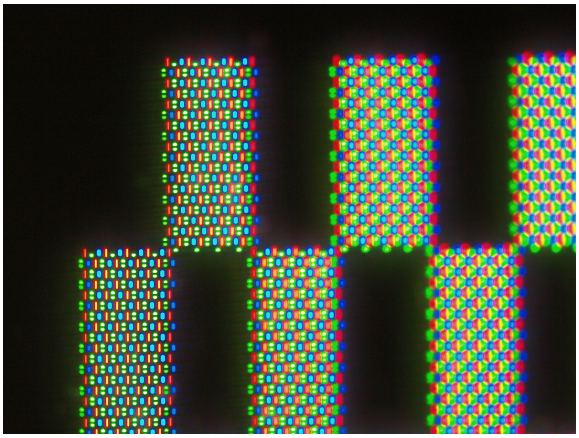
To examine the pixel alignment, we used a microscope to observe the display optically. Figure 5.34 shows close-up images of the display when displaying a chequerboard pattern with a fineness of 121 (i.e., 121 black and white segments horizontally). As can be seen, the sizes of the black and white regions are different. This effect is more extreme in in figs. 5.35b and 5.36b. In the latter case it looks as if the display is showing a white image as opposed to a really fine chequerboard pattern. Furthermore, the edges of the white regions are not as clearly delineated as they should be given the way the images were defined.

It was initially thought that these observations supported the hypothesis that the image pixels and the display pixels are misaligned and that sub-pixel rendering related artefacts were leading to the deviation in the observed power. However, it was later observed that the images themselves exhibited these artefacts.

Under this false assumption, a further investigation was carried out on how to leverage these sub-pixel rendering artefacts to minimise the display power by introducing random black pixels into the images. It was found that the addition of random black pixels did not lead to an observable decrease in power until the proportion of black pixels was substantial enough to cause visual artefacts. Furthermore, when display pixels and image pixels are aligned correctly, even a single black pixel was clearly visible and enough to degrade perceptual quality.

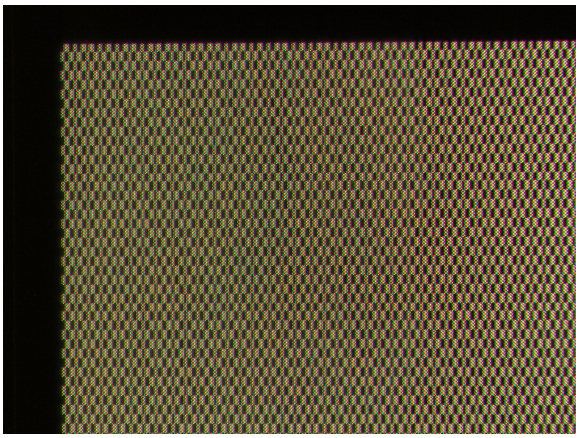


(a) Minimum Zoom



(b) 4x zoom

Fig. 5.34 Close up images of the display show the black and white regions are not aligned.

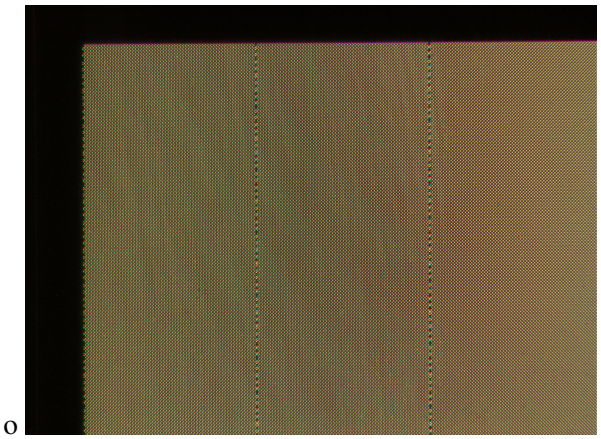


(a) Minimum Zoom



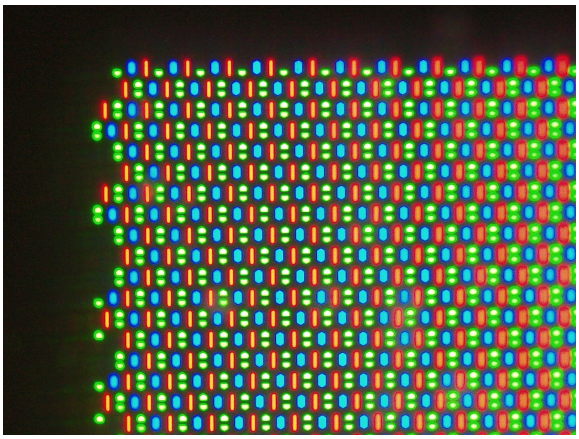
(b) 8x zoom

Fig. 5.35 The pixel misalignment is worsened when finer patterns are displayed (fineness = 371).



o

(a) Minimum Zoom



(b) 8x zoom

Fig. 5.36 The pattern where the pixel colour should change every other pixel, shows all white pixels.

5.11 Conclusions and Future Work

This chapter presented a user study backed method for calculating optimal transform parameters for the *Crayon* colour transform. We showed that we are able to predict the optimal parameter lower bounds through a simple image-colour-heuristic based model, and that it is robust to outliers and is able to potentially improve its performance beyond what is observed if a larger corpus of user scores was provided.

This method would allow the *Crayon* colour transform to be applied more efficiently by tuning it to the properties of the images being transformed. We show that in general, the ℓ_2^2 distance preserves the image quality best over a larger range of Λ values compared to the ℓ_2 distance, while achieving a significant power saving (up to 50%). The exact amount of power-saving will depend on the required image quality, but out method allows the system designer to predict the maximum possible power-saving given an image and the required image quality.

The obvious next step would be to quantify the computational cost of calculating k to calculate the potential real-life power saving of the proposed method. The work presented in this chapter assumes that the energy used in calculating the optimal parameters is negligible compared to the display power savings achieved. This is predicated on the assumption that calculating the heuristics and evaluating the parameter model is computationally trivial, which we have tried to ensure by using simple heuristics and models. However, further work will be needed to prove these assumptions and to calculate the break-even energy (i.e., the image approximation level required to save the amount of energy that would be consumed in calculating the parameter lower bound and the *Crayon* transform).

Another avenue of exploration would be to conduct a larger user study with more users and a more diverse set of images. Based on this data, it would be possible to improve beyond the lower bound prediction and give meaningful estimates for the observer score variability given a chosen Λ.

This chapter also presents an investigation carried out into evaluating the dependency of display power on the spatial arrangement of image pixels. Although a power variation was observed at different image pixel arrangements, we posit that this is due to sub-pixel artefacts caused by the image pixels not aligning with the display pixels.

Sometimes the questions are complicated and the answers are simple.

— Dr. Seuss [137, p. 71].

Chapter 6

Optimal Design of QD based High CRI Smart Lights

6.1 Introduction

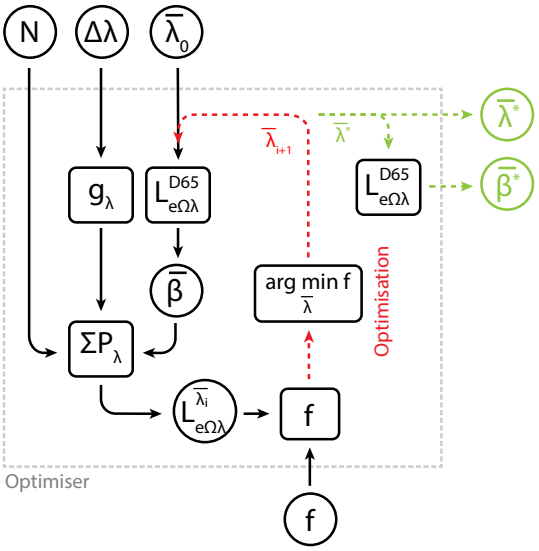


Fig. 6.1 The optimisation stage aims to find the optimal spectral shape as function of the primary wavelengths ($\overline{\lambda}^*$) and intensities ($\overline{\beta}^*$).

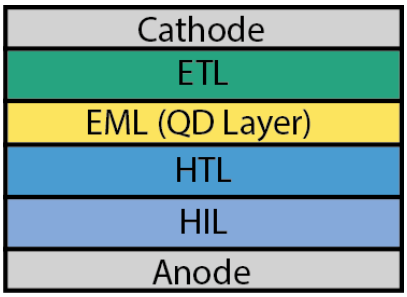


Fig. 6.2 The CTM simulates a six-layer QD device.

In the preceding chapters, the focus has been on QD based displays. This chapter, in contrast, explores the use of QDs in lighting applications and outlines the results of an investigation we carried out into designing, simulating, and fabricating QD based lighting devices. We specifically explore the use of 'multi-primary' architectures for maximising the colour rendering index (CRI). Although 'multi-primary' generally refers to displays with more than three individually controllable primaries, in this case we

use the same idea to create devices capable of generating arbitrary spectral shapes without the need for individual control circuitry.

With advances in lighting technologies, there is a growing demand for increased diversity of lighting applications and, in particular, smart lighting. Smart lighting is a rapidly developing field that is focussed on developing lighting systems that are responsive to consumer requirements, through the use of novel technologies, like QDs, flexible substrates, and machine learning [138].

With the recent developments in the field of QDs and QD based devices, the fabrication of lighting devices based on these technologies have become more viable. As outlined before, QDs are inorganic-nanoscale structures capable of exhibiting optoelectronic transduction properties by leveraging quantum mechanical effects. Compared to traditional LEDs and OLEDs, QD based LEDs (i.e., QD-LEDs) exhibit better colour purity due to narrow bandwidths of the optical emissions, have more easily tunable bandgaps, and have the ability to sustain higher current densities due to the stability of the inorganic compounds QDs are based on [139–142].

Owing to their unitary structure, QDs allow for devices to be made using solution processing techniques which are a lot more cost-effective compared to typical semiconductor fabrication technologies (e.g., epitaxial growth), especially for devices with large emissive areas. QD-LEDs are typically based on three architectures; stacked, patterned, and mixed [143, 62]. In a stacked mode device, the emissive, QD layers are fabricated on top of each-other through a sequential, multi-step spin coating process [144, 145]. In patterned mode, the different coloured QDs are patterned onto non-overlapping regions on the substrate using a transfer printing process [143, 62]. Mixed mode, by contrast, uses the QDs' ability to exist in a colloidal suspension to literally mix different QD solutions together in the required ratios. The fabrication then involves a single-layer spin-coat process [143, 146, 147].

6.1.1 What is CRI?

The particular property of importance to the present investigation is the CRI of a 'smart light'. In general terms, CRI is a metric for quantifying the effect of an illuminant on the perceived colour of objects in comparison to a reference illuminant [148]. A CRI of 100 implies that the perceived 'colour' of an object, when illuminated by the illuminant does not differ from its 'colour' when illuminated by the reference illuminant, which is usually the CIE-D65 white light spectrum. This is particularly relevant in lighting applications where the usability and the comfort of an artificial light built for general illumination depends heavily on its ability to render object colours faithfully. A CRI greater than 80 is generally preferred for use in general illumination [149]. For reference, incandescent and fluorescent lights have CRIs of around 80 and 90 respectively [149].

The process of calculating CRI involves first simulating or measuring the reflectance spectra of 14 standardised colour samples (referred to as the test colour samples (TCSs)) when illuminated by the test illuminant [43, p. 173]. This reflectance spectrum is then transformed using a von Kries transform to account for the chromatic adaptation of the eye, and the *CIELAB* coordinates for the spectrum are calculated [43, p. 174]. The Euclidean distance, ΔE_i, between the *CIELAB* coordinates of the sample

under the test illuminant and its coordinates under the reference illuminant is then calculated, for each colour sample, giving a set of 14 distance measures, $\{\Delta E_i | \forall i \in [1,14]\}$. The individual colour rendering indices, R_i, are calculated from the distance measures, where $R_i = 100 - 4.6\Delta E_i$. The *general colour rendering index*, R_a, is the average of the *first eight* samples; $R_a = \frac{1}{8}\sum_{i=1}^{8} R_i$ [150, 151]. This is usually the quoted value of the CRI for a light source.

6.1.2 Contributions

This chapter presents the following contributions:

- A framework for calculating optimised designs for multi-primary QD lighting devices
- A set of multi-primary device designs for high-CRI lighting
- Show that a CRIs in excess of 96 can be achieved with a mixture of four different QDs.

6.2 Design Framework

A key part of the work presented in this chapter was to go beyond a simple set of optimised parameters for a high-CRI light, and provide a framework for end-to-end design of QD based optical devices.

Typically, semiconductor device design is approached from the bottom up; given a set of properties the device should have, the designer tries to pick the required semiconductor material properties using their engineering judgement. These judgements are tested through simulations or by fabricating devices. The chosen material properties are then changed, based on engineering judgement as before, to modulate the device performance towards what is desired. This iterative process is very time-consuming, and the quality of the design is heavily reliant on the expertise of the designer.

In this chapter we present a top-down design framework for designing QD based optoelectronic devices that aims to automate a significant part of the iterative design process. The designers can specify the performance required of the final device, either explicitly (e.g., required luminance at the operating voltage) or as objective functions to optimise against (e.g., a heuristic function based on the CRI and the chromaticity of the device's emission spectrum). The proposed method then aims to find the optimal low-level semiconductor properties required from the device. These details can then be used to pick materials and carry out device fabrication or to direct material research.

The design process has three stages; numerical optimisation, device modelling, and validation through fabrication.

6.2.1 Numerical Optimisation

In this stage, the aim is to define optimality conditions and use numerical optimisation methods to find device properties that are 'optimal'. For example, in the present investigation, we are using numerical optimisation to find the shape of the spectrum required (the device property) to maximise the CRI (the optimality condition).

The results from this are used in the next stage, either directly as device properties (e.g., the number of primaries) and/or as a basis for calculating the device properties (e.g., the QD bandgap and vacuum energy required to generate the calculated spectra).

6.2.2 Device and Process Simulation

At this stage, the device is defined using its physical properties (e.g., doping concentrations, layer thicknesses, recombination rates, etc.), based on the results of the previous stage. A physically accurate simulator is used to model the device physics (e.g., through technology-CAD (TCAD) simulations). Results of this stage can be fed back into the previous stage for iterative, closed-loop optimisation. This can be repeated until the performance of the simulated device is as desired.

In the current investigation, the device modelling is done through a charge-transport model (CTM) based simulation. Given the QD properties, the simulator calculates the optoelectronic properties of the device (e.g., current vs voltage trace, current efficiencies, emission spectra, etc.). Due to the simplicity of the device/model and objective function used in our investigation, as will be seen later, satisfactory results can be obtained without closed-loop optimisation.

6.2.3 Validation Through Device Fabrication

This stage is simply the fabrication of the device designed in the previous stages. However, as before, the performance results of the fabricated devices can be fed back to the device modelling stage to update the simulator to better model the device performance variations caused by the manufacturing processes. This is especially relevant if a 'process simulator' is used in the device modelling stage.

6.3 Finding Optimal QD Primary Combinations for Maximum CRI

The first step in designing a device with maximised CRI is to investigate how the composition of a spectrum affects its CRI. Given that CRI is generally defined with respect to CIE-D65 spectrum, we chose that as the target spectrum. The aim was then to find a combination of QDs that give a CRI as close to the D65 spectrum while ensuring that the chromaticity of this spectrum is also identical to that of CIE-D65. The latter condition is required, not only to preserve the aesthetic property of the light, but the CRI calculation requires that the emission spectrum of the device has a correlated colour temperature (CCT) similar to the target spectrum [43, p. 173].

We defined the required optical properties through a *cost function* which we iteratively optimised to find the optimal spectral radiance required of the device. Although in this application, we used chromaticity match and high CRI as the components in the objective function, in general, this framework can be used to find optimal device properties satisfying any measurable quality metric. The device properties (band-gap, QD size, etc.) required to generate the calculated spectrum were then calculated and used in the CTM simulation to predict the optoelectronic properties of the proposed device.

6.3.1 Methodology

Problem Definition

We chose the family of spectral shapes that are a sum of N Gaussian-shaped primary emissions as our hypothesis space for the spectral optimisation. We parametrised each primary emission by its mean wavelength and peak radiance, and specified them to have the same bandwidth. Thus, for an n-primary device, we can describe the shape of the spectrum using $2n+1$ parameters; mean wavelengths and peak radiance of the n-primaries, and the FWHM of the emissions. The procedure for deriving the optimal primary emissions shapes can be summarised in the following way:

1. Choose a target spectrum: CIE-D65 $L_{e\Omega\lambda}^{D65}$
2. Choose the number of primaries: N
3. Choose the FWHM of the primaries: $\Delta\lambda$
4. Choose the initial placement of the primaries: $\left(\overline{\lambda},\overline{\beta}\right) \rightarrow L_{e\Omega\lambda,0}$
5. Define the optimality conditions using a cost function: f
6. Iteratively change the position and intensity of the primaries to minimise the cost function.
7. Use the resulting wavelength and peak radiance combinations in the device simulation: $\overline{\lambda^*},\overline{\beta^*}$

Let $g_\lambda(\lambda_i,\Delta\lambda)$ be a Gaussian distributions centred at λ_i with a FWHM $\Delta\lambda$.

$$g_\lambda(\lambda_i,\Delta\lambda) = \frac{2\sqrt{2ln(2)}}{\Delta\lambda}\frac{1}{\sqrt{2\pi}}e^{-\left[\frac{2\sqrt{2ln(2)}}{\Delta\lambda}(\lambda-\lambda_i)\right]^2} \tag{6.1}$$

Let $P_\lambda(\lambda_i,\Delta\lambda,\beta)$ be the spectrum of an individual primary with peak radiance of β_i.

$$P_\lambda(\lambda_i,\Delta\lambda,\beta_i) = \beta_i\frac{g_\lambda(\lambda_i,\Delta\lambda)}{\int_{\forall\lambda} g_\lambda(\lambda_i,\Delta\lambda)\,d\lambda} \tag{6.2}$$

Thus, for a device with N primaries, the total spectral radiance can be defined as,

$$L_{e\Omega\lambda}(\overline{\lambda}) = \sum_{i=1}^{N} P_\lambda(\lambda_i,\Delta\lambda,\beta_i) \tag{6.3}$$

In the initial evaluation, we found that using the D65 spectral shape to define the radiance of the primaries, and restricting the optimisation process to only find the primary wavelengths led to better (and more stable) results. Thus, eq. (6.3) can be rewritten as;

$$L_{e\Omega\lambda}(\overline{\lambda}) = \sum_{i=1}^{N} P_\lambda(\lambda_i,\Delta\lambda,L_{e\Omega\lambda}^{D65}(\lambda_i)) \tag{6.4}$$

where $L_{e\Omega\lambda}^{D65}(\lambda_i)$ is the radiance of the D65 spectrum at λ_i (peak wavelength of the i^{th} primary), N is the number of primaries, and $\overline{\lambda}$ is an N-tuple of the primary wavelengths. Let $L_{e\Omega\lambda,0}$ be the initial spectral radiance, calculated using the initial set of primary wavelengths, λ_0, and eq. (6.4). We set $\Delta\lambda$ as 20 nm for all the primaries. In this investigation we looked at devices with up to 8 primaries ($N=8$).

Initial Primary Placement

We investigated two initialisation strategies for the initial placement of the primaries; uniform distribution in the range 330 nm to 700 nm, and cone fundamental bound limited uniform distribution. In the former case, the N primaries were placed at uniform distances in the wavelength range 330 nm to 700 nm which is larger than the visible range. In the latter case, we restricted this range to be within the minimum and maximum peak locations of the Stockman & Sharpe 2° cone fundamentals with 5% padding on each end (i.e., the range from 420 nm to 599 nm) [49, 50]. Figure 6.3 show the initialisation ranges highlighted on the spectral range relative to the cone fundamentals. The justification was that, because the CRI calculation (implicitly) relies on weighting the spectrum of the illuminant by the spectral sensitivities of the eye, by placing the primaries in the spectral range that the eye is most sensitive to, their effect on the CRI can be maximised.

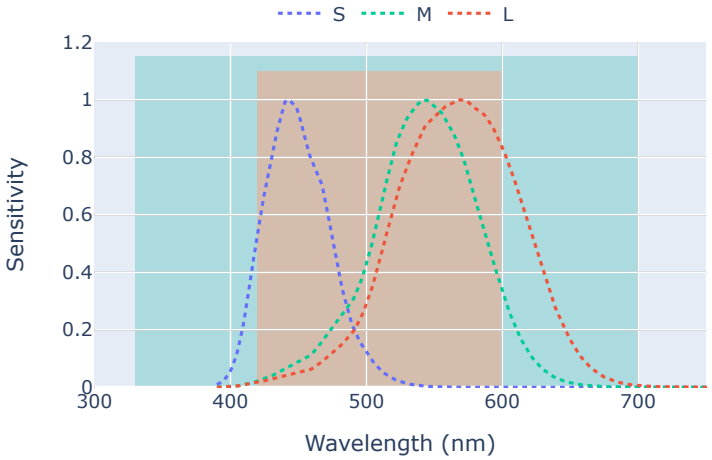


Fig. 6.3 The optimisation process used two initialisation schemes for the primaries: ② uniformly placed primaries in the range 330 nm to 700 nm (pale green shaded area) and ③ primaries placed uniformly in a narrow range centred around the LMS cone fundamental peaks (420 nm to 599 nm) (pale orange shaded area).

Optimisation Process

The optimisation procedure for finding the optimal spectrum maximising the CRI relies on three key values; the initial spectral radiance ($L_{e\Omega\lambda,0}$), the cost function (f, also referred to as the 'objective function'), and the choice of the optimisation algorithm. The cost function, f, encodes the properties required from the optimal spectrum; namely, high CRI and close chromaticity match with D65.

