Colour constancy for glossy objects under complex lighting environments



Takuma Morimoto

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2019

Acknowledgements

First of all, I would like to acknowledge my supervisor Hannah, who has been extremely encouraging all the time. It is clear that this thesis could not have been completed without her tremendous support. Even when my work slowed down, she was always patient and gave me sensible advice to overcome difficulties. I have also always appreciated that she has tried to give me chances to grow up as a scientist. For examples, asking me to give oral presentations at OxVis meetings, showing the metamerism demonstration for lab visitors, visiting Sergio's lab in Portugal, giving a lecture and running practical sessions at the RealVision Training Summer School, and attending many research conferences. Because of these opportunities, I was able to meet many excellent researchers within the UK and from other countries, from which I was able to seed the collaborative research that became a future path for my post-doc work. One of the most exciting moments that I had in my Oxford life was sitting together in front of my laptop computer, thinking about the project, and being excited about the idea we came up with. I very much enjoyed working with her, and I really hope that we will be able to work together again in the near future.

Second, I would like to thank my colleagues. It has been wonderful to be surrounded by such talented people all the time, and the intellectual conversation over lunch always interested and impressed me. They were also extremely patient with my broken English, which I always appreciated. I feel that I was very lucky to enter the DPhil programme at the same time with wonderful colleagues. It was just great that I had friends to laugh (and sometimes cry) together through the entire DPhil period.

Also, I would like to thank my friends in Oxford, who often went out together for a drink in the past three years (and probably even now). It was very nice that all of us live so close and that we were able to go out together after work. After getting together, most problems I had just shrunk, and were not so hard anymore to tackle. Spending time for years together closely in a compact city, I am confident that this is a kind of relationship I would like to maintain for years to come in my future.

Finally, of course, I want to thank my family who always warmly supported me. Without their help that I was always able to rely on, I would not have been able reach the end of the project. It was just like a security net I always had at the bottom in my mind, which gave me mental strength to keep going for three years.

Abstract

Colour constancy refers to our visual ability to identify the colour of objects under different illuminations. As a result of research over decades, our understanding of the mechanisms underlying this ability has advanced considerably.

However, a significant limitation of past studies has been that they have mostly employed oversimplified experimental stimuli that only partially reflect the properties of surfaces and illuminations in the natural environment. Stimuli were mostly matte, two-dimensional, and uniformly illuminated by single illumination. In the real world, in contrast, we know that objects sometimes include gloss (perceptual attribute of specular reflection), and their shape is usually three-dimensional rather than flat. Moreover, an object placed in the natural environment receives illumination from every direction as light is reflected from other objects in the scene, and the spectra of those lights might vary from one direction to another.

In this thesis, I report five research projects. The first three are all experiment based, and they were designed to find mechanisms that underpin human colour constancy for three-dimensional glossy objects that are placed under complex lighting environments. We conducted first an experiment where observers were asked to judge surface and illuminant colour changes for glossy objects under point light sources. The second and third experiments explored our ability to identify surface colours of glossy objects under environmental illumination that introduces a complex spatial pattern of specular highlights on the surfaces. All experiments used computer-rendered objects that take account of the complex interaction between material and illumination in the real world. Collectively these experiments suggest that observers are able to extract surface colour, even with complex stimuli in which simple heuristics are unlikely to be reliable. Instead, the results suggest that observers may build up complex internal models from which they can segregate colours of surface and illumination.

The fourth chapter is based on field work that was conducted in collaboration with Prof. Sergio Nascimento at University of Minho, Portugal. We report newly collected hyperspectral environmental illuminations recorded in eight outdoor scenes and four indoor scenes. The aim was to characterise the directional spectral variation that occurs in natural environments and its application to more precise simulation of objects using computer graphics. Subsequent fifth chapter was motivated by the striking variation of illuminant spectra in natural environments. We revisited a classical problem in human colour vision,

called illuminant metamerism where two distinct surfaces appear the same under a specific illumination. We showed theoretically that this metamerism is less of an issue than previously thought thanks to the presence of directional variation of illumination.

Contents

List of Figures **ix**

1 General Introduction **1**

1.1 Introduction to General Introduction 1
1.2 Basis of human colour vision 1
1.3 Stages of colour processing 4
1.4 Contextual effect of colour perception 7
1.5 Colour constancy 8
1.5.1 Potential cues to implement colour constancy 11
1.5.2 Problem of oversimplification - the RGB heuristic 14
1.5.3 Framework to implement colour constancy 14
1.5.4 Influence of higher-level process 15
1.5.5 Physiological substrates of human colour constancy 16
1.5.6 Experimental task 18
Asymmetric matching 18
Achromatic setting 19
Binary classification - red vs. green and blue vs. yellow 19
Colour naming and hue scaling 20
Discrimination between an illuminant-colour change and a surface-colour change 20
1.5.7 Quantification of colour constancy - constancy indices 21
1.5.8 Limitations of past colour constancy studies 23
Simplification of surface properties 23
Simplification of illuminant properties 27
1.6 Calibration of a CRT monitor 32
1.7 Overview of each chapter 34
Chapter 2 - Experiment 1 35
Chapter 3 - Experiment 2 35
Chapter 4 - Experiment 3 36
Chapter 5 - Field work 37
Chapter 6 - Theoretical work 38
Chapter 7 - Summary and Conclusion 38

2 When the brightest is NOT the best: illuminant estimation based on highlight geometry 39
2.1 Introduction . . . 39
2.2 Method . . . 41
2.2.1 Apparatus . . . 41
2.2.2 Rendering . . . 41
2.2.3 Stimuli . . . 42
2.2.4 Experimental task and creation of the animation set . . . 43
2.2.5 Observers and Ethics approval . . . 45
2.2.6 Procedure . . . 45
2.2.7 Simulation of the "brightest is white" heuristic and other models based on chromatic statistics . . . 46
2.3 Results and Discussion . . . 48
2.4 Conclusion . . . 51
2.5 Supplementary Information . . . 52
2.5.1 Rationale for why the "brightest is white" assumption fails . . . 52
2.5.2 The nature of the experimental task . . . 53