We looked at the following cost functions for this iterative optimisation process:

1. CIE-ΔE_{00} difference between the device spectrum and the D65 spectrum.
2. ℓ_2^2-norm of the xy chromaticity coordinates:

$$f(L_{e\Omega\lambda}) = \left\| \left(CIE_{xy}\left(L_{e\Omega\lambda}\right) - CIE_{xy}\left(L_{e\Omega\lambda}^{D65}\right) \right) \right\|_2^2 \tag{6.5}$$

3. Reciprocal CRI:

$$f(L_{e\Omega\lambda}) = \frac{1}{CRI(L_{e\Omega\lambda})} \tag{6.6}$$

4. Exponentially normalised reciprocal CRI:

$$f(L_{e\Omega\lambda}) = 10\ ln\left[exp\left(\frac{CRI(L_{e\Omega\lambda})}{10}\right)+1\right]^{-1} \tag{6.7}$$

5. ℓ_2^2-norm of xy chromaticity coordinates and reciprocal CRI clamped to the range [0,100]:

$$f(L_{e\Omega\lambda}) = \left\|(CIE_{xy}(L_{e\Omega\lambda}) - CIE_{xy}\left(L_{e\Omega\lambda}^{D65}\right))\right\|_2^2 + clip(CRI(L_{e\Omega\lambda}),0,100)^{-1} \tag{6.8}$$

6. As above, but with exponentially normalised reciprocal CRI:

$$f(L_{e\Omega\lambda}) = \left\|(CIE_{xy}(L_{e\Omega\lambda}) - CIE_{xy}\left(L_{e\Omega\lambda}^{D65}\right))\right\|_2^2 + 10\ ln\left[exp\left(\frac{CRI(L_{e\Omega\lambda})}{10}\right)+1\right]^{-1} \tag{6.9}$$

7. Weighted difference between ℓ_2^2-norm and linearly normalised CRI:

$$f(L_{e\Omega\lambda}) = 1.5\left\|(CIE_{xy}(L_{e\Omega\lambda}) - CIE_{xy}\left(L_{e\Omega\lambda}^{D65}\right))\right\|_2^2 - \left[\frac{CRI(L_{e\Omega\lambda})}{100}-1\right] \tag{6.10}$$

8. Weighted difference between ℓ_2^2-norm and linearly normalised CRI + CRI spread penalty:

$$\begin{aligned} f(L_{e\Omega\lambda}) =& 3\left\|(CIE_{xy}(L_{e\Omega\lambda}) - CIE_{xy}\left(L_{e\Omega\lambda}^{D65}\right))\right\|_2^2 \\ & - 2\left[\frac{CRI(L_{e\Omega\lambda})}{100}-1\right] + \left\|\frac{max(CRI(L_{e\Omega\lambda})) - min(CRI(L_{e\Omega\lambda}))}{100}\right\|_2^2 \end{aligned} \tag{6.11}$$

In all cases, $CIE_{xy}(L)$ and $CRI(L)$ are the xy-chromaticity coordinates and the CRI of the spectrum L, respectively. The optimisation process is given by eq. (6.12) and the result of which is the set of optimal primary wavelengths, $(\overline{\lambda^*})$. The spectrum with these primaries is the optimal spectrum, $L_{e\Omega\lambda}^*$.

$$\lambda^* = \arg\min_{\lambda} f\left(L_{e\Omega\lambda}\left(\overline{\lambda}\right)\right) \tag{6.12}$$

The initial placement of primaries $(\overline{\lambda_0})$, is used to calculate the initial spectrum $L_{e\Omega\lambda,0}$. The optimality of this spectrum is then calculated with the chosen cost function, f, and the optimiser picks the next set of primaries, $\overline{\lambda_1}$, such that $f(L_{e\Omega\lambda}(\overline{\lambda_1})) < f(L_{e\Omega\lambda}(\overline{\lambda_0}))$. This is repeated until the optimiser converges to a minimum value of the cost function. The final set of optimised candidate wavelengths are chosen as the optimal set of wavelengths, $\overline{\lambda^*}$. $\overline{\beta^*}$ are the radiances of the D65 spectrum at the corresponding wavelengths in $\overline{\lambda^*}$. The total luminance of the final device is defined as part of the problem as a luminous intensity and this is invariably different to the total luminance of the D65 spectrum. However, the aim here is to match the CRI of D65 not its luminance. As such, the important

quantities are the relative intensities of the primaries, which we define as the vector $\overline{w}$. These are calculated by normalising the peak primary radiances $\overline{\beta^*}$ to have a minimum value of 1. Figure 6.1 illustrates the optimisation process as a flow diagram.

We evaluated three optimisation algorithms; *Nelder-Mead*, *L-BFGS-B*, and *SHGO* [103, 105, 106, 101]. *Nelder-Mead* is a derivative-free optimisation method which works well for objective functions where the derivatives are unknown. However, it cannot explicitly incorporate optimisation bounds. *L-BFGS-B* in contrast is able to do bounded optimisation, but uses cost function gradients for the optimisation. *SHGO* is similar to *Nelder-Mead* in that it uses simplexes, but it aims to find the globally optimal value by evaluating multiple initial conditions.

6.3.2 Results

Effect of the Initialisation Scheme

Regardless of the cost function, optimising with the full range of wavelength initialisation led to poor results (e.g., the optimiser getting stuck on local optima). When using the cone fundamental bounds, we observed that the optimisers (*Nelder-Mead* in particular) spread these primaries in a way that resulted in a high CRI. Picking, initial primaries far away from the photopic LEF peak led to unstable optimisation results (e.g., overlapping primary peaks). We believe that this is because the LEF places a skewed initial weight on the different primaries at the start if they are too far apart and too far from the LEF peak, which biases the optimiser into giving more importance to the primaries at the centre too quickly.

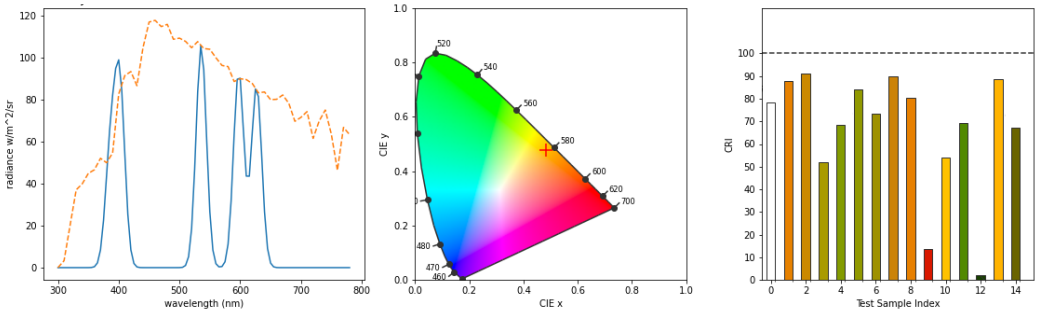
Effect of the Cost Function

As expected, just optimising for colour match (cost functions 1 and 2) or CRI (cost functions 3 and 4) led to either poor CRI values or poor colour match. However, in cases with high number of primaries (> 5), using reciprocal CRI (cost function 3) resulted in spectra with quite high CRI and chromaticity similar to D65 (see fig. 6.4). Following this observation we sequentially optimised for chromaticity (with cost function 2) and then for CRI (with cost function 3). Figure 6.5 shows these spectra and the individual CRIs. As can be seen in the 8-primary case, we can achieve a CRI of 98.2 but only 6 of the primaries are within the visible spectral range. Recalculating CRI for the spectrum with only the central primaries resulted in a CRI of 97.46 (see fig. 6.17c).

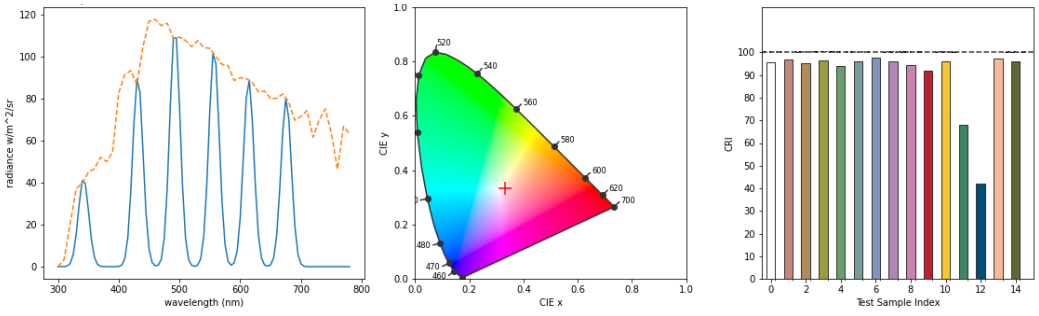
Cost functions 5 and 6 lead to inconsistent results or fail to converge altogether due to scale mismatch between the sum of squares difference of the chromaticities and the reciprocal CRI. Cost functions 7 and 8 used a different normalisation method which was more tenable to optimisation. Please refer to appendix C for a tabulated list of the results.

6.3.3 Optimised Primary Spectrum

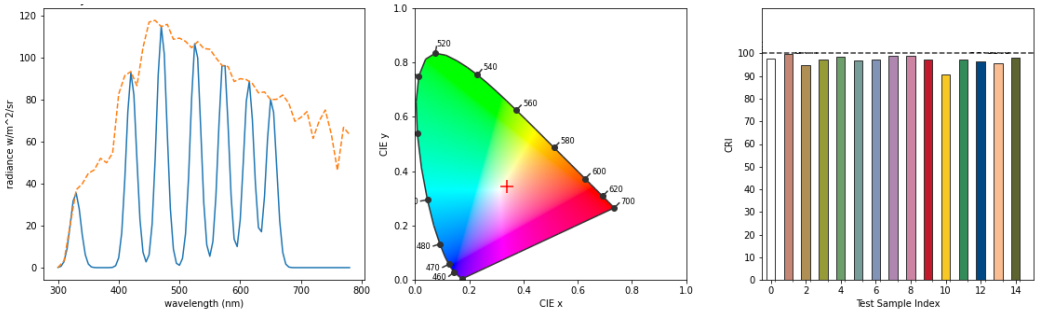
From all the simulations, we picked out the 3, 4, 5, and 6 primary combinations outlined in table 6.1 for the simulations. The 3 to 5 primary combinations were achieved with cost function 7 and the 6



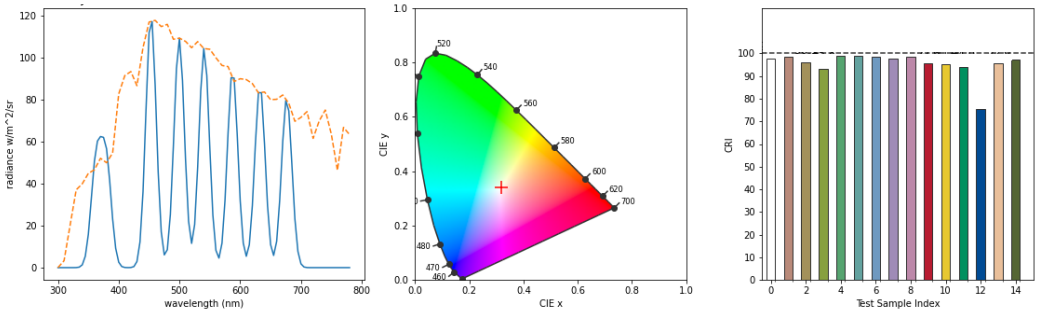
(a) 5-primaries.



(b) 6-primaries.

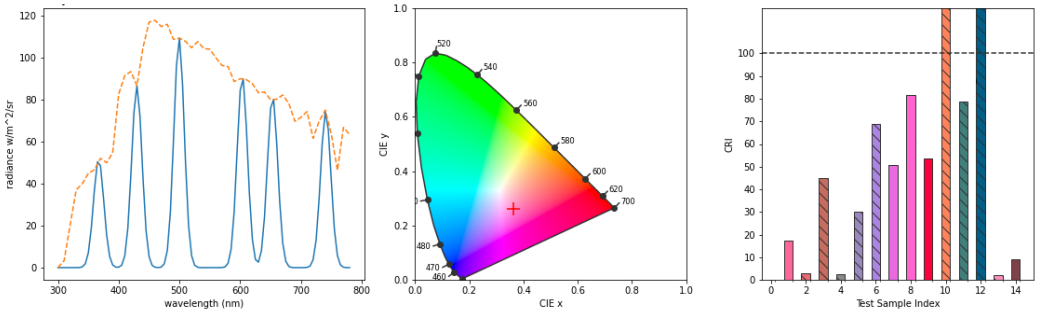


(c) 7-primaries.

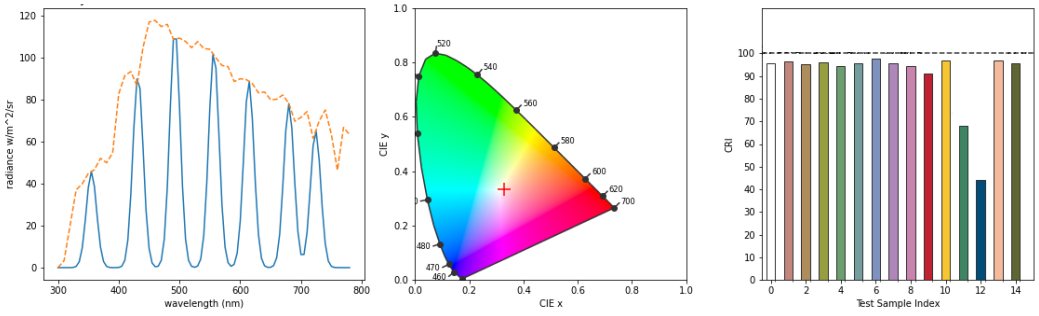


(d) 8-primaries.

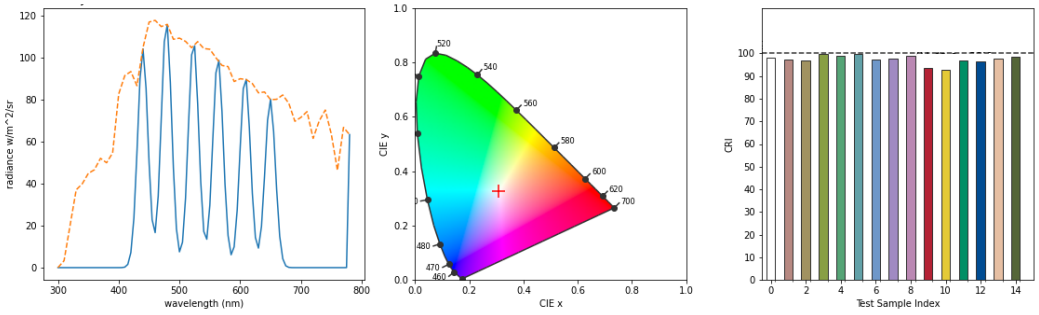
Fig. 6.4 Plots show the results of the optimisation process with the *reciprocal CRI* cost, for an increasing number of primaries from 5 to 8. Note that the peak at 400 nm in the 5-primary case and 370 nm in the 8-primary case are actually two primaries overlaid (as evident by the increased FWHM).



(a) 6-primaries.



(b) 7-primaries.



(c) 8-primaries.

Fig. 6.5 Sequential optimisation using *sum of squares* cost and the *reciprocal CRI* cost does not reliably result in spectra with both chromaticity match and high CRI.

primary spectrum was found through the sequential optimisation process outlined in the previous section. Table 6.2 summaries the properties of these spectra.

N	Primary Wavelengths, $\overline{\lambda^*}$ (nm)	Relative Intensities, $\overline{w}$
3	[463, 542, 611]	[1.308, 1.167, 1.000]
4	[450, 502, 559, 616]	[1.323, 1.233, 1.135, 1.000]
5	[433, 473, 526, 575, 623]	[1.067, 1.333, 1.233, 1.112, 1.000]
6	[440, 480, 525, 565, 610, 650]	[1.310, 1.449, 1.328, 1.227, 1.120, 1.000]

Table 6.1 Optimised primary wavelengths for 3, 4, 5, and 6 primary devices.

N	ΔE_{00}	Δxy	CIE_{xy}	CRI_{avg}	CRI_{max}	CRI_{min}
3	38.76	0.0131	[0.311, 0.319]	84.75	96.21	30.19
4	40.32	0.0227	[0.299, 0.312]	96.89	99.05	82.86
5	39.66	0.0244	[0.303, 0.310]	96.11	98.91	88.22
6	24.35	0.0035	[0.309, 0.328]	97.46	99.49	91.78

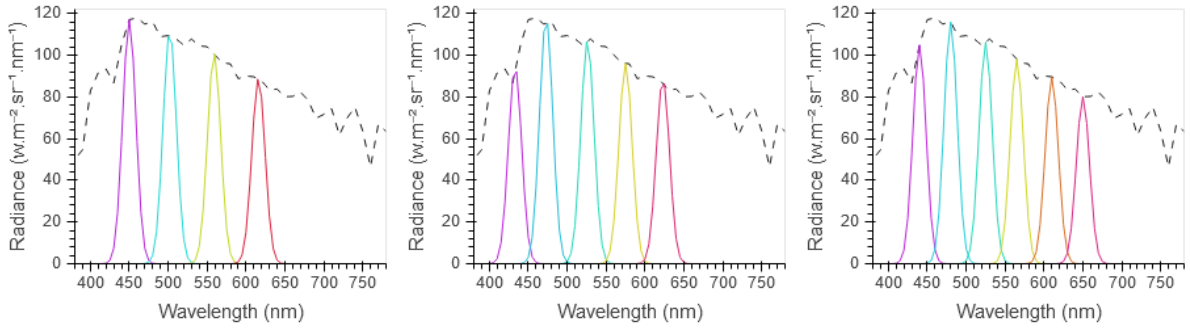
Table 6.2 Chromatic properties of the calculated 3, 4, 5, and 6 primary devices.

Figure 6.7c shows the CRI values for the different primary numbers, along with the maximum and minimum individual TCS CRI values. As can be seen, with three primaries we can create a device with an average CRI of 84.8, but the minimum CRI is an unacceptably low 30.2 (for TCS-12 or R_{12}). However, with the 4-primary device we can achieve a CRI of 96.9 with a minimum CRI of 82.9 (R_{12}). With 5 and 6 primaries the minimum CRI can be increased to 88.2 (R_{10}) and 91.8 (R_{12}) respectively, with the average CRIs reaching 96.1 and 97.5.

Figure 6.7b shows the variation of the ΔE_{00} of the calculated spectral combinations for the different number of primaries. The ΔE_{00} metric measures the Euclidean distance between two spectra in the *CIELAB* colour space and is a measure for perceptual similarity. In all cases, except the last, $\Delta E_{00} \approx 40$ when compared to the standard D65 spectrum, but for the 6-primary $\Delta E_{00} \approx 24$. It should be noted that the ΔE_{00} metric considers the luminance of the spectra as well and not just the chromaticity. As the final devices are only concerned with matching the chromaticity of D65, we also show Δxy for the different spectral combinations in fig. 6.7a. Δxy is a measure of the Euclidean distance between the CIE-xy chromaticity coordinates. As can be seen, the chromaticity difference is below 0.03 for all the optimised devices regardless of the number of primaries.Figure 6.8 shows this information graphically on a cropped in portion of the CIE-xy chromaticity diagram.

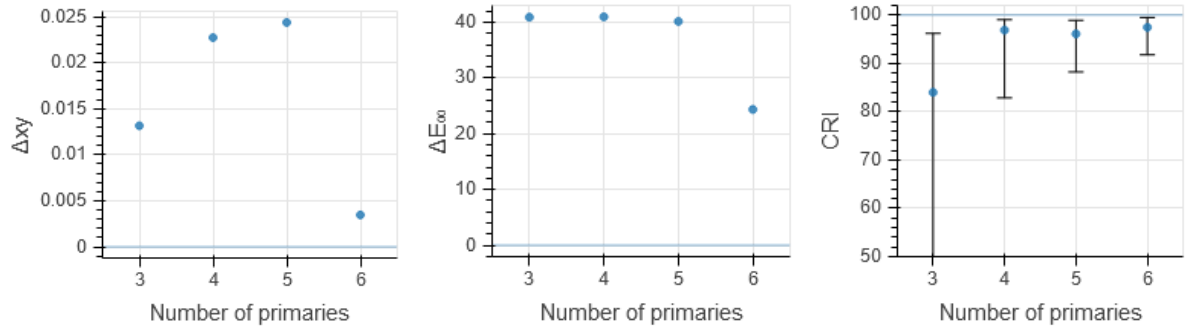
6.4 Charge-Transport Model Simulation

The QD simulator we used was designed to simulate a six layer QD device (see fig. 6.2) and uses a CTM to calculate the optoelectronic properties of the device. The physical and chemical properties of the different layers are defined in a configuration file, which the simulator uses to define the device and calculate its behaviour at different input voltages. Figure 6.14 shows how the properties of a complete QD device can be defined for this simulator using the *Python* interface we built.



(a) 4-primaries (b) 5-primaries (c) 6-primaries

Fig. 6.6 The spectral shapes of the 4, 5, and 6 primary devices show that they match the intensity of the CIE-D65 spectrum as designed.



(a) All the primary combinations match the D65 chromaticity closely. (b) Luminance differences led to a high ΔE_{00}. (c) The CRI increases with the number of primaries.

Fig. 6.7 The variation of the spectral properties of the chosen devices as a function of the number of primaries.

6.4.1 Charge-Transport Model

This model is based on work done by Kumar et al. and was modified by Dr. Sung-Min Jung to fit our device configuration to develop the QD simulator used in the sections that follow [152]. This section provides an overview of the model and the QD simulator for completeness but it should be noted that neither the model, nor the simulator are key contributions of this report. Some of the equations are omitted on purpose as they relate to unpublished work by Dr. Sung-Min Jung.