3 Reflectance discrimination under complex environmental illumination 57
3.1 Introduction . . . 57
3.2 Method . . . 61
3.2.1 Apparatus . . . 61
3.2.2 Stimuli . . . 62
Rendering . . . 62
Environmental Illumination . . . 62
Object shape and specularity . . . 63
Surface spectral reflectance . . . 64
Observers and Ethical approval . . . 65
3.2.3 Procedure . . . 65
3.3 Experiment 1 . . . 67
3.3.1 Conditions . . . 67
3.3.2 Results and Discussion . . . 69
3.4 Experiment 2 . . . 74
3.4.1 Conditions . . . 74
3.4.2 Results and Discussion . . . 74
3.5 General Discussion . . . 76
3.5.1 Overview . . . 76
3.5.2 Available cues . . . 78
3.5.3 Matte vs glossy objects . . . 79
3.5.4 Chromatically typical and atypical lighting environments . . . 80
3.5.5 Spatial signatures of lighting and reflectance changes . . . 81

4 Reflectance identification across complex environmental illuminations 83
4.1 Introduction . . . 83
4.2 General method . . . 87
4.2.1 Apparatus . . . 87
4.2.2 Experimental stimuli . . . 88
Rendering . . . 88
Environmental illumination . . . 88
Object shape and specularity . . . 89
Surface spectral reflectance . . . 89
4.2.3 Observers and ethics . . . 90
4.2.4 Procedure . . . 91
4.3 Experiment 1 . . . 93
4.3.1 Conditions . . . 93
Selection of viewing angle . . . 94
4.3.2 Results - human observers . . . 94
4.3.3 Implementation of computational observer model . . . 98
Single trial model . . . 98
Temporal integration model . . . 104
4.4 Experiment 2 . . . 108
4.4.1 Conditions . . . 108
4.4.2 Selection of environmental illumination and viewing angle . . . 110
4.4.3 Results and discussion . . . 111
4.5 General discussion . . . 114
4.5.1 Specular reflection . . . 116
4.5.2 Temporal integration of cues from surrounding context . . . 118
4.5.3 Future directions . . . 119
4.6 Appendix . . . 120

5 Hyperspectral environmental illumination maps: characterising directional spectral variation in natural environments 123
5.1 Introduction . . . 123
5.2 Method . . . 125
5.2.1 Hyperspectral imaging system . . . 125
5.2.2 Measurement . . . 126
5.2.3 Image Processing . . . 127
5.3 Results and Discussion . . . 130
5.4 Conclusion . . . 143

6 Breaking metamerism using directional spectral variation in illumination 146
6.1 Introduction . . . 146
6.2 Method . . . 148
6.3 Results and Discussion . . . 149
6.4 Supplementary materials . . . 153
6.4.1 An example of illuminant metamerism . . . 153
6.4.2 Set of surface spectral reflectances . . . 153
6.4.3 Environmental illumination . . . 154
6.4.4 Criteria to define indistinguishable surfaces . . . 156
6.4.5 Supplementary Analysis . . . 156

7 Summary and Conclusion 163
7.1 Chapter 2 . . . 163
7.2 Chapter 3 . . . 164
7.3 Chapter 4 . . . 165
7.4 Chapter 5 . . . 166
7.5 Chapter 6 . . . 167
7.6 Overall summary . . . 168
7.7 Future extensions of the current project and concluding remarks . . . 169

References 171

List of Figures

1.1 One functional advantage of colour vision 2
1.2 A schematic illustration of how human colour vision works 3
1.3 The MacLeod-Boynton chromaticity diagram that is used in this project to specify the colour coordinates of experimental stimuli 7
1.4 A typical example of contextual effects of colour perception 8
1.5 A demonstration of colour constancy 9
1.6 A schematic illustration of the colour constancy problem 10
1.7 Colour inconstancy 11
1.8 Conceptual illustration of illuminant estimation and correction from given surfaces 12
1.9 Calculation of a constancy index 22
1.10 An example of the simplified "Mondrian" stimuli typically used in colour constancy research 24
1.11 An example of a study that employed a computer-rendered 3-D scene . . 24
1.12 Two types of reflection (diffuse and specular) that occur when light hits an object 25
1.13 The spectral composition of sunlight and skylight 26
1.14 Colour distribution of rendered smooth objects under point-light sources 28
1.15 Colour distribution of rendered bumpy objects under point-light sources . 29
1.16 An RGB image of an environmental illumination map and its chromatic distribution 30
1.17 The effects of environmental illumination on the colour distributions of rendered smooth objects 31
1.18 The effects of environmental illumination on the colour distribution of rendered bumpy object 32
1.19 Measurements of the CRT monitor we used for experiments 34

2.1 Examples of different relative positions of specular highlights and the bright parts of a texture 40
2.2 The spectral composition of two test illuminants: sunlight and skylight . 42
2.3 Example trials showing a reflectance change or an illuminant change . . . 44

2.4 The chromaticities of 240 reflectance samples of natural objects and man-made objects when rendered under sunlight and skylight 45
2.5 Summary of experimental conditions 46
2.6 Schematic description of how the computational modeling works 47
2.7 Average *d'* across all observers for human observers and computational observers 49
2.8 An illustration of how chromaticity distributes on an object with spatial intensity variation 52
2.9 Histograms showing the distributions of the magnitude of chromatic change 56

3.1 Rendering objects in natural lighting environments 59
3.2 Chromatic properties of the three lighting environments used in the experiments 63
3.3 Chromatic properties of the surface spectral reflectance functions used in the experiments 66
3.4 Schematic illustration of the procedure 68
3.5 Summary of conditions in Experiment 1 68
3.6 Results from Experiment 1 70
3.7 Summary of reflectance discrimination performance across conditions of Experiment 1 72
3.8 Results from Experiment 2, with reflectance discrimination thresholds plotted in a reflectance-based plot 76
3.9 Summary of reflectance discrimination performance across conditions of Experiment 2 77