The current density in the QD device being simulated is described by three sets of equations depending on the location of the device; namely the drift/diffusion, injection, and tunnelling current densities. The drift/diffusion current density defines charge transport in the transport and injections layers and is governed by the charge density profile as well as the bulk charge mobility and the electric field in these regions. The injection current density, describes charge transport due to capture/emission processes occurring between the QD layer and the semiconductor layers (ETL and HTL). Finally, the tunnelling current density equations describe the movement of the charge carriers between the different QD layers and this is governed by the charge densities in the QD layer and coefficients describing the

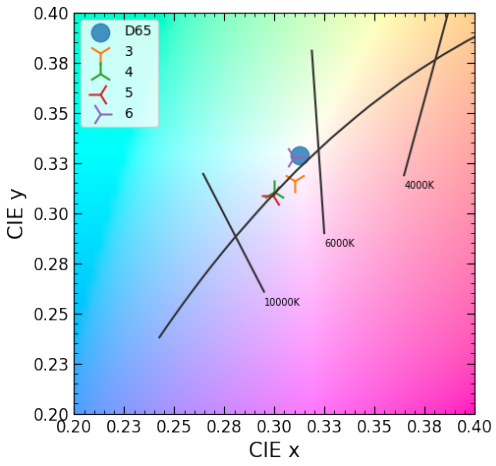


Fig. 6.8 The chromaticities of the calculated spectra closely match that of CIE-D65.

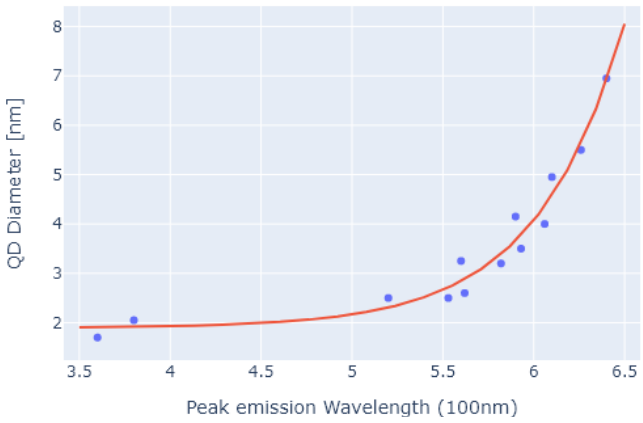


Fig. 6.9 QD diameter model was obtained by fitting an exponential to measurements in the literature.

tunnelling process. Each set of equations constitutes two separate equations for the hole and electron currents, and the sum of these currents must be constant throughout the device.

The time-evolution of these carrier densities are described in terms of the gradient of the current density at a given point in the device, and is the difference between the charge flowing into that point and the sum of radiative and non-radiative recombination processes. This is given by the following equations:

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial z} - (U_{srh} + U_{aug} + U_{rad}) \tag{6.13}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\frac{\partial J_n}{\partial z} - (U_{srh} + U_{aug} + U_{rad}) \tag{6.14}$$

where, U_{srh} and U_{aug} are the Shockley-Reed-Hall and the Auger recombination rates, and U_{rad} is the radiative recombination rate in the device. J_p and J_n are the hole and electron current densities as given by the three sets of equations described above.

The charge distribution across the device imposes an electric potential across it which is governed by the following Poisson's equation:

$$\varepsilon_0 \frac{\partial}{\partial z}\left(\varepsilon_r \frac{\partial \varphi}{\partial z}\right) = -e\left(p - n + N_d^+ - N_a^-\right) \tag{6.15}$$

where $\phi(z)$ is the electrostatic potential across the device, N_d^+ and N_a^- are the donor and the acceptor concentrations in the semiconductor layers while p and n are their hole and electron densities. These three equations form a coupled set of differential equations that needs to be solved simultaneously to find the steady-state performance of the QD device. By discretising the device geometry and iteratively solving these equations at each point through the device, the simulator calculates the electronic properties

of the device. These results are then used to calculate the emission spectrum and the other optical properties of the device.

The radiant flux of the device is equal to the total radiative recombination rate in the device. Assuming the spectral distribution follows a Gaussian shape, the spectral radiance of the device can be calculated using eq. (6.2) but with β as a function of the radiative recombination rate U_{rad} as given in eq. (6.16).

$$\beta = \frac{2\sqrt{2ln(2)}}{\Delta\lambda\sqrt{\pi}}\frac{1}{\pi}\frac{hc}{\lambda}U_{rad} \tag{6.16}$$

where $\Delta\lambda$ is the FWHM, and λ is the center wavelength of the emission.

6.4.2 Calculating QD Device Parameters

We chose ITO and Aluminium for the Anode and Cathode materials, respectively. For the hole injection layer (HIL), HTL, ETL, we chose PEDOT:PSS [153], TFB [154–156], and ZnO [157–159] materials. The target rail voltage was set to be *10V* and *100 cd* was chosen as the target luminous intensity. The numerical optimisation gives the values for the primary intensities ($\overline{\lambda^*}$), their peak radiances ($\overline{\beta^*}$), and their FWHMs. In the simulator the FWHM can be defined explicitly (see fig. 6.14) The key parameters affecting the shape of the optical emission of a QD are its bandgap (E_g^{QD}) and diameter (d^{QD}).

6.4.3 Band Gap Calculation

The band gap calculation is straightforward as the emission wavelength, λ, and the QD band gap, E_g^{QD}, are related through the speed of light, c, and the planks constant,h, as described in eq. (6.17).

$$E_g^{QD} = \frac{hc}{\lambda} \tag{6.17}$$

6.4.4 Diameter Calculation

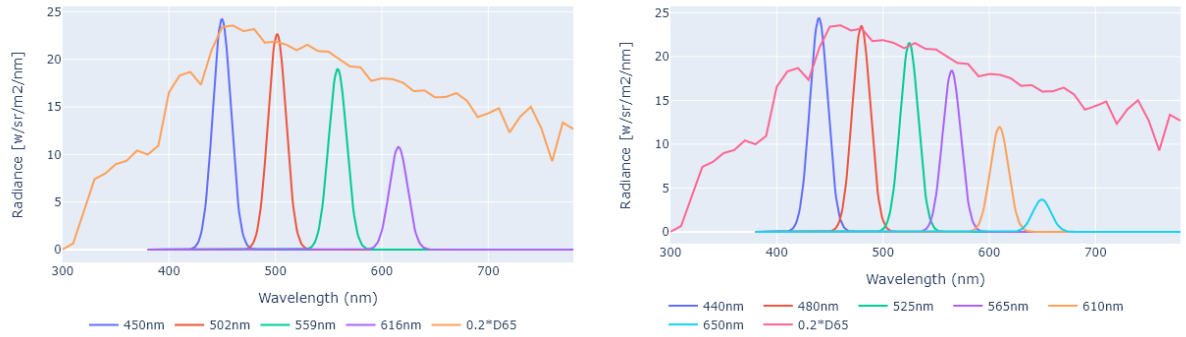
The exact value of the QD diameter depends on the QD material, passivation processes, and other process parameters involved in the fabrication of the QD. However, generally, the band gap in inversely correlated with the diameter. Using values from the literature, we fit an exponential curve to model the variation of the QD diameter (see eq. (6.18) and fig. 6.9) [160]. Note that, the values on the x-axis of fig. 6.9 are scaled to be in units of 100 nm to improve the numeric stability of the curve fitting process.

$$3.5969E-04e^{2.0847\cdot(x-1.8240)}+1.8973 \tag{6.18}$$

6.4.5 Simulation Results

When used in patterned mode, the interaction between the different QDs is minimal, thus each primary can be considered to be independent of the other primaries. We used this fact to simulate the QDs

in parallel to calculate their optoelectronic properties. For brevity, we would only be presenting the simulation results for the 4 and 6 primary devices.



(a) 4-primary device.

(b) 6-primary device

Fig. 6.10 The relative radiances of the simulated primaries do not match the required radiance ratio if the emission area was the same for all the primaries.

Figures 6.10a and 6.10b shows the raw spectral outputs at V=10 V along with 1/5th scale plot of the CIE-D65 spectrum, for the 4-primary and 6-primary, device respectively. However, the required radiance ratio is different to the natural emission ratio at 10 V. As radiance is a function of the emission area, we can modulate each primary radiance by setting the relative emission areas of the primaries to be equal to the required radiance ratios adjusted by nominal radiance at 10 V for each primary. The absolute values for the areas were then calculated using the required luminous intensity (100 cd). Table 6.3 summaries the results of these calculations.

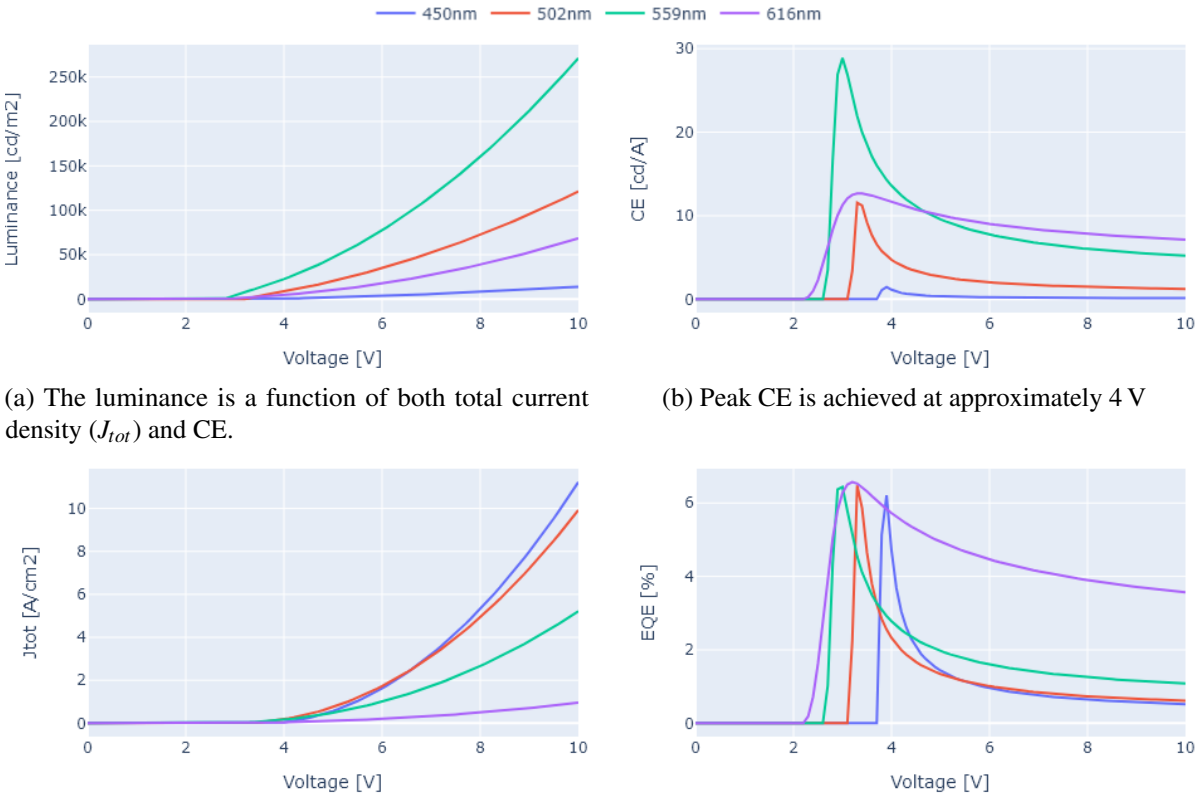
		4-Primary	6-Primary
at 10V	Radiance	76.5 $w/sr \cdot m^2$	103.7 $w/sr \cdot m^2$
	Total Area Required	1,911.7 mm^2	1,416.5 mm^2
	Sq. pixel Dimensions	43.7 mm	37.6 mm
	Primary Ratios	[1.00, 1.00, 1.10, 1.70]	[1.00, 1.15, 1.15, 1.24, 1.74, 5.06]
	Primary Currents (A)	[44.78, 39.46, 22.80, 6.50]	[13.98, 15.53, 11.99, 7.18, 2.66, 1.16]
	Total Current	113.54 A	52.51 A
at 4V	Radiance	4.80 $w/sr \cdot m^2$	6.14 $w/sr \cdot m^2$
	Total Area Required	30,518.0 mm^2	23,849.2 mm^2
	Sq. pixel Dimensions	174.7 mm	154.4 mm
	Primary Ratios	[2.12, 1.06, 1.00, 1.85]	[1.95, 1.32, 1.00, 1.04, 1.67, 7.82]
	Primary Currents (A)	[4.74, 9.81, 8.37, 3.82]	[0.94, 2.53, 2.94, 2.28, 1.22, 1.06]
	Total Current	26.75 A	10.97 A

Table 6.3 Properties of the simulated 4 and 6 primary devices.

To achieve a luminous intensity of 100 cd with the 4-primary device, we would require an area of $1,911.7\,mm^2$ which corresponds to a square with $43.7\,mm$ sides. Similarly, for the 6-primary device, we can achieve the same luminous intensity from a $1,416.5\,mm^2$ area ($37.6\,mm$ square). Figure 6.13 shows this information graphically.

To generate the required luminance, the 4-primary device consumes approximately 114 A, while the 6-primary device consumes approximately 53 A. These values are quite high in comparison to typical LED based devices. However, considering the low current efficiency (CE) and external quantum efficiency (EQE) at the target voltage, and the high target luminous intensity, this is as expected.

A simple solution to this would be to reduce the target voltage to around 4 V. At this voltage the current efficiencies of the primaries are much higher than at 10 V (see figs. 6.11b and 6.12b). At 4 V the two devices can achieve the target luminance at 26.75 A and 10.97 A respectively but with a much higher emission area (nearly 15x more, see table 6.3). Assuming a 1 Sr emission angle, this corresponds to a best-case luminous efficacy of about 2.2 lm/W. For reference, a typical commercial LED bulb would have a luminous efficacy of about 100 lm/W [161]. Figures 6.17b and 6.17d show the chromaticities and the individual CRI values for the simulated 4 and 6 primary devices at 10 V. The results show that the devices were able to achieve CRIs of 96.96 and 97.38 respectively.



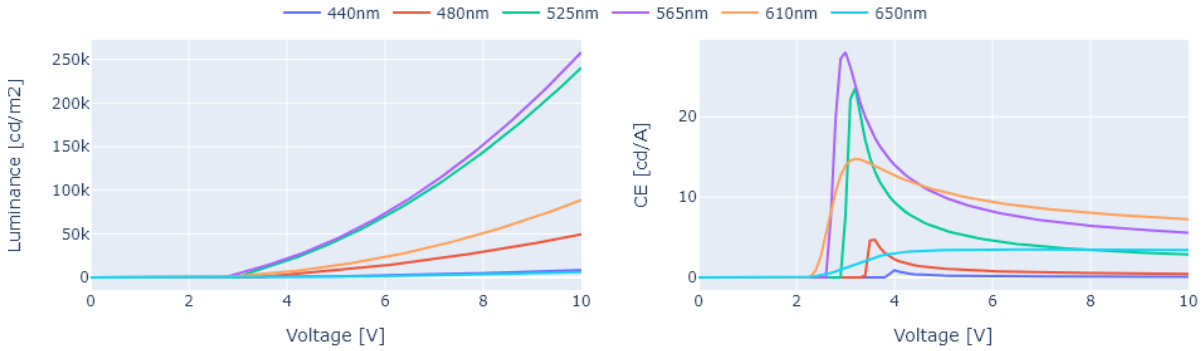
(a) The luminance is a function of both total current density (J_{tot}) and CE.

(b) Peak CE is achieved at approximately 4 V

(c) The total current density (J_{tot}) increases exponentially with increasing Voltage.

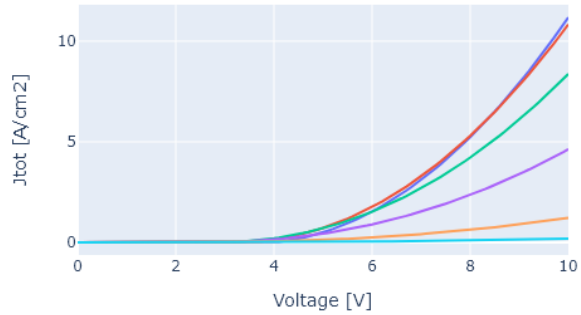
(d) The EQE trend is similar CE with a peak around 6%.

Fig. 6.11 Simulated optoelectronic properties of the 4-primary device.

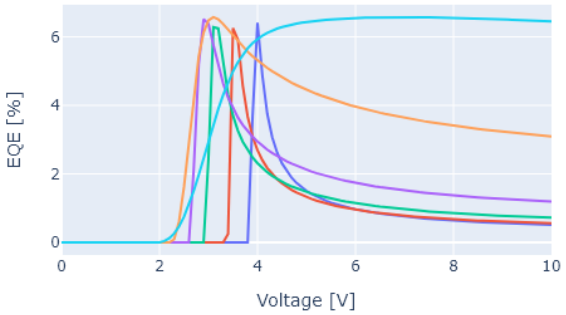


(a) The luminance is a function of both total current density (J_{tot}) and CE.

(b) The peak CE is achieved at approximately 4 V.

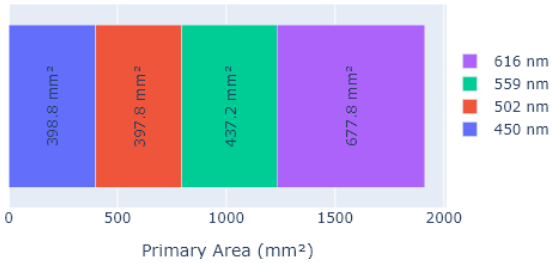


(c) The total current density (J_{tot}) increases exponentially with increasing Voltage.

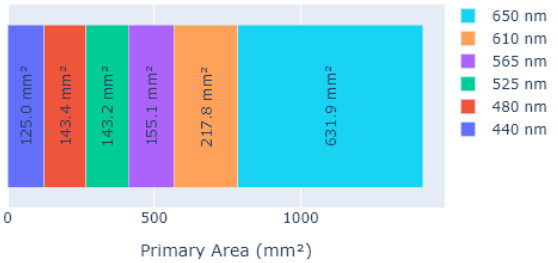


(d) The EQE has a similar trend to CE curves. Peak EQE is around 6%.

Fig. 6.12 Simulated optoelectronic properties of the 6-primary device.



(a) 4-primary device.



(b) 6-Primary device.

Fig. 6.13 Calculated sub-pixel areas for the two devices so that they achieve a colour match at 10 V.

6.5 QD Simulator Interface and Material Library

We also built a *Python* interface to the CTM simulator. This allows simulator users to easily define device properties through a hierarchical and modular interface. It also decouples the CTM simulator from the simulation definition, allowing features to be added to the simulator while maintaining backwards compatibility. We used the modularity of this setup to also lay the groundwork for developing a library of materials. The hierarchical device definitions would allow a single definition of a material to be re-used and adapted to fit new needs. For example, an interface layer is defined as given in fig. 6.15. It takes in a material parameter as defined in fig. 6.16. The interface layer is ambivalent to the material and as such can be easily swapped out with a new material, either manually or programmatically.

```python
ITO__anode = Anode(4.70)
AL__cathode = Cathode(4.06)

PEDOT_PSS__HIL = InterfaceLayer(
    Name='_HIL',
    Material=Material(Name='PDOT_PSS', Epsi=3.0,Mup=0.322e-3, Mun=0.322e-3, E0=5.0e6),
    Energy=Energy(X=3.6, Eg=1.57),
    Doping=Doping(Na=2.81e19, Nd=0.0),
    EffectiveMass=EffectiveMass(P=1.0, N=1.0),
    Thickness=Thickness(dz=1.0e-7, Ng=20),
    Recombination=Recombination(Tp=1.0e-6, Tn=11.0e-6, Cp=0.0e-31, Cn=0.0e-31)
)

TFB__HTL = InterfaceLayer(
    Name='TFB__HTL',
    Material=Material(Name='TFB', Epsi=3.5, Mup=2.0e-3, Mun=2.0e-3, E0=5.0e6),
    Energy=Energy(X=2.3, Eg=3.0),
    Doping=Doping(Na=1.0e17, Nd=0.0),
    EffectiveMass=EffectiveMass(P=1.0, N=1.0),
    Thickness=Thickness(dz=1.0e-7, Ng=20),
    Recombination=Recombination(Tp=1.0e-6, Tn=1.0e-6, Cp=0.0e-31, Cn=0.0e-31)
)

TiO2__ETL = InterfaceLayer(
    Name='TiO2__ETL',
    Material=Material(Name='TiO2', Epsi=8.5, Mup=2.0e-3, Mun=2.0e-3, E0=5.0e6),
    Energy=Energy(X=4.0, Eg=3.4),
    Doping=Doping(Na=0.0, Nd=1.0e17),
    EffectiveMass=EffectiveMass(P=1.21, N=0.29),
    Thickness=Thickness(dz=1.0e-7, Ng=40),
    Recombination=Recombination(Tp=1.0e-6, Tn=1.0e-6, Cp=0.0e-31, Cn=0.0e-31)
)

red_QD = QuantumDot(
    Name='QD_RED',
    Material=Material(Name='red_qd_material', Epsi=10.2,Mup=0.5e-6, Mun=0.5e-6, E0=5.0e6),
    Energy=Energy(X=3.695, Eg=1.86),
    FWHM=20,
    Diameter=8.0e-7,
    EffectiveMass=EffectiveMass(P=0.45, N=0.13),
    QDShell=QDShell(Thickness=0.5e-7, eMassP=0.6, eMassN=0.2, X=3.2, Eg=3.54),
    QDRecombination=QDRecombination(Gamma=0.50e-11, Tp=2.5e-6, Tn=2.5e-6, Cp=1.0e-30, Cn=1.0e-30),
    CaptureCrossSection=CaptureCrossSection(Sp=1.0, Sn=1.0)
)
```

Fig. 6.14 A code snippet showing how a complete QD-LED can be defined in with the python interface.

```python
class InterfaceLayer:
    def __init__(self, Name, Material, Energy, Doping,
                 EffectiveMass, Thickness, Recombination):

        self.Name = Name
        self.Material = Material
        self.Energy = Energy
        self.Doping = Doping
        self.EffectiveMass = EffectiveMass
        self.Thickness = Thickness
        self.Recombination = Recombination

    def __str__(self):
        return (
            f'Material {self.Material}\n'
            f'Energy {self.Energy}\n'
            f'Doping {self.Doping}\n'
            f'EffectiveMass {self.EffectiveMass}\n'
            f'Thickness {self.Thickness}\n'
            f'Recombination {self.Recombination}')

```

Fig. 6.15 Interface layers can be easily defined with its properties arranged in a hierarchical manner.

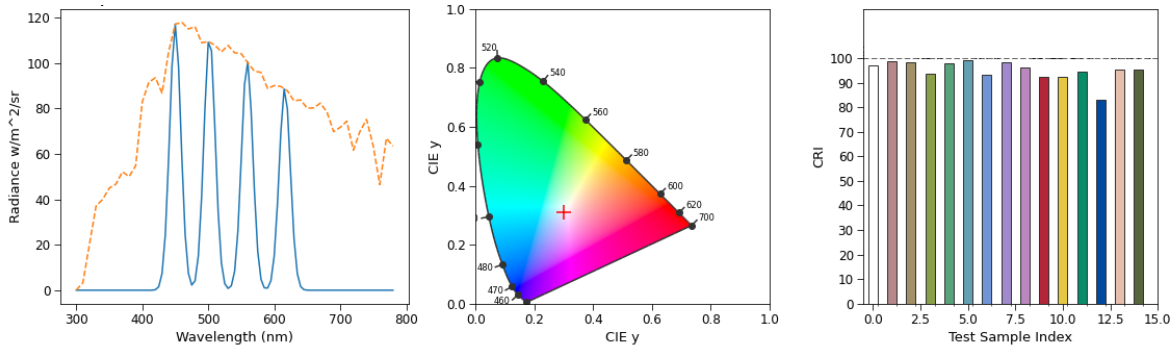
```python
class Material:
    def __init__(self, Name, Epsi, Mup, Mun, E0,
                 Mup_unit='cm2/Vs', Mun_unit='cm2/Vs',
                 E0_unit='V/cm'):

        self.Name = Name
        self.Epsi = Epsi
        self.Mup = Mup
        self.Mup_unit = Mup_unit
        self.Mun = Mun
        self.Mun_unit = Mun_unit
        self.E0 = E0
        self.E0_unit = E0_unit

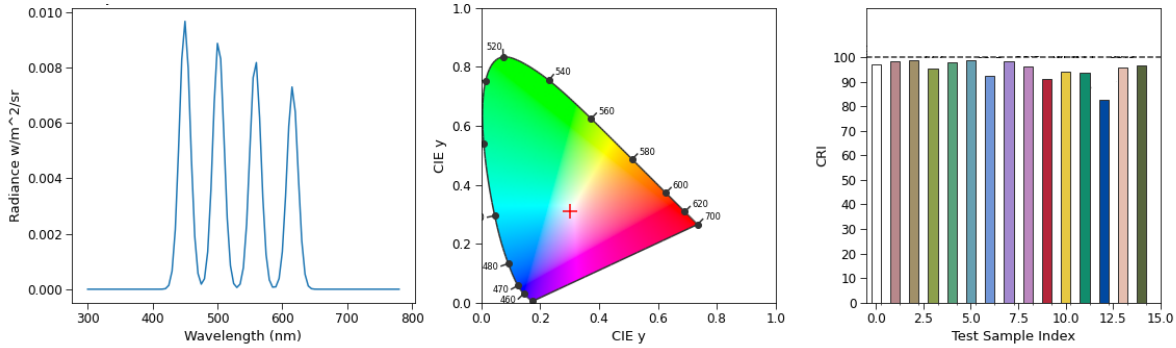
    def __str__(self):
        return f'Epsi={self.Epsi:.3f}, Mup={self.Mup:.3e}[{
self.Mup_unit}], Mun={self.Mun:.3e}[{self.Mun_unit}] at E0=
{self.E0:.3e}[{self.E0_unit}]'

```

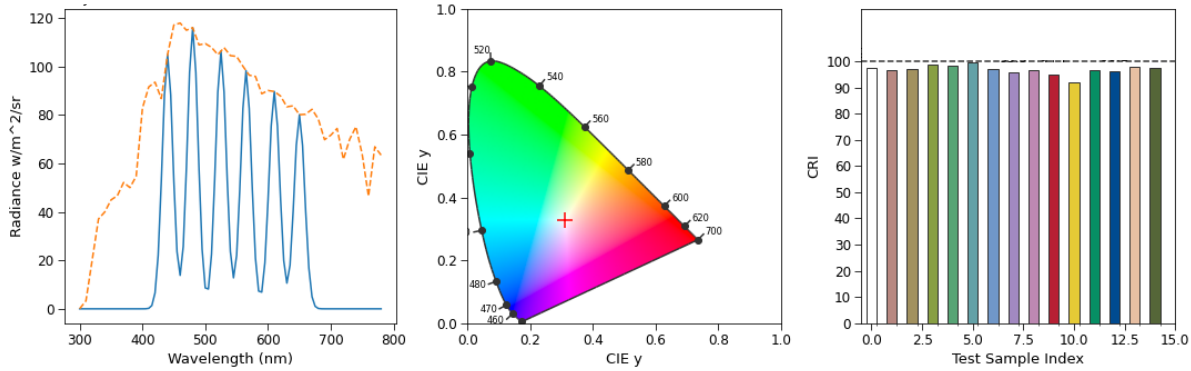
Fig. 6.16 Material definitions can be re-used and are agnostic to the application.



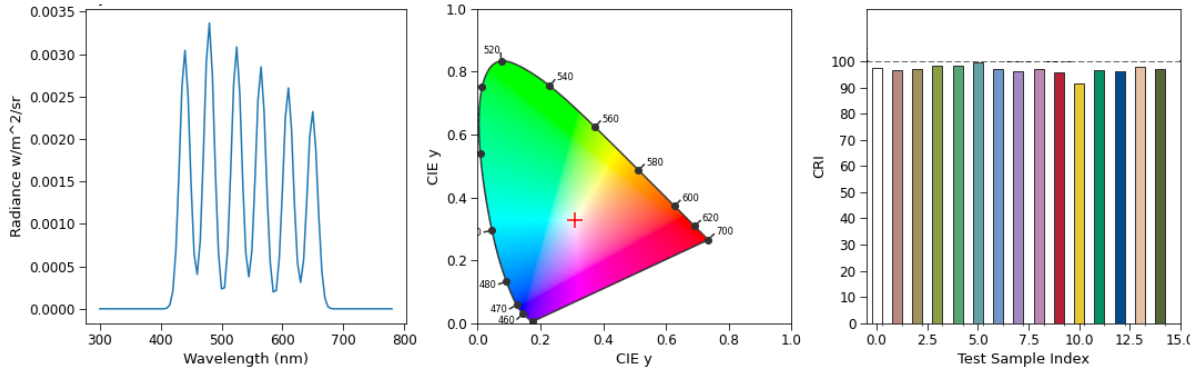
(a) Calculated high CRI spectrum for the 4-primary device.



(b) Simulated and rescaled spectrum for the 4-primary device.



(c) Calculated high CRI spectrum for the 6-primary device.



(d) Simulated and rescaled spectrum for the 6-primary device.

Fig. 6.17 The spectral properties of the simulated 4-primary and 6-primary devices match closely with their calculated spectral properties. Left to right, emission spectrum, chromaticity coordinates on the CIE-xy space, and the individual TCS CRIs.

6.6 Conclusions and Future Work

This chapter outlined how to use an optimisation based approach to design QD based lighting devices to achieve a high CRI. It showed that by following this methodology, we were able to design a QD based white light with a CRI of 96.9 using only 4 primary emissions. The chapter also presented results of 5 and 6 primary devices exhibiting even better minimum individual TCS CRIs. These devices were simulated using a CTM simulator and results of which were also presented for the 4 and 6 primary devices.

The next stage for this line of work would be to fabricate the simulated devices to compare the simulated performance with the performance of a real device. Given the flexibility of QDs, the proposed devices can be fabricated as patterned mode, stacked mode, or mixed mode devices. Patterned mode gives the best isolation between the sub-pixels with the potential for individual pixel control using a TFT layer. On the flip side, the stacked mode and mixed mode devices are much simpler to fabricate, which we hope will make the proposed design a cost effective alternative for current state of the art high-CRI illuminants. The practicality of these devices will also depend on the achievable current efficiencies and the device lifetimes which will need to be evaluated through device fabrications. Currently, arrangements are being made to fabricate these devices, but at the time of writing, we do not have any results from real devices.

As outlined before, the luminous efficacies of the designed devices were more than 50x smaller compared to commercial LED-based light bulbs. This is primarily due to low EQEs of the QDs. It is hoped that this would increase with developments in the field of QD materials. A potential design solution would be to evaluate the expected power consumption as part of the optimisation process by explicitly specifying ‘high luminous efficacy’ as one of the optimisation goals in the cost function. This would try to ensure that the primaries calculated by the optimiser are the most efficient combination in addition to being the combination that maximises the CRI and achieve a chromaticity match. However, this would require a model of the QD power as a function of the emission wavelength and intensity. We believe that the CTM can be used to generate the required data to create a lookup table of device power as a function of luminance and wavelength. This can then be used to facilitate the power efficiency optimisation.

A doctoral student and her advisor walk into a bar. The advisor orders a rough draft... and they sit in awkward silence for eight months.

— Anonymous.

Chapter 7

Conclusions and Future Work

The core of the work presented in this dissertation explored how to use the limited colour acuity of human vision and advantages of QDs, to achieve power savings for QD based displays.

In chapter 3, we explored global colour vision statistics through a dataset of global colour matching data from a mobile video game. We presented how the players' vision acuity was affected by different colour properties. We showed how this video game could be thought to be analogous to traditional colour matching experiments and how to use this fact to calculate individual player CMFs from this data. After evaluating the feasibility of the numerical optimisation process, we showed how to use a strong physiologically accurate prior to extract sensible estimates for the CMFs.

The key contribution of this chapter is the derivation and evaluation of a method for extracting individual observer CMFs from a colour matching experiment carried out on a digital display. We demonstrate that this method is able to calculate realistic CMFs using the colour mismatches observed during the experiment and a simple numerical optimisation method. The use of a physiologically relevant prior in this optimisation process allows this method to achieve realistic results. The work presented in this chapter can be used as the first stage (user characterisation) in a perception aware optimisation framework where digital colour matching data can be used to infer a user's spectral sensitivities. However, before these results can be used in practice, they will need to be validated through user studies with traditional colour matching experiments to evaluate whether the learned CMFs match the user's CMFs.

To that end we propose the following avenues for future investigations:

1. Minimising Epistemic uncertainty: The some of the key sources of uncertainty in the data from the *Specimen* dataset are on the display model and the user age. To resolve the first, an approach similar to what is outlined in appendix A could be used to measure the power dynamics of an *iPhone X* display. For the latter, the information could potentially be obtained from the authors of the game. The evaluation presented in chapter 3 should be repeated with the updated data. The output from this could be used in the user study to facilitate more robust measurements.
2. Exploring numerical instabilities and gradient based optimisation: As outlined in chapter 3, we were unable to use gradient based optimisation methods and some results showed numerical

instabilities. A detailed exploration into the sources of these instabilities and potential fixes to improve the stability of the process is recommended.

3. Method validation with a user study: To verify the CMFs calculated by the method are accurate, user studies will need to be carried out, involving both digital and traditional colour matching experiments. The results of the digital experiment (done analogous to the *Specimen* game) will be used to infer the CMFs using the presented method, and compared against the CMFs obtained from the traditional experiment, which serves as the ground truth.
4. Optimising colour sample presentation: The work was based on the *Specimen* dataset and the colours it presented during play. However, if one were to pick these colours more intelligently, the the proposed system would be able to learn the user's spectral sensitivities much faster, more robustly, while reducing the number of required colour mismatches. This avenue would involve investigating how to use Bayesian optimisation or a similar approach to dynamically pick which colours are shown to the user such that the information gain about their spectral response (i.e., minimise the uncertainty in the CMFs).
5. Publishing the Specimen data: Currently the dataset used in this work is not available publicly. With the author's permission and ethics approval, this data could be made available publicly. A key point of contention is the usefulness of the information about microtransaction events and potential harm they could do. Thus a rigorous ethical approval process is recommended.

Chapter 4 showed how to use metamerism to achieve a significant display power saving while generating the same colour sensation and highlighted how an MPD architecture could be used to enable the spectral flexibility required to leverage metamerism in a real device. We presented a mathematical description of an MPD and presented a pair of indices for measuring how well a given display can leverage metamerism; one for metameric capacity and the other for how much power saving can be achieved using metamerism. Although the work presented used standard observer CMFs, in a perception-aware framework, the output of the user characterisation can be used instead to create optimisations specific to each user.

The takeaway from this work is that, *if* display pixels with increased spectral flexibility can be made, a substantial amount of power saving could be achieved by dynamically picking the most optimal spectral shapes for each colour sensation to leverage the limited colour acuity of human vision and the wavelength dependent power consumption of the pixels. We show that, in an ideal case, an order of magnitude lower power can be achieved. The multi-primary architecture presented in this chapter is a way to achieve this spectral flexibility and the indices presented allows one to quantify the flexibility of any given multi-primary display with respect to a target colour space.

For future investigations based on this work, we suggest the following directions:

1. Metamere search with more flexible spectral shapes: One of the limits of the metamere search presented in this chapter is that it only examined hypothetical displays with three primaries. Re-running this search with four or more primaries would allow for a better understanding of the metamere landscape and the potential energy savings of an MPD. However, this would

probably require an optimisation based method as opposed to a grid search due to the increased dimensionality of the search space.

2. Evaluate the metameric indices: The metameric indices presented in this work have not been evaluated on hypothetical MPDs. Validating these indices by using real or hypothetical MPDs would be required to assess the usefulness of the indices.
3. Fabricate a multi primary display: All the work presented in chapter 4 have been based on numerical analysis of hypothetical displays. To concretely assess the hardware overheads of an MPDs, such a device will need to be designed and fabricated using standard display fabrication methodologies. The output of the two directions outlined above could be used for this to inform the optimal design parameters.
4. Explore QD anisotropy: In section 4.3 we briefly outlined an idea for using QD polarisation or piezoelectricity for modulating the emission spectrum of a QD. Although this was not explored in detail due to the reasons outlined, it would be an interesting avenue to explore for changing the emission spectrum of a QD sub-pixel after device fabrication.

In chapter 5 we highlighted an algorithmic method reported in the literature for achieving display power savings by imperceptibly changing the colours of the content shown on it. We carried out two user studies to evaluate the effectiveness of this method and presented their results, highlighting that up to 50% power saving can be achieved with 'acceptable' quality degradation. Using the user study results, we showed how to efficiently calculate a parameter lower bound that is a function of the colour properties of the image being transformed to find the optimal transform parameters for a given image. The chapter also outlined how this modified framework can be used in an end-to-end application.

The method presented in this chapter allows the users of *Crayon* to intelligently pick the optimal transform parameters for each image by specifying the minimum quality score. The method uses the models trained with the user study data to find a trade-off between maximising power saving while achieving the target minimum quality score, which was not possible with the original *Crayon* formulation. This is further supported by a parameter normalisation scheme which allows transform parameters to be made independent of the display power model, improving the inter-operability of Crayon.

We propose the following action points for future work in this area:

1. Detailed evaluation of computational cost: The key unanswered question in this chapter is the computational cost of calculating the optimal parameters. The current implementation relies on an offline optimisation stage. Thus, the priority should be on creating such an optimised implementation.
2. Develop *Crayon* as a easy-to-use library: Currently there exists not implementation of *Crayon* that potential users can easily integrate into their mobile applications. Creating such a library, either as a standalone implementation or an extension to the Android OS, would allow the benefits of *Crayon* to be leveraged by potential users and improve its impact.
3. Larger user study: Although the results of the user study used in this work have been sufficient to develop useful models for predicting optimal transform parameters, a larger user study with

more participants and a wider diversity of images would allow these models to be more accurate and robust. Additionally, a richer set of results would allow the models to be modified to output confidence bounds on the potential score distribution of an unseen image.

4. Use in an MPD architecture: Although the work presented in this chapter assumed a target device with three primaries, it can also be applied to an MPD architecture. An exploration of how Crayon can work on an MPD and benefit from the power savings outlined in chapter 4 would be another avenue for future work, bringing together the findings of the different chapters to maximise their impact.

Finally, in chapter 6 we presented an optimisation based approach for calculating optimised designs for QD based devices. We showed how to use this method to calculate the device parameters for a set of high CRI QD lights. We show that through an MPD-like architecture, we can achieve a CRI of up to 97. We simulated the electrical performance of these devices through a CTM simulator and show how the radiance can be modulated by adjusting the 'pixel' area for each primary in a patterned mode QD device.

We propose two avenues for future work:

1. Fabricating the designed devices: Although the design was shown to be potentially groundbreaking, for it to be validated, it will need to be fabricated using standard display fabrication techniques. Fabricating the designed devices and assessing their performance against the simulation results would allow us to validate the designs and the effect of fabrication tolerances.
2. Modifying the cost function to include efficiency: The devices presented in this chapter have a much lower efficiency compared to commercial LED-based devices. A potential solution for this would be to evaluate whether incorporating device efficiency as an optimisation criterion in the cost functions would lead to improved device efficiency.

We believe that there is an opportunity to create better optoelectronic devices and systems that interact with humans by allowing them to be aware of the limits of human colour vision. This body of work has been successful in exploring some of these limits and evaluating how they could be used to inform design decisions of display and lighting systems, and presented the results of investigations into several potential avenues for optimisation. We hope that the results showcased in this dissertation and the tools developed during the course the work presented in it will be helpful in facilitating future research into this area.

References

[1] S.-M. Jung, T. H. Lee, S. Y. Bang, S. D. Han, D.-W. Shin, S. Lee, H. W. Choi, Y.-H. Suh, X.-B. Fan, J.-W. Jo, S. Zhan, J. Yang, C. Samarakoon, Y. Kim, L. G. Occhipinti, G. Amaratunga, and J. M. Kim, "Modelling charge transport and electro-optical characteristics of quantum dot light-emitting diodes," *npj Computational Materials*, vol. 7, Jul 2021.

[2] K. Kumar and contributors, "kks32/phd-thesis-template: A latex / xelatex / lualatex phd thesis template for cambridge university engineering department (cued)," Dec 2020.

[3] E. Carter, J. Schanda, R. Hirschler, S. Jost, M. Luo, M. Melgosa, Y. Ohno, M. Pointer, D. Rich, F. Vienot, L. Whitehead, and J. Wold, "CIE 015:2018 Colorimetry, 4th Edition," tech. rep., International Commission on Illumination (CIE), Oct. 2018.

[4] Colour Developers, "Colour Science for Python." Available via Github and PyPi, Feb. 2021.

[5] OpenStax College, "Anatomy of the Eye," May 2013. Released under Creative Commons Attribution 3.0 Unported license. Available through Wikimedia Commons.

[6] H. Fischer, "Schematic Diagram of the Human Eye," 31 January 2013. Released under Creative Commons Attribution 3.0 Unported license. Available through Wikimedia Commons.

[7] T. Goodman, "Light Emission and Photometry," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 217–228, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[8] A. Stockman and L. Sharpe, "Cone Fundamentals." Colour and Vision Research laboratory, UCL. Available at http://www.cvrl.org/cones.htm, 2000.

[9] Wikimedia Commons, "CIE1931 RGB Coordinate Space Plot." In the Public Domain. Available through Wikimedia Commons.

[10] J. Schanda, *Colorimetry: Understanding the CIE System*. Hoboken, NJ, USA: John Wiley & Sons, Inc., July 2007.

[11] SharkD, "RGB Cube Low-gamma Cutout." Available under Creative Commons Attribution-Share Alike 3.0 Unported License through Wikimedia Commons, Mar. 2010.

[12] SharkD, "HSV Color Solid Cylinder Saturation Gray." Available under Creative Commons Attribution-Share Alike 3.0 Unported License through Wikimedia Commons, Mar. 2010.

[13] D.-M. Kim, *Introductory Quantum Mechanics for Semiconductor Nanotechnology*. Weinheim: Wiley-VCH, 2015.

[14] Oak Ridge National Laboratory, "Demystifying Quantum Dot Conundrums." Report available at https://www.olcf.ornl.gov/2015/05/05/demystifying-quantum-dot-conundrums/, May 2015.

[15] V. I. Klimov, *Nanocrystal Quantum Dots*. Boca Raton: CRC Press, second edition. ed., 2010.

[16] Sigma-Aldrich, "Photoluminescence in Alloyed Quantum Dots." Report available at https://www.sigmaaldrich.com/technical-documents/articles/materials-science/nanomaterials/quantum-dots.html.

[17] LG, "OLED Introduction." Avaliable at https://www.lg.com/global/business/technology-solution-oled.