4.1 Images of four environmental illuminations and the chromatic distribution in the MacLeod-Boynton chromaticity diagram 89
4.2 Environmental illumination, selected surface spectral reflectances, and chromatic distribution of reflectances when rendered under equal energy white 91
4.3 All stimuli used for Experiments 1 and 2 92
4.4 Stimulus presentation for one trial 93
4.5 The selection of viewing angles in Experiment 1 95
4.6 Performance across all trials in Experiment 1. 96
4.7 Performance for individual pairs of target and distractor reflectances in Experiment 1 97
4.8 A schematic description of the model used in the present study 99
4.9 Average human performance across all trials compared to model performance 101
4.10 Performance across all individual reflectance pairs in Experiment 1 . . . 103
4.11 A comparison of human and model performance 105

4.12 A diagram to show the weighting of each trial for the temporal-integration model . . . 107
4.13 Percentage correct values and d-prime values across all conditions as a function of time constant . . . 109
4.14 Generation of the "congruent" and "incongruent" conditions of Experiment 2 111
4.15 Examples of stimulus presentation for Experiment 2 . . . 111
4.16 Model and human performance for congruent and incongruent conditions 114
4.17 d-prime across all time constants in Experiment 2 . . . 115
4.18 Simulation results for Experiment 1 with no noise in estimation of surface colour and illuminant colour . . . 121
4.19 Simulation results for Experiment 1 with double-strength noise in estimation of surface colour and illuminant colour . . . 122

5.1 How to measure environmental illumination . . . 127
5.2 The image-processing pipeline . . . 131
5.3 sRGB images of hyperspectral environmental illumination for 8 outdoor scenes and 4 indoor scenes . . . 133
5.4 Directional spectral variation as a function of azimuth ϕ and elevation θ for one outdoor and one indoor scene . . . 134
5.5 Average RMSE between the measured spectra and reconstructed spectra 136
5.6 Cone signals, opponent signals, and correlated colour temperature as a function of elevation θ and azimuth ϕ . . . 138
5.7 Heatmaps to show the spectral variation in outdoor scenes and indoor scenes 139
5.8 Application of spherical-harmonic expansion to environmental illuminations 141
5.9 Directional spectral variation for the remaining 7 outdoor scenes . . . 144
5.10 Directional spectral variation for the remaining 3 indoor scenes . . . 145

6.1 Spectral variation in reflected light from a white planar surface under an environmental illumination . . . 150
6.2 An illustration of the conceptual idea of breaking metamerism using directional spectral variation in illumination . . . 151
6.3 The proportion of solved metameric pairs for outdoor and indoor scenes . 152
6.4 An example of illuminant metamerism . . . 154
6.5 Chromatic distributions and histograms of the reflectance datasets for evaluation when rendered under equal energy white . . . 155
6.6 sRGB representations of the 10 environmental illumination maps used in this study . . . 155
6.7 Absolute numbers of indistinguishable pairs on a log scale . . . 157
6.8 The relative proportion of solved metameric pairs as a function of ΔE and the cut-off correlation coefficient . . . 159

6.9 Heatmaps to show which orientations of the samples solve indistinguishable surfaces more efficiently 161
6.10 The average proportion of solved metameric pairs using illumination from the upper and lower fields 162

1 General Introduction

1.1 Introduction to General Introduction

The purpose of this chapter is to review concepts that are important to understand this thesis. We start by introducing the basis of human colour vision, colour representation, and related phenomenon. Then, we introduce colour constancy research, with a particular emphasis on the history of proposed methods for illuminant estimation, the methodologies, and experimental stimuli used to assess colour constancy, and most importantly the various limitations in past studies.

1.2 Basis of human colour vision

Colour plays a crucial role in everyday life. The world might appear boring without colour, but this is of course not the only reason why we have evolved to have colour vision [1]. Figure 1.1 shows an example to highlight the functional importance of colour. Panel (a) presents red ripe berries surrounded by foliage with only luminance information. Although shape information is available, it is hard to detect the berries. Panel (b), on the other hand, provides chromatic information in addition to luminance information. The berries now pop out from their surroundings, and we can effortlessly and quickly detect them. As this example shows, colours help us to identify objects, in this case

by segregating objects from their surroundings. Therefore, colour is a useful perceptual attribute for humans to interact with our surrounding world.



Figure 1.1: One functional advantage to have colour vision. (a) A monochromatic image where red berries are hidden in foliage. Although shape information is presented in this figure, it is hard for us to detect the red ripe berries from the foliage. (b) The same image with colours. The red berries now pop out from the background. As this example shows, colour plays an important role to segregate objects from their background.

What then is colour? When we look at an object, it is tempting to think that the object has a colour. However, colour is not a physical quantity, and thus it does not physically exist outside the observer. Instead, all colours we see are percepts that our visual system creates. Figure 1.2 shows the basis of how human colour vision works. Illumination I(λ) emitted from a light source hits a surface that has a spectral reflectance R(λ), and the modified spectrum I(λ)R(λ) is reflected back to observer's eye. The reflected light that enters the eye is absorbed by three classes of cone, which are differently sensitive to short-, middle-, and long-wavelengths and the cone signals L, M, and S are generated as shown in equation 1.1 to 1.3. Here, L(λ), M(λ) and S(λ) denote the sensitivity of long-, middle-, and short-wavelength sensitive cones.

$$L = \int R(\lambda)I(\lambda)L(\lambda)d\lambda \tag{1.1}$$

$$M = \int R(\lambda)I(\lambda)M(\lambda)d\lambda \tag{1.2}$$

$$S = \int R(\lambda)I(\lambda)S(\lambda)d\lambda \tag{1.3}$$

Because cone signals are determined by the integration of reflected light and cone fundamentals L(λ), M(λ), and S(λ), when the lights are transformed to cone signals, information about the wavelengths is all lost (known as the principle of univariance [2]). Thus, comparisons between cone classes are the key to separating wavelength and intensity information. This has been identified as the basis of human colour vision [3], which has allowed us to model various functions of the human colour vision system in a biologically plausible way.