[18] T. Tsujimura, *OLED Display: Fundamentals and Applications : Fundamentals and Applications*. Hoboken, United States of America: John Wiley & Sons, Incorporated, 2012.

[19] DisplayMate Technologies, "iPhone-XOLED Display Technology Shoot-Out," 2017.

[20] Lumileds Holding, "DS105 LUXEON Z Color Line Product Datasheet," 2017.

[21] P. Stanley-Marbell, V. Estellers, and M. Rinard, "Crayon: Saving Power through Shape and Color Approximation on next-Generation Displays," in *Proceedings of the Eleventh European Conference on Computer Systems*, EuroSys '16, (New York, NY, USA), Association for Computing Machinery, 2016.

[22] Y. Asano, M. D. Fairchild, and L. Blondé, "Individual Colorimetric Observer Model," *PLOS ONE*, vol. 11, p. e0145671, Feb. 2016.

[23] G. Moore, "Cramming More Components Onto Integrated Circuits," *Proceedings of the IEEE*, vol. 86, pp. 82–85, Jan. 1998.

[24] R. Y. Pai, "Efficiency Limits for Fluorescent Lamps and Application to LCD Backlighting," *Journal of the Society for Information Display*, vol. 5, no. 4, pp. 371–374, 1997.

[25] B. Geffroy, P. le Roy, and C. Prat, "Organic Light-Emitting Diode (OLED) Technology: Materials, Devices and Display Technologies," *Polymer International*, vol. 55, no. 6, pp. 572–582, 2006.

[26] A. Salehi, X. Fu, D.-H. Shin, and F. So, "Recent Advances in OLED Optical Design," *Advanced Functional Materials*, vol. 29, no. 15, p. 1808803, 2019.

[27] 4D Systems, "Introduction to OLED displays design guide for active matrix OLED (AMOLED)," tech. rep., 2008.

[28] Qualcomm, "Competitive Display Technologies," White Paper, June 2009.

[29] S. Reineke, "Complementary LED Technologies," *Nature Materials*, vol. 14, pp. 459–462, May 2015.

[30] ITU Radiocommunication Bureau, "ITU-R BT.2020. Parameter Values for Ultra-High Definition Television Systems for Production and International Programme Exchange," tech. rep., International Telecommunication Union, 2012.

[31] ITU Radiocommunication Bureau, "ITU-R BT.2100. Image Parameter Values for High Dynamic Range Television for Use in Production and International Programme Exchange," tech. rep., 2016.

[32] VESA, "VESA DisplayHDR specification : Performance criteria 1.1." Available at https://displayhdr.org/performance-criteria-cts1-1/, 2020.

[33] D. R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*. London: Penguin, 20th-anniversary ed., 2000.

[34] D. A. McIntyre, *Colour Blindness: Causes and Effects*. Dalton Publishing, 2002.

[35] J. J. Koenderink, *Color for the Sciences*. Cambridge, Mass: MIT Press, 2010.

[36] M. Bloj and M. Hedrich, "Color Perception," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 171–178, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[37] D. Bohm, *Quantum Theory*. Prentice-Hall Physics Series, Englewood Cliffs, N. J: Prentice-Hall, 16. print ed., 1951.

[38] L. Carroll, *Alice's Adventures in Wonderland*. Rosalind Series, W.B. Conkey, 1920.

[39] M. D. Fairchild, *Color Appearance Models*. The Wiley-IS&T Series in Imaging Science and Technology, Chichester, West Sussex: John Wiley & Sons, Inc, third edition ed., 2013.

[40] Z. Li, *Understanding Vision: Theory, Models, and Data*. Oxford, United Kingdom ; New York, NY, United States of America: Oxford University Press, first edition ed., 2014.

[41] D. J. Felleman, "Visual System in the Brain," in *International Encyclopedia of the Social & Behavioral Sciences* (N. J. Smelser and P. B. Baltes, eds.), pp. 16278–16285, Oxford: Pergamon, Jan. 2001.

[42] M. Bass and Optical Society of America, eds., *Handbook of Optics*. New York: McGraw-Hill, 2nd ed ed., 1995.

[43] G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data, and Formulae*. Wiley Classics Library, New York: John Wiley & Sons, wiley classics library ed ed., 2000.

[44] W. D. Wright, "A Re-Determination of the Trichromatic Coefficients of the Spectral Colours," *Transactions of the Optical Society*, vol. 30, no. 4, pp. 141–164, 1929.

[45] J. Guild, "The Colorimetric Properties of the Spectrum," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 230, pp. 149–187, June 1931.

[46] CIE, "Commission Internationale de l'Eclairage Proceedings, 1931," (Cambridge), Cambridge University Press, 1932.

[47] W. S. Stiles and J. M. Burch, "N.P.L. Colour-Matching Investigation: Final Report," *Optica Acta: International Journal of Optics*, vol. 6, pp. 1–26, Jan. 1959.

[48] N. I. Speranskaya, "Determination of Spectrum Color Coordinates for Twenty-Seven Normal Observers," *Optics and Spectroscopy*, vol. 7, pp. 424–429, 1959.

[49] A. Stockman and L. T. Sharpe, "The Spectral Sensitivities of the Middle- and Long-Wavelength-Sensitive Cones Derived from Measurements in Observers of Known Genotype," *Vision Research*, vol. 40, pp. 1711–1737, June 2000.

[50] A. Stockman, L. T. Sharpe, and C. Fach, "The Spectral Sensitivity of the Human Short-Wavelength Sensitive Cones Derived from Thresholds and Color Matches," *Vision Research*, vol. 39, pp. 2901–2927, Aug. 1999.

[51] CIE Tchnical Committee TC 1-36, "Fundamental Chromaticity Diagram with Physiological Axes. Part 1," Tech. Rep. 170,1, CIE Central Bureau, Vienna, 2006.

[52] D. L. MacAdam, "Visual Sensitivities to Color Differences in Daylight," *Journal of the Optical Society of America*, vol. 32, pp. 247–274, May 1942.

[53] S. Westland and V. Cheung, "RGB Systems," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 147–153, Berlin, Heidelberg: Springer, 2012.

[54] ITU Radiocommunication Bureau, "BT.709:Parameter Values for the HDTV Standards for Production and International Programme Exchange," Recommendation ITU-R BT.709-6, International Telecommunication Union, June 2015.

[55] M. R. Krames, "Light Emitting Diodes: Fundamentals," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1155–1167, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[56] R. Ma, "Organic Light Emitting Diodes (OLEDs)," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1209–1222, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[57] H. Sasabe and J. Kido, "Development of High Performance OLEDs for General Lighting," *Journal of Materials Chemistry C*, vol. 1, pp. 1699–1707, Feb. 2013.

[58] W. W. Chow and F. Jahnke, "On the Physics of Semiconductor Quantum Dots for Applications in Lasers and Quantum Optics," *Progress in Quantum Electronics*, vol. 37, pp. 109–184, May 2013.

[59] S. B. Brichkin and V. F. Razumov, "Colloidal Quantum Dots: Synthesis, Properties and Applications," *Russian Chemical Reviews*, vol. 85, p. 1297, Dec. 2016.

[60] J. Ko, J. H. Chang, B. G. Jeong, H. J. Kim, J. F. Joung, S. Park, D. H. Choi, W. K. Bae, and J. Bang, "Direct Photolithographic Patterning of Colloidal Quantum Dots Enabled by UV-Crosslinkable and Hole-Transporting Polymer Ligands," *ACS Applied Materials & Interfaces*, vol. 12, pp. 42153–42160, Sept. 2020.

[61] L. Kim, P. O. Anikeeva, S. A. Coe-Sullivan, J. S. Steckel, M. G. Bawendi, and V. Bulović, "Contact Printing of Quantum Dot Light-Emitting Devices," *Nano Letters*, vol. 8, no. 12, pp. 4513–4517, 2008. PMID: 19053797.

[62] T.-H. Kim, K.-S. Cho, E. K. Lee, S. J. Lee, J. Chae, J. W. Kim, D. H. Kim, J.-Y. Kwon, G. Amaratunga, S. Y. Lee, B. L. Choi, Y. Kuk, J. M. Kim, and K. Kim, "Full-Colour Quantum Dot Displays Fabricated by Transfer Printing," *Nature Photonics*, vol. 5, pp. 176–182, Mar. 2011.

[63] S. Zhang, S. Jiao, Y. Li, W. Yuan, S. Zhang, S. Iguchi, Y. Wu, and H. Zhou, "Efficient Top-Emitting Quantum Dot Light Emitting Diodes via Inkjet Printing," in *2019 IEEE 2nd International Conference on Electronics Technology (ICET)*, pp. 423–426, May 2019.

[64] P. Yang, L. Zhang, D. J. Kang, R. Strahl, and T. Kraus, "High-Resolution Inkjet Printing of Quantum Dot Light-Emitting Microdiode Arrays," *Advanced Optical Materials*, vol. 8, no. 1, p. 1901429, 2020.

[65] P. O. Anikeeva, J. E. Halpert, M. G. Bawendi, and V. Bulović, "Quantum Dot Light-Emitting Devices with Electroluminescence Tunable over the Entire Visible Spectrum," *Nano Letters*, vol. 9, pp. 2532–2536, July 2009.

[66] Samsung Global, "Samsung Premium UHD QLED TV catalogue." Catalogue available at https://www.samsung.com/uk/tvs/qled-tv/.

[67] T. Ishida, Y. Nakanishi, M. Izumi, and V. Berryman-Bousquet, "How Will Quantum Dots Enable Next-Gen Display Technologies?," *Information Display*, vol. 36, no. 6, pp. 14–18, 2020.

[68] H. Hong, "In-Plane Switching (IPS) Technology," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1469–1483, Berlin, Heidelberg: Springer, 2012.

[69] P. Raynes, "Twisted Nematic and Supertwisted Nematic LCDs," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1433–1444, Berlin, Heidelberg: Springer, 2012.

[70] H. Yoshida, "Vertically Aligned Nematic (VAN) LCD Technology," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1485–1505, Berlin, Heidelberg: Springer, 2012.

[71] M. Klasen-Memmer and H. Hirschmann, "Liquid Crystal Materials for Devices," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1315–1342, Berlin, Heidelberg: Springer, 2012.

[72] G. Boyd, "LCD Backlights," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1609–1623, Berlin, Heidelberg: Springer, 2012.

[73] R. Thielemans, "LED Display Applications and Design Considerations," in *Handbook of Visual Display Technology* (J. Chen, W. Cranton, and M. Fihn, eds.), pp. 1169–1180, Berlin, Heidelberg: Springer, 2012.

[74] J. Guild, "The Geometrical Solution of Colour Mixture Problems," *Transactions of the Optical Society*, vol. 26, pp. 139–174, Feb. 1925.

[75] E. Gorochow, C. Whitney, S. Randazzo, and PepRally, "Specimen: An iPhone Game about Color," 2015.

[76] K. St Clair, "The Science of Colour Is Upending Our Relationship with Screens," *Wired UK*, May 2020.

[77] G. E. P. Box, "Robustness in the Strategy of Scientific Model Building," in *Robustness in Statistics* (R. L. Launer and G. N. Wilkinson, eds.), pp. 201–236, Academic Press, Jan. 1979.

[78] J. Cambronero, P. Stanley-Marbell, and M. Rinard, "Incremental Color Quantization for Color-Vision-Deficient Observers Using Mobile Gaming Data," Mar. 2018.

[79] "TinyPNG – Compress PNG Images While Preserving Transparency." Accessible via https://tinypng.com/.

[80] "Pngquant – Lossy PNG Compressor." Available via Github.

[81] P. Stanley-Marbell and M. Rinard, "Perceived-Color Approximation Transforms for Programs That Draw," *IEEE Micro*, vol. 38, pp. 20–29, July 2018.

[82] Statista, "UK: Market Share Apple iOS 2011-2019." Report available at https://www.statista.com/statistics/271195/apple-ios-market-share-in-the-united-kingdom-uk/.

[83] P. Chandro, "Countryinfo." Available via Github and PyPi, Nov. 2020.

[84] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder, and P. Morvan, "Modern Displays: Why We See Different Colors, and What It Means?," in *2010 2nd European Workshop on Visual Information Processing, EUVIP2010*, pp. 1–6, IEEE, July 2010.

[85] M. Ebner, *Color Constancy*. Wiley-IS&T Series in Imaging Science and Technology, Chichester: John Wiley, 2007.

[86] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, International Conference on Learning Representations, ICLR, Dec. 2015.

[87] A. Stockman and U. Color and Vision Research Lab, "CVRL Colour Matching Functions Database." Colour and Vision Research laboratory, UCL. Data available at http://www.cvrl.org/, 2020.

[88] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[89] A. N. Tikhonov, V. Y. Arsenin, and R. A. Willoughby, "Solutions of Ill-Posed Problems," *SIAM Review*, vol. 21, pp. 266–267, Apr. 1979.

[90] A. Stockman, "Cone Fundamentals and CIE Standards," *Current Opinion in Behavioral Sciences*, vol. 30, pp. 87–93, Dec. 2019.

[91] H. R. Wilson, "A transducer function for threshold and suprathreshold human vision," *Biological Cybernetics*, vol. 38, p. 171–178, Oct 1980.

[92] J. Ross, H. D. Speed, and F. W. Campbell, "Contrast adaptation and contrast masking in human vision," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 246, no. 1315, pp. 61–70, 1991.

[93] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Am.*, vol. 70, pp. 1458–1471, Dec 1980.

[94] C. Yu and D. M. Levi, "Surround modulation in human vision unmasked by masking experiments," *Nature Neuroscience*, vol. 3, p. 724–728, Jul 2000.

[95] M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial-frequency channels in human vision∗," *J. Opt. Soc. Am.*, vol. 61, pp. 1176–1186, Sep 1971.

[96] R. Watt and M. Morgan, "A theory of the primitive spatial code in human vision," *Vision Research*, vol. 25, no. 11, pp. 1661–1674, 1985.

[97] J. Pokorny, V. C. Smith, and M. Lutze, "Aging of the Human Lens," *Applied Optics*, vol. 26, p. 1437, Apr. 1987.

[98] J. Moreland and E. Alexander, "Effect of Macular Pigment on Colour Matching with Field Sizes in the 1° to 10° Range," in *Colour Vision Deficiencies XIII* (C. R. Cavonius, ed.), vol. 59, pp. 363–368, Dordrecht: Springer Netherlands, 1997.

[99] J. Pokorny and V. C. Smith, "Effect of Field Size on Red–Green Color Mixture Equations," *Journal of the Optical Society of America*, vol. 66, p. 705, July 1976.

[100] A. Stockman, "CIE-1964 10-degree CMFs." Colour and Vision Research laboratory, UCL. Data available at http://www.cvrl.org/database/text/cienewxyz/cie2012xyz10.htm, 2012.

[101] S. C. Endres, C. Sandrock, and W. W. Focke, "A Simplicial Homology Algorithm for Lipschitz Optimisation," *Journal of Global Optimization*, vol. 72, pp. 181–217, Oct. 2018.

[102] I. M. Sobol, "On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 86–112, Jan. 1967.

[103] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, vol. 7, pp. 308–313, Jan. 1965.

[104] Y. Xiang, D. Y. Sun, W. Fan, and X. G. Gong, "Generalized Simulated Annealing Algorithm and Its Application to the Thomson Model," *Physics Letters A*, vol. 233, pp. 216–220, Aug. 1997.

[105] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing*, vol. 16, pp. 1190–1208, Sept. 1995.

[106] D. Kraft, "A Software Package for Sequential Quadratic Programming," tech. rep., DFVLR, Braunschweig, 1988.

[107] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, J. Stauder, and P. Morvan, "How Modern Displays Push Conventional Colorimetry to its Limit," in *Colour in art, science, design, conservation, research, printmaking, digital technologies, textiles (CREATE)*, (Norway), June 2010.

[108] B. Oicherman, M. R. Luo, B. Rigg, and A. R. Robertson, "Adaptation and colour matching of display and surface colours," *Color Research & Application*, vol. 34, no. 3, pp. 182–193, 2009.

[109] B. Oicherman, M. R. Luo, B. Rigg, and A. R. Robertson, "Effect of observer metamerism on colour matching of display and surface colours," *Color Research & Application*, vol. 33, no. 5, pp. 346–359, 2008.

[110] J. Fang and Y. J. Kim, "78-2: A matrix-based method of color correction for metamerism failure between lcd and oled," *SID Symposium Digest of Technical Papers*, vol. 49, no. 1, pp. 1044–1047, 2018.

[111] H. Xie, S. P. Farnand, and M. J. Murdoch, "Observer metamerism in commercial displays," *J. Opt. Soc. Am. A*, vol. 37, pp. A61–A69, Apr 2020.

[112] P. Pardo, A. Pérez, and M. Suero, "A new colour vision test in a pc-based screening system," *Displays*, vol. 21, no. 5, pp. 203–206, 2000.

[113] Y. Park, C.-M. Yang, D. Lee, C.-W. Kim, H. S. Lee, S. Park, and D. Kim, "Estimation of color matching functions for tiled lcds based on genetic algorithm," *Electronic Imaging*, vol. 2016, no. 20, pp. 1–4, 2016.

[114] S. M. Sze and M. K. Lee, *Semiconductor Devices, Physics and Technology*. Hoboken, N.J: Wiley, 3rd ed ed., 2012.

[115] X. Luo and R. Hu, "Chip Packaging: Encapsulation of Nitride LEDs," in *Nitride Semiconductor Light-Emitting Diodes (LEDs)* (J. Huang, H.-C. Kuo, and S.-C. Shen, eds.), Woodhead Publishing Series in Electronic and Optical Materials, pp. 491–528, Woodhead Publishing, second edition ed., Jan. 2018.

[116] J. K. Rowling, *Harry Potter and the Deathly Hallows*. London: Bloomsbury, 1st ed ed., 2007.

[117] M. Dong, Y.-S. K. Choi, and L. Zhong, "Power-Saving Color Transformation of Mobile Graphical User Interfaces on OLED-Based Displays," in *Proceedings of the 2009 ACM/IEEE International Symposium on Low Power Electronics and Design*, ISLPED '09, (New York, NY, USA), pp. 339–342, Association for Computing Machinery, 2009.

[118] T. K. Wee and R. K. Balan, "Adaptive Display Power Management for OLED Displays," *SIGCOMM Computer Communication Review*, vol. 42, p. 485–490, Sept. 2012.

[119] M. Dong and L. Zhong, "Chameleon: A Color-Adaptive Web Browser for Mobile OLED Displays," *IEEE Transactions on Mobile Computing*, vol. 11, pp. 724–738, May 2012.

[120] B. Anand, L. Kecen, and A. L. Ananda, "PARVAI – HVS Aware Adaptive Display Power Management for Mobile Games," in *2014 Seventh International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, pp. 21–26, 2014.

[121] D. Li, A. H. Tran, and W. G. J. Halfond, "Making Web Applications More Energy Efficient for OLED Smartphones," in *Proceedings of the 36th International Conference on Software Engineering*, ICSE 2014, (New York, NY, USA), pp. 527–538, Association for Computing Machinery, 2014.

[122] J. Fig, *Inside the Painter's Studio*. Chronicle Books, June 2012.

[123] M. R. Luo, G. Cui, and B. Rigg, "The Development of the CIE 2000 Colour-Difference Formula: CIEDE2000," *Color Research & Application*, vol. 26, no. 5, pp. 340–350, 2001.

[124] The University of Southern California Signal and Image Processing Institute, "The USC-SIPI Image Database." Dataset available at http://sipi.usc.edu/database/.

[125] Amazon Inc., "Amazon Mechanical Turk." Accessible via https://www.mturk.com/.

[126] F. Long, V. Ganesh, M. Carbin, S. Sidiroglou, and M. Rinard, "Automatic Input Rectification," in *Proceedings - International Conference on Software Engineering*, pp. 80–90, 2012.

[127] W. Mason and S. Suri, "Conducting Behavioral Research on Amazon's Mechanical Turk," *Behavior Research Methods*, vol. 44, pp. 1–23, Mar. 2012.

[128] Amazon Inc., "Amazon SageMaker." Accessible via https://aws.amazon.com/sagemaker/.

[129] Unsplash, "Beautiful Free Images & Pictures," 2020.

[130] G. Albaum, "The Likert Scale Revisited," *Market Research Society. Journal.*, vol. 39, pp. 1–21, Mar. 1997.

[131] R. Likert, "A Technique for the Measurement of Attitudes," *Archives of Psychology*, vol. 22 140, pp. 55–55, 1932.

[132] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, Apr. 2004.

[133] S. Feng, "Color-Thief-Py." Available via Github and PyPi, Oct. 2020.

[134] Z. Li and C. G. Bampis, "Recover Subjective Quality Scores from Noisy Measurements," in *2017 Data Compression Conference (DCC)*, pp. 52–61, Apr. 2017.

[135] Netflix, Inc., "Sureal: Subjective Recovery Analysis." Available via Github and PyPi, Oct. 2020.

[136] P. Felzenszwalb and D. Huttenlocher, "Efficient Graph-Based Image Segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.

[137] E. Wesemann, *Looking Tall by Standing Next to Short People: Other Techniques for Managing a Law Firm*. AuthorHouse, Mar. 2007.

[138] E. F. Schubert and J. K. Kim, "Solid-State Light Sources Getting Smart," *Science*, vol. 308, pp. 1274–1278, May 2005.

[139] W. Bae, Y. S. Park, J. Lim, D. Lee, L. A. Padilha, H. McDaniel, I. Robel, C. Lee, J. M. Pietryga, and V. I. Klimov, "Controlling the Influence of Auger Recombination on the Performance of Quantum-Dot Light-Emitting Diodes," *Nature Communications*, vol. 4, pp. 2661–2661, Jan. 2013.

[140] C. R. Kagan, C. B. Murray, M. Nirmal, and M. G. Bawendi, "Electronic Energy Transfer in CdSe Quantum Dot Solids," *Physical Review Letters*, vol. 76, pp. 1517–1520, Feb. 1996.

[141] M. Achermann, M. A. Petruska, D. D. Koleske, M. H. Crawford, and V. I. Klimov, "Nanocrystal-Based Light-Emitting Diodes Utilizing High-Efficiency Nonradiative Energy Transfer for Color Conversion," *Nano Letters*, vol. 6, pp. 1396–1400, July 2006.

[142] H. Kim, M. Shin, H. Hong, B. Song, S. Kim, W. Koo, J. Yoon, S. Yoon, and Y. Kim, "Enhancement of Optical Efficiency in White OLED Display Using the Patterned Photoresist Film Dispersed With Quantum Dot Nanocrystals," *Journal of Display Technology*, vol. 12, pp. 526–531, June 2016.

[143] M. K. Choi, J. Yang, K. Kang, D. C. Kim, C. Choi, C. Park, S. J. Kim, S. I. Chae, T.-H. Kim, J. H. Kim, T. Hyeon, and D.-H. Kim, "Wearable Red-Green-Blue Quantum Dot Light-Emitting Diode Array Using High-Resolution Intaglio Transfer Printing," *Nature Communications*, vol. 6, May 2015.

[144] T.-H. Kim, D.-Y. Chung, J. Ku, I. Song, S. Sul, D.-H. Kim, K.-S. Cho, B. L. Choi, J. Min Kim, S. Hwang, and K. Kim, "Heterogeneous Stacking of Nanodot Monolayers by Dry Pick-and-Place Transfer and Its Applications in Quantum Dot Light-Emitting Diodes," *Nature Communications*, vol. 4, p. 2637, Nov. 2013.

[145] W. K. Bae, J. Kwak, J. Lim, D. Lee, M. K. Nam, K. Char, C. Lee, and S. Lee, "Multicolored Light-Emitting Diodes Based on All-Quantum-Dot Multilayer Films Using Layer-by-Layer Assembly Method," *Nano Letters*, vol. 10, pp. 2368–2373, July 2010.

[146] W. K. Bae, J. Lim, D. Lee, M. Park, H. Lee, J. Kwak, K. Char, C. Lee, and S. Lee, "R/G/B/Natural White Light Thin Colloidal Quantum Dot-Based Light-Emitting Devices," *Advanced Materials*, vol. 26, no. 37, pp. 6387–6393, 2014.

[147] P. O. Anikeeva, J. E. Halpert, M. G. Bawendi, and V. Bulović, "Electroluminescence from a Mixed Red-Green-Blue Colloidal Quantum Dot Monolayer," *Nano Letters*, vol. 7, pp. 2196–2200, Aug. 2007.

[148] J. Schanda, "CIE Color-Rendering Index," in *Encyclopedia of Color Science and Technology* (R. Luo, ed.), pp. 1–5, Berlin, Heidelberg: Springer, 2014.

[149] H. Yang, Y. Shi, Y. Zhao, Y. Meng, W. Hu, J. Hou, and S. Liu, "High Colour Rendering Index White Organic Light-Emitting Devices with Three Emitting Layers," *Displays*, vol. 29, pp. 327–332, Oct. 2008.

[150] X. Guo and K. Houser, "A Review of Colour Rendering Indices and Their Application to Commercial Light Sources," *Lighting Research & Technology*, vol. 36, pp. 183–197, Sept. 2004.

[151] Y. Ohno, "Color Rendering of Light Sources." National Institute of Standards and Technology, July 2009.

[152] B. Kumar, S. A. Campbell, and P. Paul Ruden, "Modeling Charge Transport in Quantum Dot Light Emitting Devices with NiO and ZnO Transport Layers and Si Quantum Dots," *Journal of Applied Physics*, vol. 114, p. 044507, July 2013.

[153] B. Xu, S.-A. Gopalan, A.-I. Gopalan, N. Muthuchamy, K.-P. Lee, J.-S. Lee, Y. Jiang, S.-W. Lee, S.-W. Kim, J.-S. Kim, H.-M. Jeong, J.-B. Kwon, J.-H. Bae, and S.-W. Kang, "Functional Solid Additive Modified PEDOT:PSS as an Anode Buffer Layer for Enhanced Photovoltaic Performance and Stability in Polymer Solar Cells," *Scientific Reports*, vol. 7, p. 45079, Apr. 2017.

[154] Ossila, "TFB Polymer for OLEDs & Devices | 220797-16-0." Datasheet available at https://www.ossila.com/products/tfb.

[155] D. A. Ahn, S. Lee, J. Chung, Y. Park, and M. C. Suh, "Impact of Interface Mixing on the Performance of Solution Processed Organic Light Emitting Diodes—Impedance and Ultraviolet Photoelectron Spectroscopy Study," *ACS Applied Materials & Interfaces*, vol. 9, pp. 22748–22756, July 2017.

[156] S. Wang, Y. Guo, D. Feng, L. Chen, Y. Fang, H. Shen, and Z. Du, "Bandgap Tunable Zn1-xMgxO Thin-Films as Electron Transport Layers for High Performance Quantum Dot Light-Emitting Diodes," *Journal of Materials Chemistry C*, vol. 5, pp. 4724–4730, May 2017.

[157] F. Wang, W. Sun, P. Liu, Z. Wang, J. Zhang, J. Wei, Y. Li, T. Hayat, A. Alsaedi, and Z. Tan, "Achieving Balanced Charge Injection of Blue Quantum Dot Light-Emitting Diodes through Transport Layer Doping Strategies," *The Journal of Physical Chemistry Letters*, vol. 10, pp. 960–965, Mar. 2019.

[158] L. Qian, Y. Zheng, J. Xue, and P. H. Holloway, "Stable and Efficient Quantum-Dot Light-Emitting Diodes Based on Solution-Processed Multilayer Structures," *Nature Photonics*, vol. 5, pp. 543–548, Sept. 2011.

[159] Azo Materials, "Zinc Oxide (ZnO) Semiconductors." Datasheet available at https://www.azom.com/article.aspx?ArticleID=8417, Apr. 2013.

[160] P. Guyot-Sionnest, "Introduction to QD Properties and Applications," *Material Matters*, vol. 2, no. 1, 2007.

[161] Phillips, "LED Classic: 60W E27 CDL A60 FR ND 1PF/10." Phillips Commercial LED Catalogue. Specification available at https://www.lighting.philips.co.uk/consumer/p/led-bulb/8718699762551/specifications.

[162] A. Developers, "Android Battery Manager API Reference." Documentation available at https://developer.android.com/reference/android/os/BatteryManager/.

[163] Unsplash, "The Unsplash License." Available at https://unsplash.com/license/.

Appendix A

OLED Power Characterisation and Modelling

A.1 Introduction

The approximate colour transforms outlined in chapter 5 depend on having an accurate power model of the display (see eq. (5.1)). This appendix explains the process we used to measure the display power characteristics of the *Huawei P30 Pro* device used in the contemporary user studies outlined in chapter 5. We also discuss the results of few related experiments we carried out into evaluating the temperature dependence of the display's power consumption.

A.2 Power measurement

The first stage of creating the *Crayon* power model involves measuring the display power at different pixel values. One method would be to connect a Voltmeter and an Ammeter to the display being tested, send different R,G,B signals to the display, and measure the power draw. This was what the original authors did for their evaluation of *Crayon*. However, in this work, we wanted to evaluate the performance of the *Crayon* on a real device (i.e., the *Huawei P30 Pro* device we chose). As such, we needed to ensure parity between the display we measured and the display of the physical device.

The body of the device had a glue sealed construction and opening up the device and getting access to the power terminals on the display would have been impossible to do without potentially damaging the device. This restricted our ability to measure the display power as outlined above.

Instead, we chose a more holistic approach by measuring the power of the display as a part of the entire mobile system. We devised a method that uses the battery management subsystem in the *Android* operating system to measure the overall system power consumption. The idea was that, by interspersing active display states (i.e., showing a particular colour on the screen) with inactive display states (i.e., display set to show a black screen), we would be able to get baseline readings of the system power

without the power draw of the display. We can use this information to calculate the power draw of the active display. We used two *Huawei P30 Pro* mobile phones for this process to check for parity.



Fig. A.1 We used a *Binder MK56* thermal chamber (left) to measure display parameters for the Huawei P30 Pro devices (inside the chamber, right) at 25°C.

A.2.1 Methodology

The *Android BatteryManager API* provides access to information about the state of the device's battery, including, but not limited to, the following metrics [162]:

- Battery charging status
- Battery capacity (in microampere-hours)
- Average battery current (in microamperes)
- Instantaneous battery current (in microamperes)
- Remaining battery capacity (in nanowatt-hours)
- Battery temperature
- Battery voltage (in millivolts)

We wrote a *Kotlin* application that used this API to periodically read the above metrics and store them on a file. The application uses a configuration file with a list of display states along with how long to display each state. Each display state is defined as a tuple of ['state type','state parameters']. The display 'state type' is either 'rgb' or 'image'. For an 'rgb' type, the parameters are the exact RGB values to show on the screen as a hyphen separated string (e.g., "0-50-12" for the RGB triple [0,50,12]). For 'image' type, the parameter is a string of the path to the image to be displayed on the screen.

The application goes through the list of display states and records the aforementioned metrics. Between display state changes, the application turns the display off by setting all the pixel values to zero. It then makes multiple readings of the power draw of the system. These power readings serve as the baseline power reading of the system. The power consumption above this baseline is taken to be the power consumption of the display.

Since both the logic for controlling the display states and the power measurements are integrated into the same application, it uses the built-in software clock of the device for timing and to accurately match changes in display state to the changes in the system power. This allows the app to be packaged and distributed, to allow in-situ display power measurement by 3^{rd} parties (e.g., other research groups and consumers).

We used a *Binder MK56* thermal chamber set to $25°C$ to control the ambient temperature of the displays and carried out the power measurements inside this chamber. Once the display power for all the possible individual R,G,B states (3x255 states) were measured, the results were transferred from the mobile device for processing. The raw measurements were filtered using different methods to extract stable power readings. A quadratic function was fit to this data using a 'least squares' fit which was then used as the parametric power model for the *Crayon* image transforms.

A.2.2 Results

Raw Data

Figure A.2 shows the raw current measurements as the display cycled through the RGB values from 0 to 255, one channel at a time. Figure A.3 shows a zoomed-in view of the measurements, highlighting the measurement noise and the transient behaviour of the current. From this noisy data, we needed to extract a stable current measurement for each RGB value before fitting a quadratic function to the data.

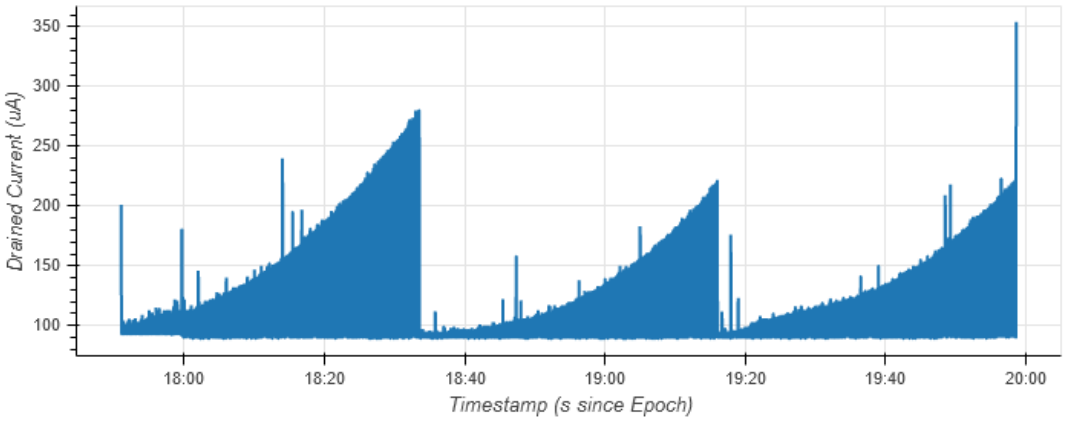


Fig. A.2 Raw current readings from the devices are very noisy and needs filtering.

The first step of the processing is to correlate the screen states to the power readings by comparing the timestamps on the power readings with the experiment log, which records the timestamps for when each display state was active.

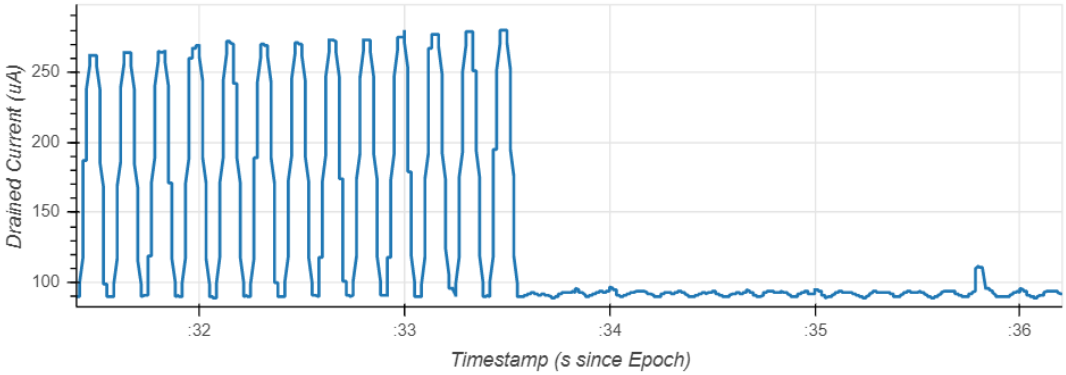
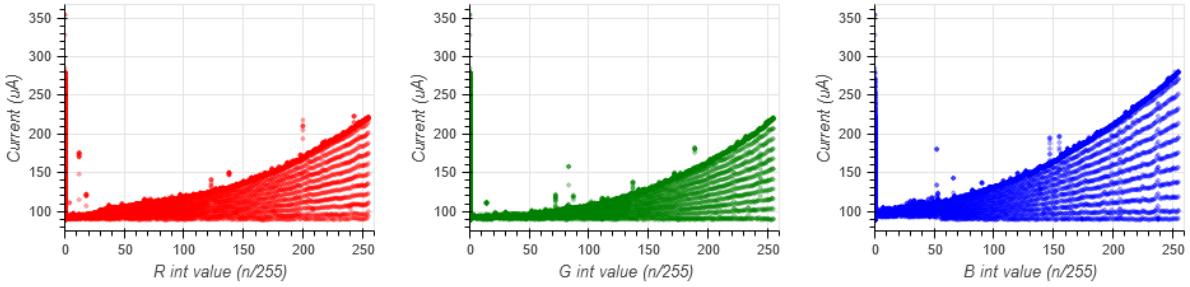


Fig. A.3 The measurement variations are more apparent when zoomed in.



(a) Red (b) Green (c) Blue

Fig. A.4 By correlating the display current with the display state, the variation of current as a function of pixel value for the three channels can be calculated.

Zero Filtering

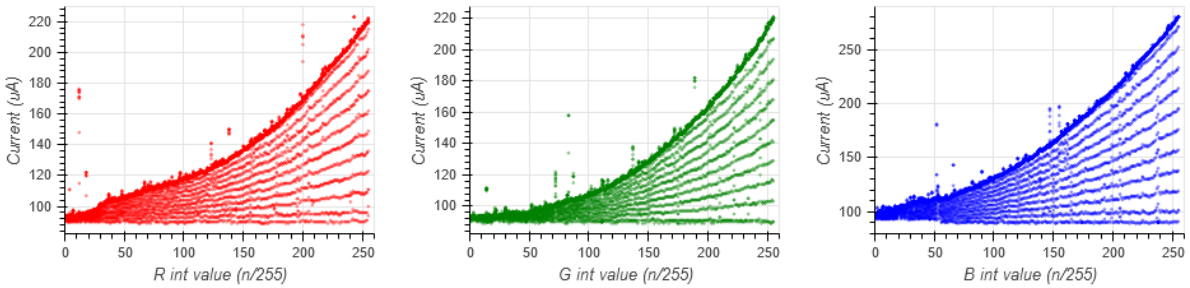
As the first filtering step is to filter out all the zero values (i.e. the baseline readings). These appear as the vertical line of measurements at the left most end of fig. A.4. Figures A.5 and A.14 show the results of this process.

Neighbourhood Average

We then calculated the average current value of the zero-filtered data for the different display states. The idea was that this would smooth out the current fluctuations that occur even when the display state is unchanging. As expected this resulted in much smoother current curves (see Figures A.6 and A.15a).

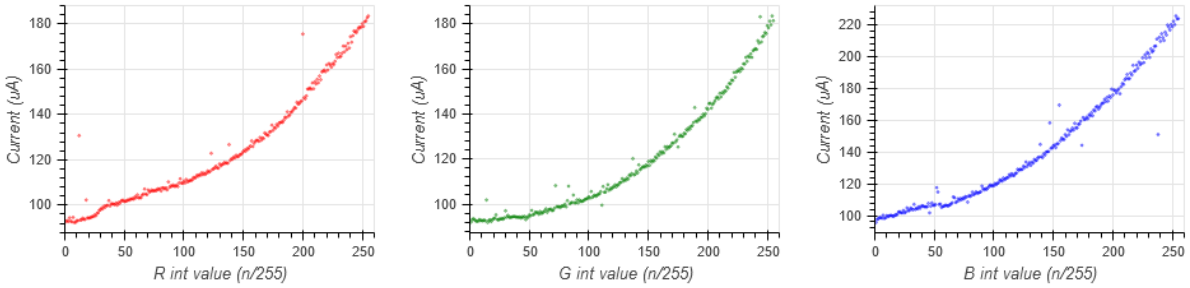
Neighbourhood Average - Pairwise Re-normalisation

As can be seen in fig. A.15a, even at an R/G/B value of zero, there is non-zero current draw. As the display is completely turned off at this stage, this value corresponds to the power draw of the rest of the system. We adopted a pairwise normalisation strategy instead of using the mean baseline reading to ensure that our normalisation process takes into account variations in the system power over the duration of the measurement experiment. For this process, we first averaged the baseline current draw



(a) Red (b) Green (c) Blue

Fig. A.5 Zero-filtering removes the baseline power readings and make the current variation more apparent.



(a) Red (b) Green (c) Blue

Fig. A.6 'Neighbourhood averaging' removes the observed variation at a given display state and results in a clear current trend.

immediately before the screen state was changed to a given RGB value. We then subtracted this value from the current draw recorded when the display state was changed to the chosen RGB value. Figure A.7 shows the baseline current reading highlighted in blue and the non-zero display state current highlighted in red. The corresponding pairs are shown with matching numbers in the highlighted regions.

Figures A.8 and A.16a show the results of this process, indicating that the baseline current draw has been factored out from the aggregate current reading.

Step Filtered

We observed that the baseline values the neighbourhood averaging process extracted was correlated with the display state (see fig. A.9). Although some variation in the baseline is expected, the strong trend shown in this strong correlation indicated that the filtering process was inadequate in extracting a stable baseline reading and this would cause the display power to be underestimated.

We investigated the behaviour of the current as the display changed state. Figure A.10 shows a section of the raw current measurements with vertical lines indicating the display state changes. At the red dashed lines, the display state changed from black (baseline) to the designated RGB value. At the black dotted lines, the display state reverted back to black (baseline). As can be seen, there is a significant rise and fall time associated with these transitions. Given a long enough measurement at

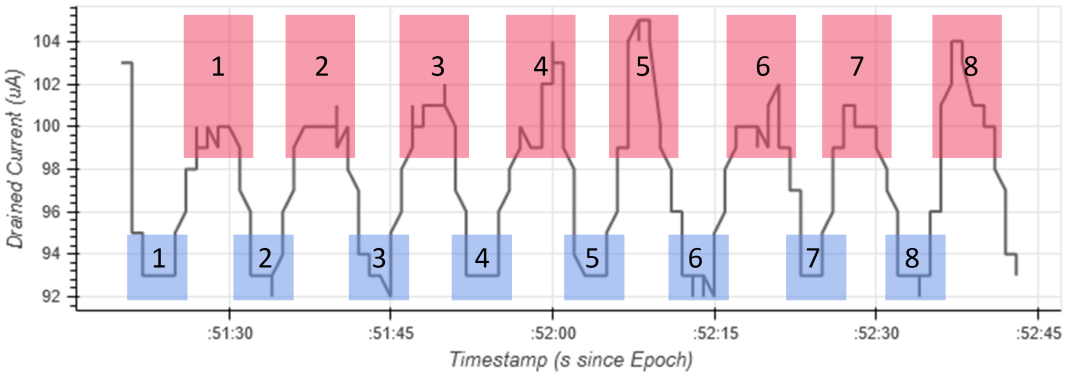
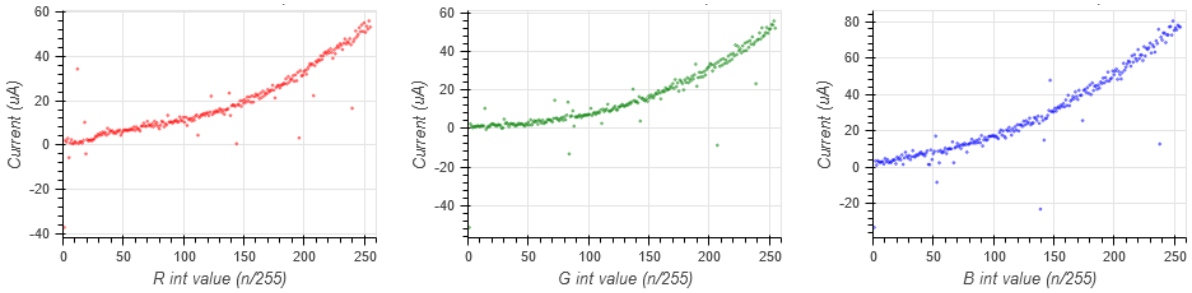


Fig. A.7 Pariwise re-normalisation uses the average baseline reading (blue) immediately before the active display state (red) to isolate the current draw of the display from the total system current draw.