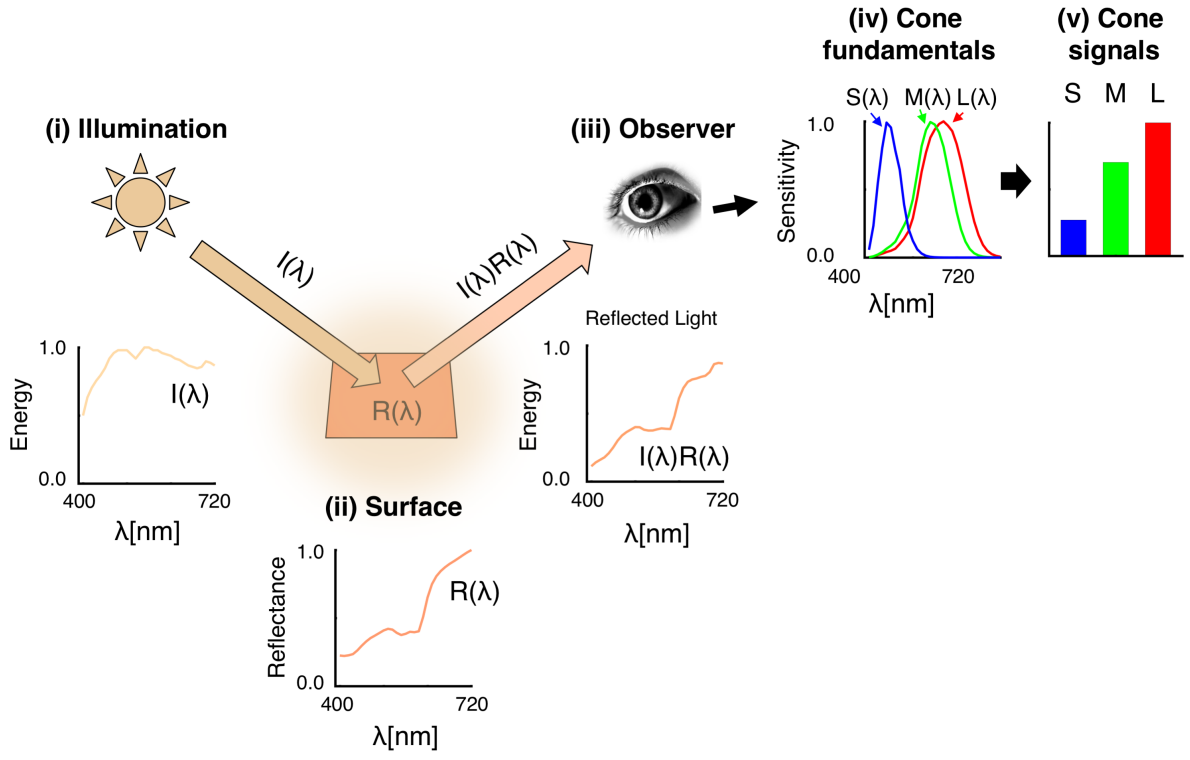


Figure 1.2: A schematic illustration of how human colour vision works. Illumination I(λ) emitted from (i) a light source hits (ii) a surface that has a surface spectral reflectance R(λ) and the modified spectrum I(λ)R(λ) is reflected back to (iii) the observer's eye. The reflected light is absorbed by (iv) three classes of cones to generate (v) cone signals L, M, and S. Based on the combination of these three cone signals, wavelength and intensity information can be disentangled.

1.3 Stages of colour processing

Neural processing of colour information progresses through several phases of transformation from one physiological layer to the next, starting from the signals in three types of cone photoreceptors in the retina. At the level of individual cones, wavelength and intensity information are mixed and not encoded separately. Cone signals are then sent via bipolar cells to retinal ganglion cells (RGCs) where the next stage of colour processing is thought to begin, and signals from RGCs are sent via the lateral geniculate nucleus (LGN) to the cortex. The post-receptoral phase of colour processing is chromatically opponent meaning that post-receptoral colour mechanisms are excited by some wavelengths and inhibited by others [4]. Physiologically speaking, chromatic opponency is implemented by comparing cone responses. There is good evidence that this chromatic opponency starts at latest at the level of RGCs. The midget ganglion cells, which have centre-surround receptive fields, receive opposed signals from the L and M-cones, and their signals are sent via the parvocellular layers in LGN. The parvocellular pathway is the primary candidate for generating the L-M, "red-green" chromatic channel [5]. Small bistratified ganglion cells, that were identified relatively recently [6], do not have centre-surround receptive fields, but receive inputs from all three types of cones with the S-cone signals opposed to L-cone and M-cone signals. Signals from small bistratified ganglion cells are sent to the koniocellular layer in the LGN and this pathway is known as the S-(L+M), "blue-yellow" channel. Finally, parasol ganglion cells receive summed inputs from L- and M-cones and send signals to the magnocellular layers of the LGN, and are thought to underlie the L+M or luminance channel, which is achromatic [7]. These three channels (L-M, S-(L+M), and L+M channels) are thought to be independent, and together form the "cardinal" directions of the human colour vision system [8]. This hierarchical processing from trichromacy (retina) to chromatic opponency (RGCs, and LGN) has been acknowledged as a multi-stage colour model [9].

Neurons that have hue selectivity are identified in several different cortical areas. In primary visual cortex (V1), double-opponent cells contained in so-called blobs are found to show chromatic preferences and are thought to contribute to detection of

chromatic edges [10]. V2 has an anatomical feature known as "thin stripes" and cells there seem to show preferences to specific hues [11]. The potential role of V4, posterior inferotemporal area (TEO), and dorsal posterior inferotemporal cortex (PITd) are further complicated, but they have been claimed to have "glob" cells that are narrowly tuned to specific hues [12]. Note that the cortical processing of colours is less clear than earlier stages of colour vision, such as in retina or LGN. Past findings are systematically summarised by several review papers [13]–[15].