(a) Red (b) Green (c) Blue

Fig. A.8 The pairwise re-normalisation process isolates the display power as a function of the pixel RGB value.

each state, the effect of this transient could be minimised, but as the data shows, the duration of the transient is in a similar time scale to the steady-state observations. This causes the baseline readings to be artificially inflated and the display power to be underestimated.

As a solution, we carried out a simple filtering method (which we refer to as “step-filtering”), where the measured power was determined by the value immediately prior to the state change. Thus, this method uses the current readings immediately before the black dotted lines as the current readings for the on-state display current and the readings immediately before the red dashed lines were used as the baseline current readings. Figures A.11 and A.15b shows the results of this filtering process. Figure A.12 shows the baseline power reading as calculated by this method. As intended, this leads to a more stable baseline reading.

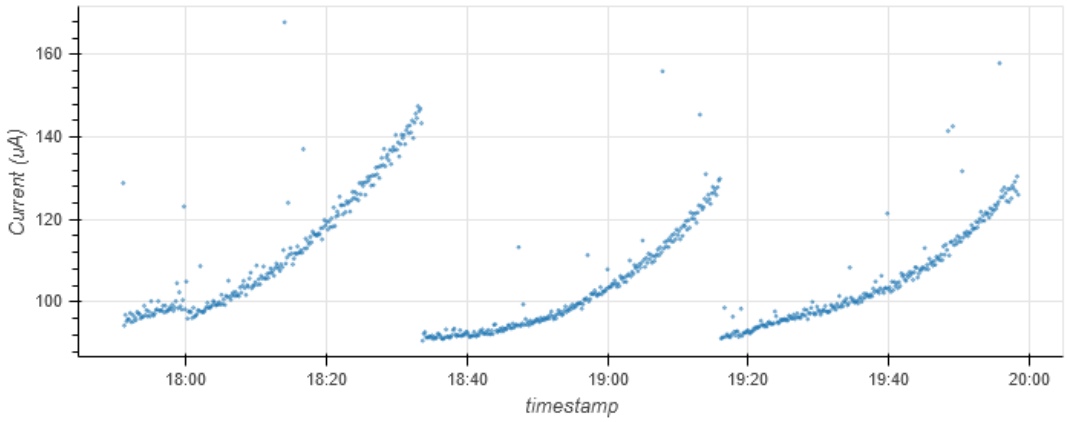


Fig. A.9 The baseline current extracted by the 'neighbourhood averaging' process is highly correlated with the display state. This indicates that the rise and fall times associated with the display sate changes are causing the baseline current to be overestimated.

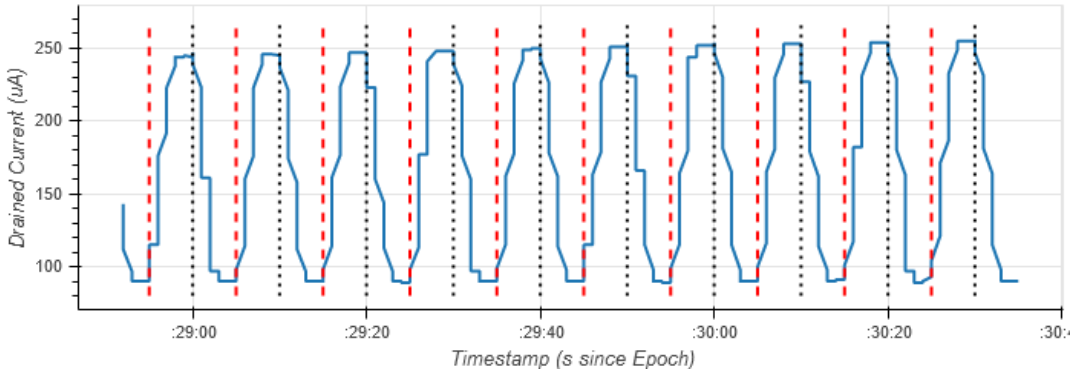


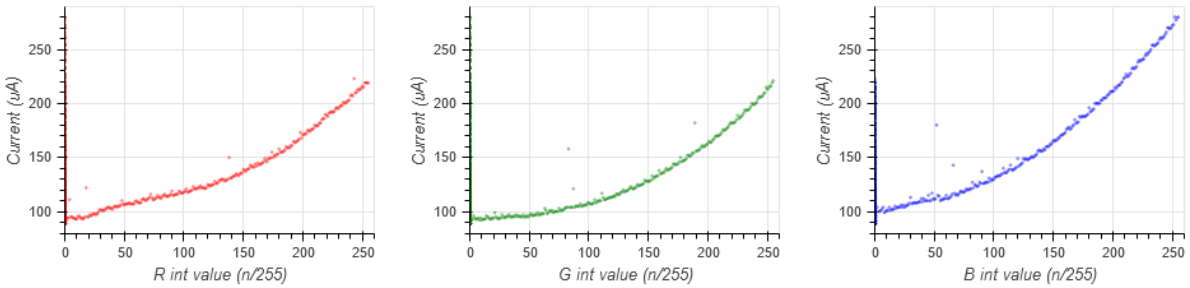
Fig. A.10 Marking the timestamps for display state changes makes it clear that the calculated display current is biased by the rise and fall times of the system current.

Step Filtered Re-normalised

To isolate the display power, we normalised the step-filtered values by subtracting the baseline power readings, pairwise, as before (see figs. A.13 and A.16b).

A.2.3 Finding *Crayon* Model parameters

The final step of creating the power model for *Crayon* involves fitting the observed data with three quadratic functions. We used a least-squares optimisation process for this. Figure A.17 shows the results of this process.



(a) Red (b) Green (c) Blue

Fig. A.11 Step-filtering leads to a more accurate estimate of the system current draw compared to the neighbourhood averaging and shows the clear trend even before the pairwise re-normalisation process.

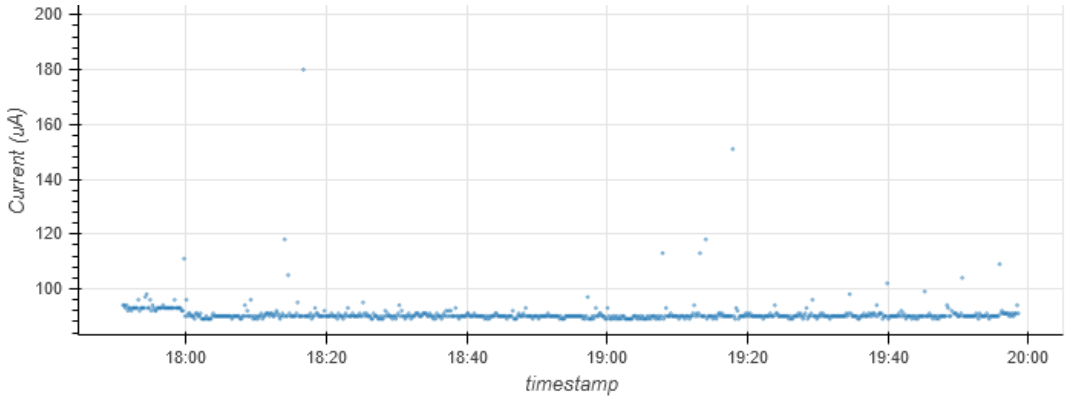
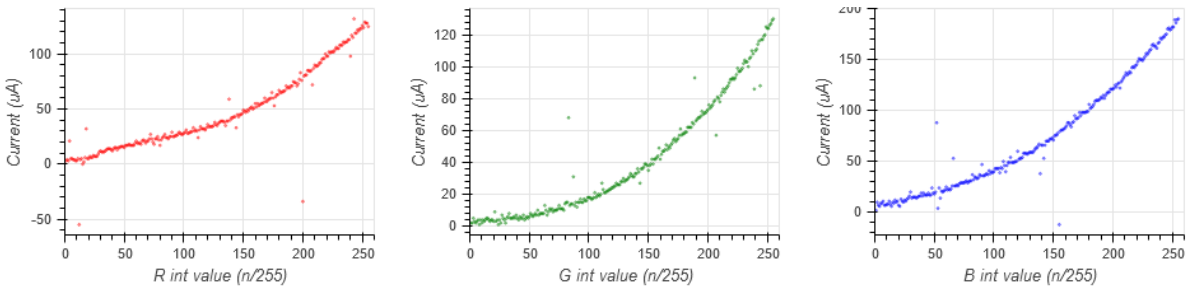


Fig. A.12 Baseline current draw extracted by the step-filtering method is more stable and shows no correlation with the display state.

A.2.4 Temperature dependence of power

To quantify the effect of display panel temperature on its power consumption, we repeated the RGB sweep experiment in the thermal chamber at different temperatures. We did the measurements at $0°C$, $15°C$, $25°C$ and $35°C$. This temperature range represents the maximum and minimum rated operational temperatures of the device. To ensure that the surface of the device was at the same temperature as the internal temperature of the thermal chamber, we allowed the thermal chamber to reach the target temperature with the device in it and left the device inside for 5 minutes before carrying out the measurements. Each display state was measured for 10 seconds, as before, and it took 5 hours to obtain a complete set of measurements for each temperature setting. As before, we carried out these measurements for two *Huawei P30 Pro* devices.

Figures A.18a and A.18b illustrate data from one of the devices for the two experiments at $0°C$ and $35°C$. The results show that the display power is not dependent on its temperature to a significant degree.



(a) Red (b) Green (c) Blue

Fig. A.13 Step-filtered and pairwise re-normalised values show a clear quadratic relationship between the display current and the RGB value.

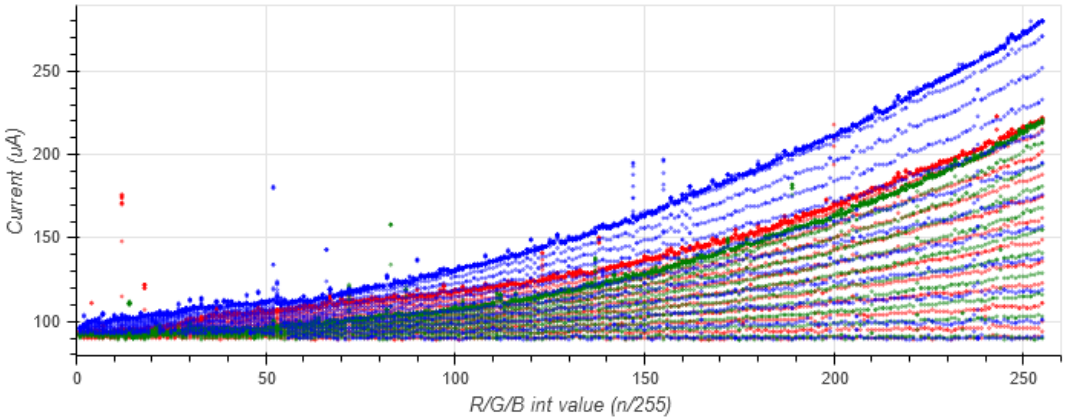
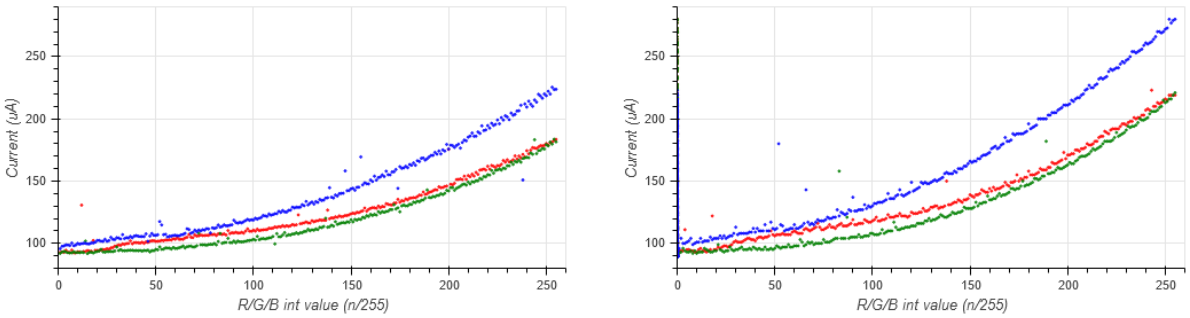
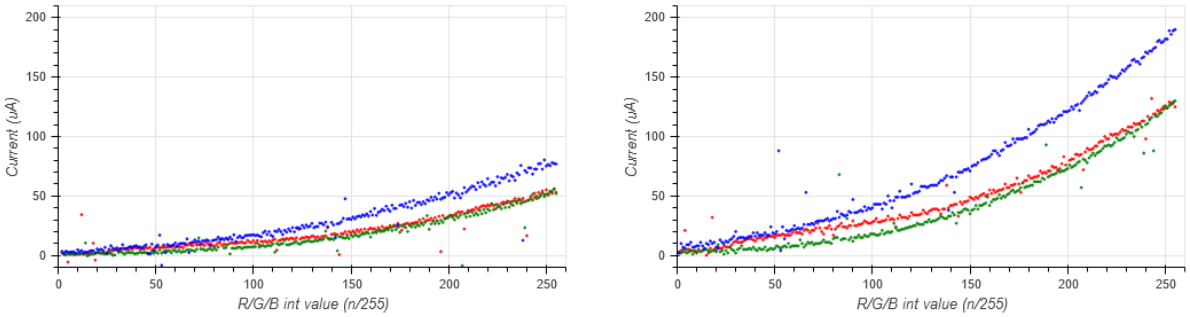


Fig. A.14 The results of simply zero-filtering the raw current measurements are inadequate for calculating the display power model.



(a) Neighbourhood Average (b) Step-Filtering

Fig. A.15 Due to the effect of rise/fall times, the neighbourhood averaging process significantly underestimates the overall system current draw.



(a) Neighbourhood Average (b) Step-Filtering

Fig. A.16 The error caused in the neighbourhood averaging propagates to the pairwise re-normalisation step where it is exacerbated by the overestimated baseline current reading leading to a calculation error $> 100\mu A$ (for the blue pixel at 255).

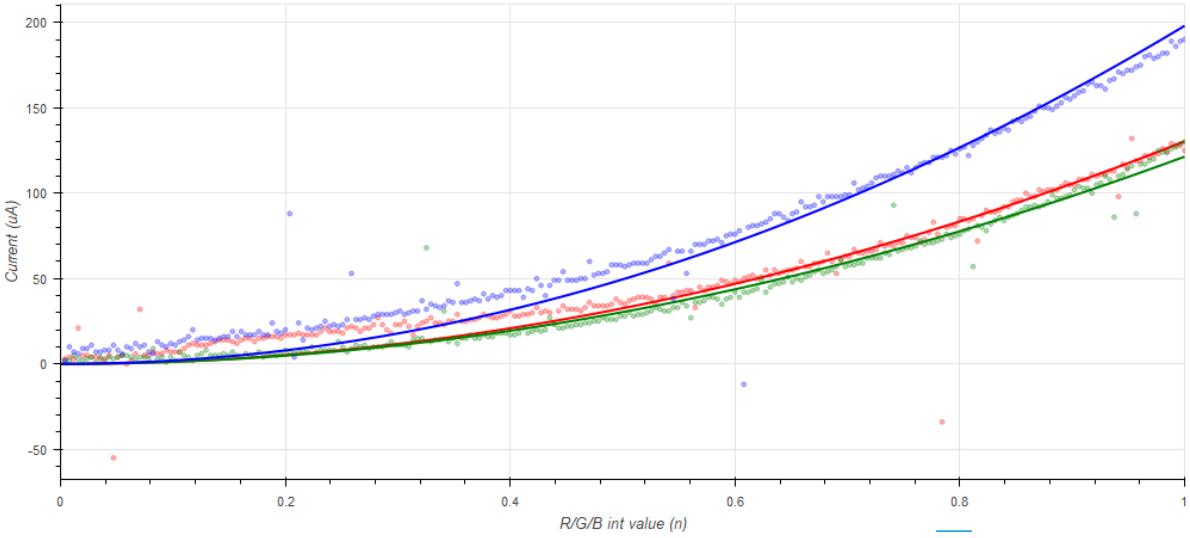
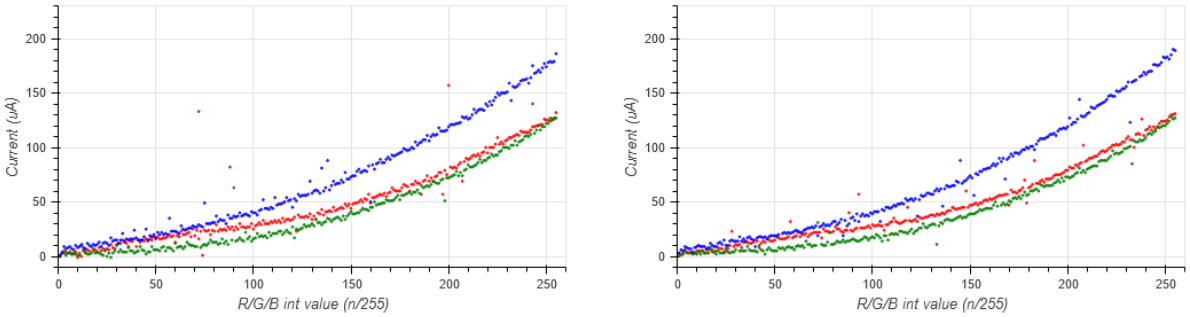


Fig. A.17 The display power model is calculated by fitting a triple of quadratics to the step-filtered and pairwise re-normalised data.



(a) At T = $0°C$ (minimum rated operational temperature) (b) At T = $35°C$ (maximum rated operational temperature)

Fig. A.18 The current draw of the display is not significantly affected by the ambient temperature.

Appendix B

Supplimentary Results from the Crayon User Studies

B.1 Predicting user score

		Linear		Cubic		SVM	
Colour Space	Distance	MSE	Variance	MSE	Variance	MSE	Variance
RGB	L2 (ℓ_2)	0.03400	0.00112	0.04011	0.00144	0.04307	0.00105
RGB	L2 Sq. (ℓ_2^2)	0.01446	0.00031	0.01068	0.00027	0.01880	0.00049
CIE-LAB	L2 (ℓ_2)	0.02862	0.00049	0.02651	0.00058	0.02753	0.00077
CIE-LAB	L2 Sq. (ℓ_2^2)	0.00904	0.00009	0.00795	0.00008	0.01023	0.00006
CIE-UVW	L2 (ℓ_2)	0.01357	0.00002	0.01464	0.00005	0.01486	0.00006
CIE-UVW	L2 Sq. (ℓ_2^2)	0.01262	0.00014	0.01084	0.00007	0.01355	0.00013

Table B.1 Performance of the models using Λ to predict the user adjusted scores (UAS).

		Linear		Cubic		SVM	
Colour Space	Distance	MSE	Variance	MSE	Variance	MSE	Variance
RGB	L2 (ℓ_2)	0.03411	0.00162	0.03749	0.00198	0.04161	0.00180
RGB	L2 Sq. (ℓ_2^2)	0.01446	0.00021	0.00774	0.00008	0.01056	0.00010
CIE-LAB	L2 (ℓ_2)	0.02660	0.00064	0.02385	0.00059	0.02782	0.00069
CIE-LAB	L2 Sq. (ℓ_2^2)	0.00980	0.00006	0.00708	0.00005	0.00840	0.00003
CIE-UVW	L2 (ℓ_2)	0.01319	0.00009	0.01325	0.00006	0.01474	0.00008
CIE-UVW	L2 Sq. (ℓ_2^2)	0.01238	0.00005	0.01125	0.00010	0.01486	0.00010

Table B.2 Performance of the models using Λ to predict the mean opinion scores (MOS).

			Linear		Cubic		SVM	
	Colour Space	Distance	MSE	Variance	MSE	Variance	MSE	Variance
MOS	RGB	L2 (ℓ_2)	0.03411	0.00162	0.03749	0.00198	0.04602	0.00203
	RGB	L2 Sq. (ℓ_2^2)	0.01446	0.00021	0.00774	0.00008	0.01068	0.00009
	CIE-LAB	L2 (ℓ_2)	0.02862	0.00049	0.02651	0.00058	0.02542	0.00054
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.01446	0.00021	0.00774	0.00008	0.01068	0.00009
	CIE-UVW	L2 (ℓ_2)	0.01357	0.00002	0.01464	0.00005	0.01412	0.00007
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.01238	0.00005	0.01125	0.00010	0.01314	0.00010
UAS	RGB	L2 (ℓ_2)	0.03400	0.00112	0.04011	0.00144	0.05493	0.00309
	RGB	L2 Sq. (ℓ_2^2)	0.01446	0.00031	0.01068	0.00027	0.00833	0.00014
	CIE-LAB	L2 (ℓ_2)	0.02660	0.00064	0.02385	0.00059	0.02107	0.00057
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.00904	0.00009	0.00795	0.00008	0.00852	0.00006
	CIE-UVW	L2 (ℓ_2)	0.01319	0.00009	0.01325	0.00006	0.01431	0.00010
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.01262	0.00014	0.01084	0.00007	0.01326	0.00009

Table B.3 Performance of the models using normalised Λ to predict the mean opinion scores (MOS).

			Linear		Cubic		SVM	
	Colour Space	Distance	MSE	Variance	MSE	Variance	MSE	Variance
Raw MOS	RGB	L2 (ℓ_2)	0.63155	0.23803	4.59797	14.56214	0.83579	1.23580
	RGB	L2 Sq. (ℓ_2^2)	0.38238	0.09590	3.93264	63.33686	0.31055	0.13054
	CIE-LAB	L2 (ℓ_2)	1.27895	0.82855	7.27373	11.30107	0.49846	0.34544
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.39849	0.05317	1.56376	14.76341	0.13216	0.01653
	CIE-UVW	L2 (ℓ_2)	0.22148	0.02763	32.12062	1078.10	0.17934	0.01158
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.31222	0.02242	2.73416	5.04147	0.39671	0.03810
Norm. MOS	RGB	L2 (ℓ_2)	0.02526	0.00038	0.20351	0.02285	0.03253	0.00032
	RGB	L2 Sq. (ℓ_2^2)	0.01530	0.00015	0.03324	0.00288	0.02814	0.00016
	CIE-LAB	L2 (ℓ_2)	0.05116	0.00133	0.06917	0.00286	0.03216	0.00018
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.01594	0.00009	0.01007	0.00009	0.01978	0.00003
	CIE-UVW	L2 (ℓ_2)	0.00886	0.00004	0.53268	0.29284	0.02127	0.00006
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.01249	0.00004	0.12410	0.00965	0.02579	0.00008
UAS	RGB	L2 (ℓ_2)	0.02435	0.00043	0.24227	0.06663	0.02507	0.00047
	RGB	L2 Sq. (ℓ_2^2)	0.01843	0.00022	0.09682	0.00834	0.02748	0.00026
	CIE-LAB	L2 (ℓ_2)	0.05098	0.00092	0.23682	0.05210	0.03429	0.00027
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.01183	0.00007	0.05718	0.00652	0.01236	0.00003
	CIE-UVW	L2 (ℓ_2)	0.00911	0.00002	712.72135	1286030	0.01957	0.00005
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.01198	0.00008	0.08496	0.00249	0.02313	0.00017

Table B.4 Performance of the models using normalised Λ, and mean and variance of HSV values to predict the mean opinion scores (MOS).

			Linear		Cubic		SVM	
	Colour Space	Distance	MSE	Variance	MSE	Variance	MSE	Variance
Raw MOS	RGB	L2 (ℓ_2)	0.4220	0.05988	1,985.7	$> 100,000$	1.2685	1.96651
	RGB	L2 Sq. (ℓ_2^2)	0.5263	0.22461	466.7	$> 100,000$	0.3109	0.10050
	CIE-LAB	L2 (ℓ_2)	1.2452	0.81225	454.2	$> 100,000$	0.4977	0.35179
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.5602	0.19705	207.7	73,089	0.1956	0.03473
	CIE-UVW	L2 (ℓ_2)	0.1692	0.00639	161,343.3	$> 100,000$	1.4078	0.41703
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.3332	0.02430	8,767.6	$> 100,000$	0.5534	0.10679
Norm. MOS	RGB	L2 (ℓ_2)	0.0168	0.00010	66.6	34,679	0.0815	0.00185
	RGB	L2 Sq. (ℓ_2^2)	0.0210	0.00036	581.2	$> 100,000$	0.0279	0.00018
	CIE-LAB	L2 (ℓ_2)	0.0498	0.00130	1,212.7	$> 100,000$	0.0927	0.01065
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.0224	0.00032	13.0	444	0.0197	0.00003
	CIE-UVW	L2 (ℓ_2)	0.0067	0.00001	5,656.9	$> 100,000$	0.0571	0.00036
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.0133	0.00004	329.3	18,2657	0.0234	0.00013
UAS	RGB	L2 (ℓ_2)	0.0239	0.00035	161.6	$> 100,000$	0.0962	0.00245
	RGB	L2 Sq. (ℓ_2^2)	0.0200	0.00033	17.0	2,257	0.0284	0.00021
	CIE-LAB	L2 (ℓ_2)	0.0332	0.00038	9,075.5	$> 100,000$	0.0983	0.01246
	CIE-LAB	L2 Sq. (ℓ_2^2)	0.0122	0.00011	62.3	24,399	0.0123	0.00003
	CIE-UVW	L2 (ℓ_2)	0.0082	0.00001	159,178.6	$> 100,000$	0.0554	0.00042
	CIE-UVW	L2 Sq. (ℓ_2^2)	0.0115	0.00006	183.2	$> 100,000$	0.0189	0.00023

Table B.5 Performance of the models using normalised Λ, and mean and variance of CIE-Lab values to predict the mean opinion scores (MOS).

Appendix C

Tabulated results of the CRI optimisation process

Abbreviation.	Explanation
Opt.	Optimisation method
Init.	Initialisation scheme
NM	Nelder-Mead
B	L-BFGS-B
S	SHGO
LMS	Cone fundamental range based initialisation
-	No explicit initilisation (SHGO by design tries multiple initial positions)
xy+cri	Weighted difference between ℓ_2^2-norm and linearly normalised CRI (cost function 7)
xy+cri+div	Weighted difference between ℓ_2^2-norm and linearly normalised CRI + CRI spread penalty (cost function 8)
xyCRI seq	Sequential optimisation with ℓ_2^2-norm of the xy chromaticity coordinates (cost function 2) followed by Inverse CRI (cost function 3)

Table C.1 Summary of abbreviations used in the data tables

N	Opt.	Cost	Init.	ΔE_{00}	Δxy	CIE_{xy}	CRI_{avg}	CRI_{max}	CRI_{min}
3	NM	xy+cri	LMS	38.756	0.0099	[0.311, 0.319]	84.75	97.88	28.38
4	NM	xy+cri	LMS	40.317	0.0216	[0.299, 0.312]	97.23	99.52	83.37
5	NM	xy+cri	LMS	39.664	0.0213	[0.303, 0.310]	96.57	98.80	92.05
6	NM	xy+cri	LMS	41.563	0.0436	[0.281, 0.299]	97.71	99.23	91.25
7	NM	xy+cri	LMS	36.975	0.0225	[0.296, 0.314]	97.92	99.65	76.83
3	NM	xy+cri+div	LMS	42.750	0.0233	[0.330, 0.345]	62.03	84.61	26.29
4	NM	xy+cri+div	LMS	31.517	0.0050	[0.308, 0.327]	88.99	95.45	71.73
5	NM	xy+cri+div	LMS	33.462	0.0074	[0.312, 0.322]	97.25	98.62	92.58
6	NM	xy+cri+div	LMS	42.089	0.0457	[0.281, 0.296]	97.72	99.68	93.61
7	NM	xy+cri+div	LMS	41.468	0.0494	[0.278, 0.294]	97.75	99.09	90.61
3	NM	xyCRI seq	LMS	41.434	0.0174	[0.296, 0.326]	84.62	99.30	-68.08
4	NM	xyCRI seq	LMS	40.379	0.0217	[0.299, 0.312]	97.23	99.50	83.49
5	NM	xyCRI seq	LMS	39.933	0.0229	[0.300, 0.310]	97.14	99.45	91.79
6	NM	xyCRI seq	LMS	30.299	0.0084	[0.319, 0.335]	96.77	99.00	75.33
7	NM	xyCRI seq	LMS	38.591	0.0285	[0.292, 0.310]	98.32	99.65	86.20
3	B	xy+cri	LMS	42.692	0.0227	[0.291, 0.322]	84.62	98.55	-103.37
4	B	xy+cri	LMS	41.105	0.0144	[0.327, 0.326]	95.54	97.59	43.17
5	B	xy+cri	LMS	27.370	0.0031	[0.312, 0.326]	94.75	99.51	91.03
6	B	xy+cri	LMS	33.763	0.0123	[0.301, 0.324]	96.34	99.51	75.09
7	B	xy+cri	LMS	36.047	0.0194	[0.299, 0.315]	98.41	99.61	93.07
3	B	xy+cri+div	LMS	42.624	0.0229	[0.330, 0.345]	61.90	83.83	27.39
4	B	xy+cri+div	LMS	41.291	0.0190	[0.304, 0.312]	92.07	97.67	71.81
5	B	xy+cri+div	LMS	31.550	0.0068	[0.319, 0.330]	96.79	98.53	91.38
6	B	xy+cri+div	LMS	29.967	0.0067	[0.310, 0.323]	94.58	96.50	81.62
7	B	xy+cri+div	LMS	39.011	0.0299	[0.291, 0.308]	98.14	99.17	93.43
3	B	xyCRI seq	LMS	40.531	0.0181	[0.297, 0.320]	83.84	96.55	-13.74
4	B	xyCRI seq	LMS	43.802	0.0235	[0.336, 0.329]	94.13	97.60	34.96
5	B	xyCRI seq	LMS	45.617	0.0549	[0.364, 0.350]	-0.00	81.60	-394.75
6	B	xyCRI seq	LMS	32.885	0.0116	[0.302, 0.323]	96.32	99.38	74.59
7	B	xyCRI seq	LMS	25.354	0.0032	[0.313, 0.332]	96.94	99.46	61.51
3	S	xy+cri	-	43.364	0.0238	[0.290, 0.324]	84.91	98.64	-118.86
4	S	xy+cri	-	40.886	0.0242	[0.296, 0.311]	96.91	99.51	80.03
5	S	xy+cri	-	42.522	0.0281	[0.297, 0.305]	96.78	98.92	85.58
6	S	xy+cri	-	40.330	0.0216	[0.299, 0.312]	97.23	99.52	83.43
7	S	xy+cri	-	28.230	0.0049	[0.311, 0.324]	98.41	99.85	94.79
3	S	xy+cri+div	-	46.199	0.0288	[0.305, 0.357]	77.52	83.50	58.41
4	S	xy+cri+div	-	33.600	0.0095	[0.306, 0.323]	88.82	94.96	63.00
5	S	xy+cri+div	-	31.856	0.0055	[0.312, 0.323]	94.69	98.69	87.55
6	S	xy+cri+div	-	32.726	0.0072	[0.310, 0.322]	97.10	98.29	90.23
7	S	xy+cri+div	-	42.555	0.0256	[0.302, 0.306]	96.78	99.42	91.72

Table C.2 The properties of the calculated optimal spectra for the different cost functions and optimisation methods.

N	Opt.	Cost	Init.	Primary Wavelengths, $\overline{\lambda^*}$ (nm)	Relative Intensities, $\overline{w}$
3	NM	xy+cri	LMS	[463.13, 542.50, 610.51]	[1.306, 1.166 , 1]
4	NM	xy+cri	LMS	[449.56, 502.14, 558.56, 616.26]	[1.317, 1.233 , 1.138, 1]
5	NM	xy+cri	LMS	[432.53, 473.43, 525.75, 575.36, 622.50]	[1.054, 1.331 , 1.229, 1.109, 1]
6	NM	xy+cri	LMS	[436.52, 464.42, 501.37, 537.71, 583.66, 627.49]	[1.168, 1.381 , 1.293, 1.246, 1.104, 1]
7	NM	xy+cri	LMS	[422.36, 454.34, 486.13, 521.77, 556.95, 596.90, 632.48]	[1.101, 1.407 , 1.338, 1.263, 1.210, 1.074, 1]
3	NM	xy+cri+div	LMS	[433.01, 532.37, 608.13]	[1.028, 1.192 , 1]
4	NM	xy+cri+div	LMS	[423.81, 481.10, 545.07, 611.27]	[1.017, 1.289 , 1.166, 1]
5	NM	xy+cri+div	LMS	[425.83, 471.79, 527.81, 580.32, 627.50]	[1.061, 1.363 , 1.269, 1.132, 1]
6	NM	xy+cri+div	LMS	[432.00, 461.16, 498.18, 535.96, 582.92, 628.32]	[1.075, 1.398 , 1.300, 1.258, 1.115, 1]
7	NM	xy+cri+div	LMS	[432.49, 455.22, 484.43, 517.31, 552.50, 594.50, 632.50]	[1.094, 1.408 , 1.352, 1.266, 1.236, 1.071, 1]
3	NM	xyCRI seq	LMS	[462.74, 532.50, 600.90]	[1.301, 1.188 , 1]
4	NM	xyCRI seq	LMS	[449.52, 502.09, 558.56, 616.25]	[1.317, 1.233 , 1.138, 1]
5	NM	xyCRI seq	LMS	[432.65, 474.39, 527.61, 580.68, 629.57]	[1.096, 1.382 , 1.282, 1.142, 1]
6	NM	xyCRI seq	LMS	[406.63, 458.55, 509.56, 557.50, 606.06, 650.75]	[1.106, 1.470 , 1.348, 1.262, 1.121, 1]
7	NM	xyCRI seq	LMS	[429.87, 457.87, 490.63, 524.67, 559.96, 599.65, 637.50]	[1.038, 1.407 , 1.300, 1.270, 1.196, 1.076, 1]
3	B	xy+cri	LMS	[460.64, 527.50, 597.85]	[1.311, 1.192 , 1]
4	B	xy+cri	LMS	[432.76, 493.93, 558.76, 617.34]	[1.040, 1.236 , 1.139, 1]
5	B	xy+cri	LMS	[421.42, 469.48, 524.91, 577.34, 627.50]	[1.096, 1.363 , 1.259, 1.137, 1]
6	B	xy+cri	LMS	[420.08, 460.66, 502.90, 543.09, 588.99, 629.16]	[1.116, 1.406 , 1.302, 1.247, 1.069, 1]
7	B	xy+cri	LMS	[422.95, 456.37, 488.48, 526.72, 562.64, 603.67, 639.02]	[1.093, 1.405 , 1.314, 1.276, 1.184, 1.074, 1]
3	B	xy+cri+div	LMS	[433.18, 532.61, 608.23]	[1.031, 1.191 , 1]
4	B	xy+cri+div	LMS	[434.04, 486.39, 548.86, 611.67]	[1.053, 1.248 , 1.166, 1]
5	B	xy+cri+div	LMS	[419.85, 469.48, 527.79, 580.97, 627.50]	[1.107, 1.363 , 1.269, 1.127, 1]
6	B	xy+cri+div	LMS	[418.72, 459.94, 503.45, 545.20, 590.23, 627.50]	[1.104, 1.396 , 1.290, 1.235, 1.051, 1]
7	B	xy+cri+div	LMS	[423.54, 454.45, 485.46, 521.44, 557.50, 599.70, 637.50]	[1.089, 1.404 , 1.340, 1.259, 1.208, 1.076, 1]
3	B	xyCRI seq	LMS	[461.41, 532.50, 603.90]	[1.307, 1.189 , 1]
4	B	xyCRI seq	LMS	[427.86, 489.65, 559.35, 617.83]	[1.000, 1.238 , 1.138, 1]
5	B	xyCRI seq	LMS	[433.04, 401.45, 563.30, 583.86, 737.48]	[1.250, 1.139 , 1.339, 1.261, 1]
6	B	xyCRI seq	LMS	[419.90, 460.16, 502.45, 543.13, 588.66, 628.64]	[1.114, 1.404 , 1.299, 1.243, 1.069, 1]
7	B	xyCRI seq	LMS	[417.51, 455.24, 489.87, 526.10, 559.87, 598.56, 627.66]	[1.102, 1.393 , 1.292, 1.264, 1.187, 1.065, 1]

Table C.3 Primary wavelengths and the relative intensities of the primaries for the calculated optimal spectra using the different cost functions and optimisation methods.

N	Opt.	Cost	Init.	Primary Wavelengths, $\overline{\lambda^*}$ (nm)	Relative Intensities, $\overline{w}$
3	S	xy+cri	-	[527.50, 596.92, 461.17]	[1.194, 1 , 1.310]
4	S	xy+cri	-	[556.59, 500.98, 615.48, 449.80]	[1.145, 1.233 , 1 , 1.319]
5	S	xy+cri	-	[556.92, 499.34, 615.37, 448.13, 322.92]	[4.026, 4.347 , 3.522, 4.5622, 1]
6	S	xy+cri	-	[502.13, 330.94, 765.94, 616.25, 558.57, 449.54]	[2.921, 1 , 1.568, 2.369, 2.695, 3.120]
7	S	xy+cri	-	[439.52, 480.65, 663.57, 389.84, 616.26, 573.22, 527.62]	[1.906, 2.116 , 1.483, 1 , 1.62 , 1.762, 1.961]
3	S	xy+cri+div	-	[470.80, 532.37, 609.11]	[1.282, 1.193 , 1]
4	S	xy+cri+div	-	[507.50, 623.03, 444.59, 563.98]	[1.253, 1 , 1.279, 1.141]
5	S	xy+cri+div	-	[622.28, 417.28, 521.77, 465.32, 574.32]	[1 , 1.072 , 1.215, 1.341, 1.109]
6	S	xy+cri+div	-	[528.49, 634.55, 421.69, 583.84, 306.09, 470.67]	[53.07, 41.310, 45.670, 46.05, 1 , 56.87]
7	S	xy+cri+div	-	[500.31, 330.95, 765.96, 615.73, 558.56, 446.91, 319.69]	[5.547, 1.894 , 2.972, 4.491, 5.104 , 5.747, 1]

Table C.4 Primary wavelengths and the relative intensities of the primaries for the calculated optimal spectra using the different cost functions and optimisation methods (continued).

Glossary

A

Amazon Mechanical Turk (AMT)

A crowdsourcing website for hosting tasks that require human intelligence (e.g., image annotation). The AMT public workforce can volunteer to complete these tasks and get paid by the task posters, based on the task.

Application Programming interface (API)

A particular set of rules and specifications that a software program can follow to access and make use of the services and resources provided by another software program that implementing the API.

C

CC0

CC0 is a creative commons license where the author of the content attributed this license relinquishes their copyright, making the work part of the public domain, allowing for modification and re-distribution of the said material with no restrictions. It is also referred to as the "The Creative Commons Public Domain Dedication".

Charge-Transport Model (CTM)

A mathematical model that uses the mechanics of how current flows in solids to calculate the electrical properties of a (semiconductor) device.

Chromaticity

The property of a colour stimulus described by its chromaticity coordinates.

Colour Matching Function (CMF)

Set of functions (of wavelength) describing the proportions of the different primaries required to give a perceptual match with a given wavelength monochromatic light.

Colour Rendering Index (CRI)

A numerical measure of how well a light source preserves the colour appearance of objects compared to a reference light source (usually 'daylight').

Correlated Colour Temperature (CCT)

Correlated Colour Temperature of a spectrum is the temperature of the Plankian radiator with the closest chromaticity to that spectrum.

Current Efficiency (CE)

The ratio between output luminance and the input current of an opto-electric device.

E

Epitaxial Growth

The process of growing crystalline layers on top of a crystalline substrate.

Extended Graphics Array (XGA)

A display standard proposed by IBM, and is typically used to refer to displays with a resolution of 1024 by 768 pixels.

External Quantum Efficiency (EQE)

The ratio between the number of electrons injected into a device and the number of photons extracted from the device.

F

FM100

FM100 is the shorthand name used to refer to the Farnsworth-Munsell 100 Hue colour vision test. It is a hue discrimination test where the participants are asked to arrange a set of colour samples in an order such that the hue variation from the first sample to the last sample looks continuous.

Full Width At Half Maximum (FWHM)

Width of a function measured between the locations where the function takes half of its maximum value. For Gaussian shapes this is equal to $2\sqrt{2ln(2)}\sigma$.

I

Image Quality Metric (IQM)

An algorithm used to measure perceptual quality of an image.

L

Lateral Geniculate Nucleus (LGN)

The region in the human brain that collects impulses from the retina (through the optic nerve) and projects them to higher order visual processing in Occiptal Lobe.

Luminous Efficiency Function (LEF)

Function describing the brightness sensitivity of the human eye to different wavelengths of light.

M

Mean Opinion Score (MOS)

Mean value of the subjective opinion scores.

Mean Squared Error (MSE)

The average difference between squared values.

Metal-Organic Chemical Vapour Deposition (MOCVD)

A fabrication process where precursor chemicals in the gaseous phase are fed into a pressurised chamber where they undergo thermal decomposition and deposit onto the substrate semiconductor material.

Metamere

Metameres are spectra that have the same tristimulus values.

Metamerism

Metamerism is the effect where quantifiably different spectra are perceptually identical (i.e., have the same tristimulus values).

Multi Primary Display (MPD)

Displays with more than three primaries.

R

Root Mean Squared Error (RMSE)

Square root of the MSE.

S

Structural Similarity Index Measure (SSIM)

A popular IQM used for measuring the similarity between two images.

Support Vector Machine (SVM)

A class of machine learning models used in supervised learning that uses support-vectors to carry out classification or regression.

T

Technology-CAD (TCAD)

A branch of semiconductor design that uses models of the device circuitry, fabrication processes, and the fabricated devices to design and verify semiconductor devices.

Trichromatic Generalisation

The theory based on the existence of three types of cones, that state that a range of colour stimuli can be achieved by a mixture of three different colour stimuli.

U

Ultra High Definition (UHD)

A term used to refer to displays with resolutions of 3840 × 2160 ("4K") and 7680 × 4320 ("8K") pixels as standardised by the ITU recommendation BT-2020 (also known as rec.2020).

Unsplash license

Unsplash License is the license under which images on the image-sharing platform *Unsplash* can be used by 3rd parties. In its verbose form, it states that, "Unsplash grants you an irrevocable, non-exclusive, worldwide copyright license to download, copy, modify, distribute, perform, and use photos from Unsplash for free, including for commercial purposes, without permission from or attributing the photographer or Unsplash. This license does not include the right to compile photos from Unsplash to replicate a similar or competing service" [163].

User adjusted score (UAS)

Opinion score calculated after factoring out the user biases and inconsistencies.

V

Vacuum Thermal Evaporation (VTE)

A coating process where a solid material is evaporated in a vacuum chamber and allowed to deposit onto the substrate.

W

Wide Colour Gamut (WCG)

Usually refers to colour gamuts with more colours than the rec.709 colour space (e.g., Adobe RGB and DCI-P3 colour spaces).