Here, we note that some colour spaces have been specifically inspired by the properties of the underlying neural substrate. For example, the Derrington-Krauskopf-Lennie (DKL) colour space was proposed based on cell recordings in macaque LGN [16]. In this DPhil thesis, we use the MacLeod-Boynton (MB) chromaticity diagram [17], which also represents colour processing at opponent stages. The horizontal axis is L/(L+M) and expresses the ratio of signals in the long- and middle-wavelength sensitive cones, while the vertical axis is S/(L+M) and expresses the signal in the short-wavelength cones in relation to a combination of long- and middle-wavelength sensitive cone signals. We used the Stockman and Sharpe cone fundamentals [18] to calculate LMS cone excitations. The MacLeod-Boynton chromaticity diagram forms an equiluminant plane, which allows us to specify chromaticity and luminance independently. Specifically, when we calculate MB coordinates and luminance from L, M, and S cone signals, we used following equations. Here, L, M, and S signifies the integrated L, M, and S cone signals assuming spectral sensitivity functions that are scaled to peak at one.

$$redness = \frac{0.68990L}{0.68990L + 0.34832M} \tag{1.4}$$

$$blueness = \frac{1.9354S}{0.68990L + 0.34832M} \tag{1.5}$$

$$luminance = 0.68990L + 0.34832M \tag{1.6}$$

The scaling factor for each cone signal was defined as follows. For L and M cone signals, scaling factors 0.68990 and 0.34832 were determined so that 0.68990L(λ)+0.34832M(λ)

matches the luminous efficiency function, which accounts for the human visual system's sensitivity to luminance [19]. For S cone signals, the factor 1.9354 was determined so that the *blueness* of equal energy white becomes 1.000. Because of the arbitrary scaling of each axis, absolute values of MB coordinates do not generally convey concrete meaning. Rather they need to be used in relative sense. Under the calculations defined by equations 1.4 - 1.6, the *redness* and *blueness* of equal energy white corresponds to 0.7078 and 1.000, respectively (marked by black cross symbol in Figure 1.3). It should also be noted that this diagram is defined on a physiological basis and not designed to be perceptually uniform. In the rest of the thesis, we use L/(L+M) and S/(L+M) when referring to "redness" and "blueness" respectively, as defined in equations 1.4 and 1.5.

Figure 1.3 (a) shows the whole MB chromaticity diagram. We calculated the chromaticities of all monochromatic lights from 380 nm to 830 nm in 1 nm steps. This so-called spectrum locus is shown by a black line. The black cross symbol indicates the chromaticity of equal energy white. Panel (b) shows an enlarged figure around equal energy white and the panel is pseudo-coloured by sRGB to show how colour appearance changes as the chromaticity coordinates change. Although, as seen in Figure 1.3 (a), the chromaticity diagram extends widely especially in the S/(L+M) direction, the chromaticities that can be presented on an experimental display are severely limited. Towards the end of the General Introduction, we introduce the gamut of the CRT monitor that we used for experiments.

We also note that these physiology-based chromaticity coordinates do not consider the perceptual discriminability of colours. Thus, equal distances picked from different regions in the chromaticity diagram are not guaranteed to have equal subjective difference. For cases where perceptual equality is desired, CIE has developed several perceptually uniform colour spaces over decades. The current standard is CIECAMO2-UCS [20], which considers observer's adaptation state as well as the white point.

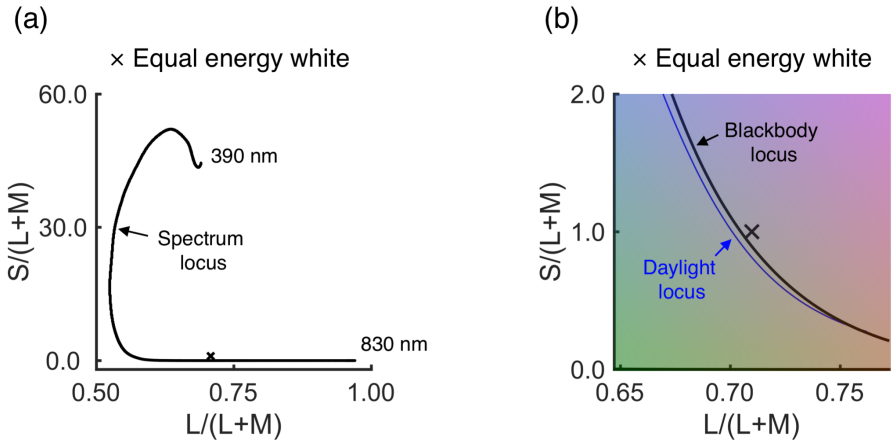


Figure 1.3: The MacLeod-Boynton chromaticity diagram that is used in this project to specify the colour coordinates of experimental stimuli. (a) The whole region of the chromaticity diagram. The black solid line indicates the so-called spectrum locus, which is the set of chromaticities produced by monochromatic lights, ranging from 390 nm to 830 nm in this case. (b) A narrower region of the diagram in which the background shows the sRGB representation of how colour appearance changes within MacLeod-Boynton colour space. The black and blue solid lines indicate the black body locus and the daylight locus, respectively.

1.4 Contextual effect of colour perception

As suggested by the preceding discussion of the stages of chromatic processing, our perception of colour is not purely determined by cone signals in our retina. One common way to show this is to consider a contextual effect on colour appearance [21]. The left image in Figure 1.4 appears as a purple square surrounded by a red square and the right image appears as a more reddish purple square surrounded by a blue square. The centre squares on the left and the right side probably appear different to us. However, they have exactly the same RGB values, so the lights coming out from these small two square regions to your eyes are physically identical. This example is called simultaneous contrast and demonstrates that human colour perception can be modified when the surrounding context is changed. There are many other known chromatic contextual effects that modulate our perception of colour in complex ways [22]–[27], through which various underlying mechanisms have been identified.

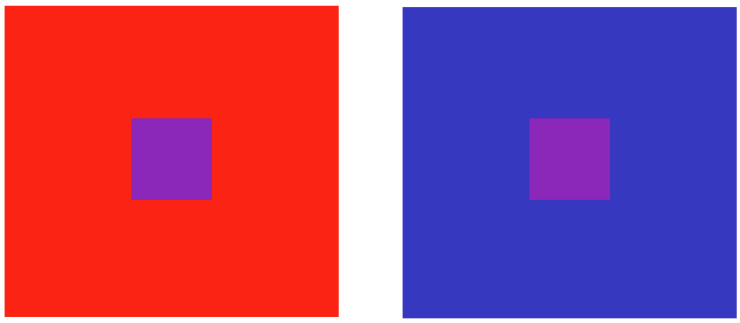


Figure 1.4: A typical example of contextual effects of colour perception, called simultaneous contrast. Two small centre squares might appear very different to us, but they have exactly the same RGB values. Thus, the lights coming from these regions are identical, but their appearance is modified influenced by surrounding colours. (This figure is best seen on a computer display)

1.5 Colour constancy

When the scene illumination changes, there is a dramatic change in the light reflected from the surfaces of objects in the scene. Thus, the physical input to the visual system depends largely on the illumination. Nevertheless, perceived surface colours change little. This ability of the human colour vision system is called colour constancy, and it helps us to perceive an object's colour stably under different illuminations. Figure 1.5 shows two rooms illuminated by two different light sources. The left room is under white illumination and the right room is under blue illumination. Yet, we perceive the sheet of paper on the table as being white in both cases. However, the light reflected from the paper on the right table is actually bluish, influenced by the bluish illumination, as shown by the square below the right-hand figure. This simple demonstration shows that the visual system discounts the influence of illumination to allow us to see the white paper as white under different illuminations.

However, this discounting of illumination is not a simple matter to solve. Let's consider the situation depicted in Figure 1.6. Here, (i) the illumination has changed from $I_1(\lambda)$ to $I_2(\lambda)$. (ii) The surface spectral reflectance $R(\lambda)$ stays the same, but as a result of the illuminant change, the reflected light from the surface to (iii) the observer accordingly changes from $I_1(\lambda)R(\lambda)$ to $I_2(\lambda)R(\lambda)$ which generates a different set of (v) cone signals. Now, in order to see the surface colour stably, we need to discount the influence of illumination. However, from the cone signals alone, we are not able to work

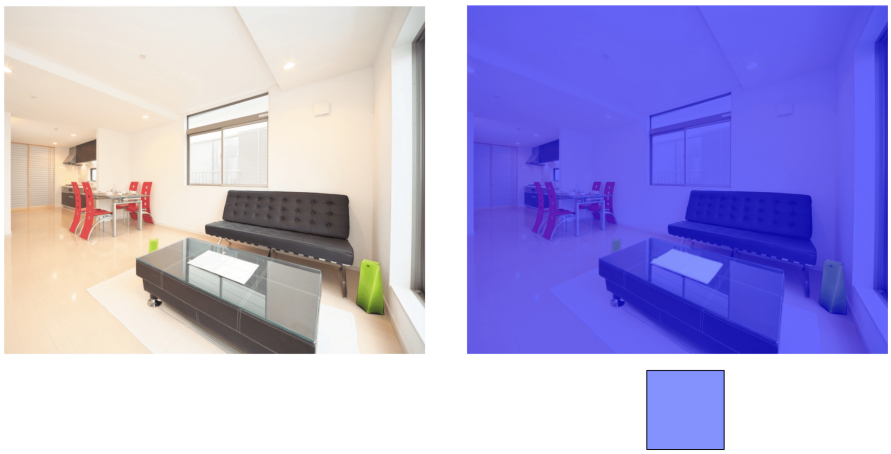


Figure 1.5: A demonstration of colour constancy. Two rooms are illuminated by different light sources, but we can recognise that the papers on the focal tables as white in both rooms. However, the blue square below the right panel has exactly the same colour as the "white" paper in the right-hand room. (This figure is best seen on a computer display)

out how much of the change in cone signals is caused by the illuminant change, since the same change in cone signals could arise from a change in surface reflectance. Thus, the influence of illumination cannot be uniquely extracted from cone signals. This is the nature of the colour constancy problem, which is a mathematically ill-posed problem. The phenomenon of colour constancy was identified three centuries ago [28]. The first behavioural study was conducted decades ago [29], and since then colour constancy has been historically a rich research domain as summarised in review papers [30], [31]. Yet, the mechanisms of colour constancy are not fully understood.

Because of these computational difficulties, it should not be very surprising if colour constancy fails. Nevertheless, in daily life, colour constancy works reasonably well and we do not even notice that our surface colour perception is supported by the success of colour constancy. Thus, it is easy to think that colour constancy is a complete mechanism that does not make a mistake. However, it is important to remember that many lab-based experiments showed that colour constancy does not work perfectly in general. Figure 1.7 demonstrates rare examples that shows the failure of colour constancy. Panel (a) shows a picture of grey building which is illuminated by orange illumination [32]. However, for most people, this picture appears a grey building the top of which is painted by orange paint. In other words, our brain has attributed that orangeness to a surface property

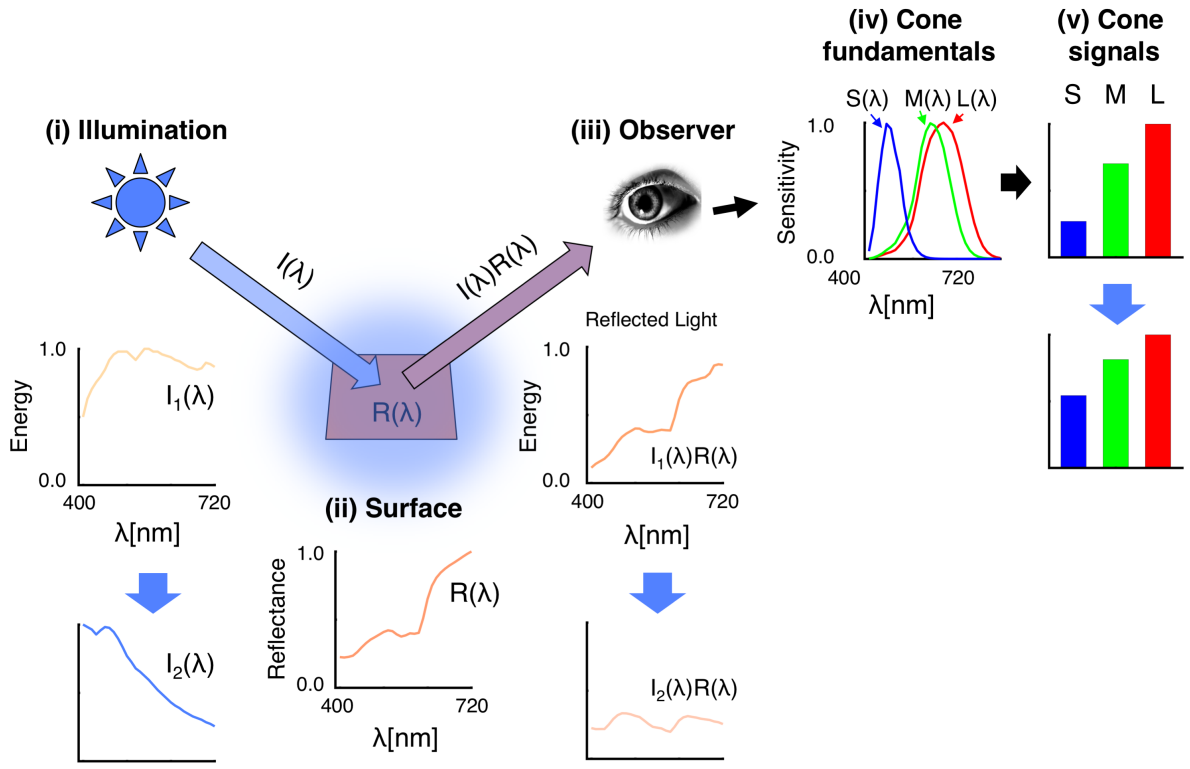


Figure 1.6: A schematic illustration of the colour constancy problem. An illumination change produces different reflected lights from the surface, which in turn generate different cone signals. To achieve colour constancy, we need to subtract the influence of the illumination. However, with only information about the cone signal change, it is mathematically impossible to extract the influence of illuminant uniquely.

rather than to an illumination property. Panel (b) shows a relatively recent picture which spread as a viral phenomenon via the internet, known as #TheDress. The world's population was divided into mainly two groups depending on whether they saw this picture as a blue and black dress or as a white and gold dress. One potential reason why this phenomenon happened is that the illumination falling on this dress was interpreted differently by different individuals. Those who saw dress as blue and black estimated the illumination as white and intense. On the other hand, if the dress appeared white and gold, the illumination was interpreted as the dark bluish. In summary, the dress picture revealed that there are surprising individual differences in human colour constancy. The failure of the elaborate colour vision mechanism surprised us, but at the same time it gave us a hint to understand the underlying mechanisms of colour constancy. Since this image was posted on social media, many researchers have attempted to understand this image

and offered a variety of potential explanation for this mysterious phenomenon [33]–[51].

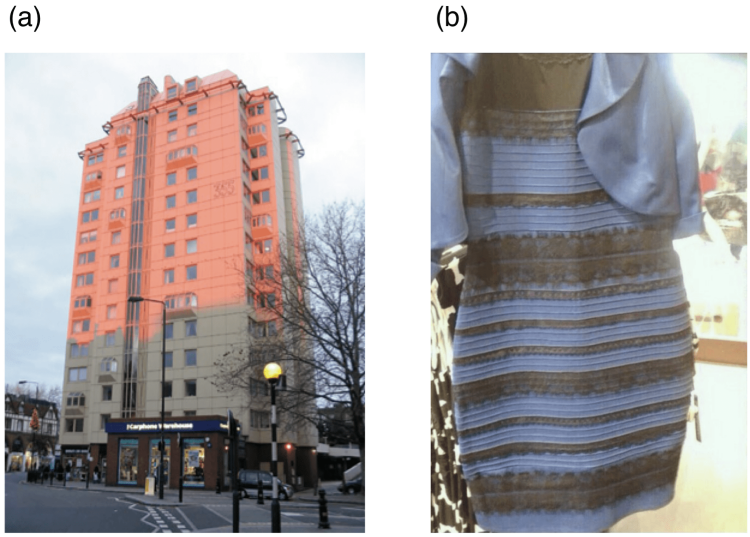


Figure 1.7: Colour inconstancy. (a) A gray building illuminated by orangish illumination [32]. However, most people report this building appears to be painted orange. (b) An ambiguous image known as #TheDress. Some people perceive this dress as white and gold while others perceive it as blue and black. This phenomenon quickly spread via the Internet and provoked discussion among vision scientists.

1.5.1 Potential cues to implement colour constancy

One common way to conceptualise colour constancy is that the visual system first estimates the colour of scene illumination based on cues available in the scene, and then discount its influence from the whole scene. Although the exact mechanism for colour constancy is not fully understood, past colour constancy research has been successful in identifying a number of cues thought to underlie colour constancy. Perhaps the simplest algorithm would be to find the brightest surface in the scene. This is based on the observation that a white surface that has a completely flat spectral reflectance reflects any lights in a spectrally unmodified way, meaning that it carries direct information about the colour of illumination [52]. Another simple but powerful transform would be to calculate mean chromaticity across the whole scene and assume that it is a good estimate of the illuminant colour [53]. This is based on the idea that the average colour across all objects in a scene is typically grey (grey-world hypothesis [53]), and thus the deviation from grey can be assumed to be due to the influence of illumination. It is also

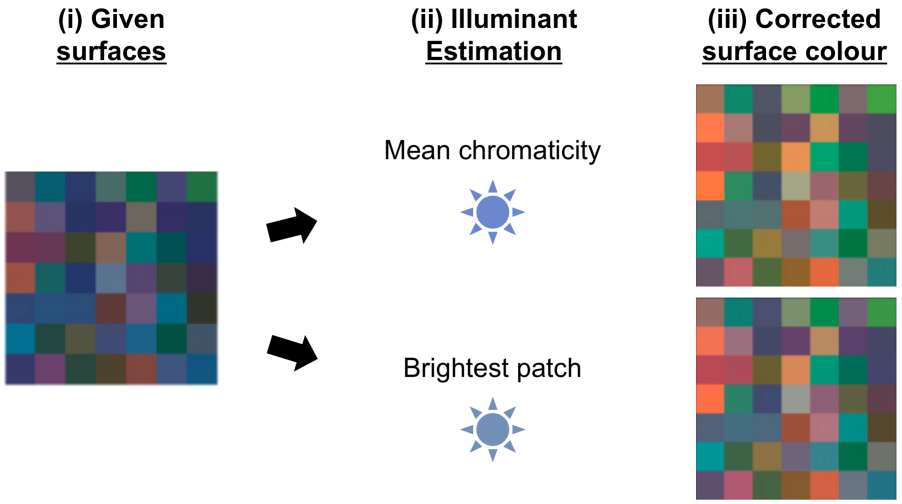


Figure 1.8: Conceptual illustration of illuminant estimation and its correction from given surfaces. From (a) given surfaces, (ii) two candidate models based on mean chromaticity and brightest patch estimate different colours of illumination. Estimated illuminant colour was subtracted from the original surfaces to produce (iii) the corrected surface colour.

known that spatial mean cone signals (rather than mean chromaticity) would be a good predictor of an illuminant colour [54]. The notion behind the use of these global statistics is that illumination typically lights a scene globally. Thus, when the scene contains many different surface spectral reflectances, the global mean estimates the influence of illumination. Figure 1.8 shows (i) example surfaces and subtraction of the influence of (ii) illumination estimated by the mean chromaticity and the brightest patch from those surfaces. (iii) The resultant image would appear as surfaces under neutral illumination.

Using global scene statistics (e.g. mean chromaticity) to estimate the colour of illumination has an attractive simplicity. However, whilst the use of global scene statistics reasonably account for experimental data, one of the longstanding problems in the field of human colour vision was how our visual system distinguishes a white scene illuminated by a reddish illumination and a reddish scene illuminated by whitish illumination when both cases produce the same reddish mean chromaticity. If our visual system relies purely on mean chromaticity, estimation of an illuminant colour for these two situations should both be the same red. However, we empirically know that colour constancy does not collapse in such a situation. Thus, we seem to be able to hold colour constancy even for a scene that consists of a biased surface colour set (e.g. dominantly reddish surfaces).

A milestone study [55] provided a solution to this question. Firstly, they theoretically showed that red surfaces tend to have higher luminance when the illumination gets reddish compared to when the illumination is white. In contrast, when the illumination is white, this chromatic-dependent luminance shift does not occur. Therefore, by looking at the correlation between luminance and chromaticity of given surfaces in a scene, it is possible to work out the colour of illumination lighting the scene. Secondly, importantly, they have shown psychophysically that human observers are able to use this cue to solve such an ambiguity. Similarly, other researchers pointed out that the luminances and chromaticities of surface colours do not randomly distribute in a scene. Instead, they distribute in a systematic way depending on the colour of illumination. They have also experimentally shown that this relation between luminance and chromaticity plays a crucial role to solve ambiguities [54], [56], [57]. Although some studies argued that the use of luminance-chromaticity correlation might be limited [58], [59], luminance information seems to play a key role as well as chromaticity of given surfaces to implement colour constancy.

Colour constancy is also a large research domain in computer vision, and thus numerous computational algorithms that use higher order image statistics have been proposed in past research [60]–[62]. Also, recent convolutional networks have also been used to tackle the problem, and are outperforming conventional heuristics-based algorithms [63].

Similarly, statistical models such as Bayesian estimation [64], [65] or low-level linear models to recover surface spectral reflectance [66] were reported to be good candidate models of human colour constancy. Interestingly, these theoretical models make use of statistical regularities of surface spectral reflectance and illuminant spectra that occur in the natural environment. This might imply that the human visual system utilises such constraints in the natural environment to solve mathematically ill-posed visual problems such as colour constancy.

In either case, however, to answer whether those models can be a good predictor of our visual system's behaviour, we would need to conduct experimental verification. This is simply because finding something that is in theory useful does not guarantee that our visual system can or does use it.

1.5.2 Problem of oversimplification - the RGB heuristic

We argued above that colour constancy can be thought to be a process to first estimate the cone signals of the illumination, and then to use that to transform the cone signals of the surface to the values they would be if seen under neutral illumination (e.g. equal energy white). However, to what extent is this approach mathematically plausible? Let's assume that we are looking at a surface with surface reflectance $R(\lambda)$ illuminated by an illuminant that has spectral composition $I(\lambda)$. L cone signals generated by the reflected light $R(\lambda)I(\lambda)$ from the surface is L_{RI}. If we have access to the illuminant itself, for example by directly looking at it, we get L_I. Now, our goal here is to estimate the cone signal L_R generated by the surface spectral reflectance itself or seen under equal energy white. One might imagine this L_R can be obtained by calculating L_{RI}/L_I. However, there is no mathematical reason that this should work since the illuminant is mixed with the reflectance in a wavelength-dependent way. Therefore, even if our visual system can access the exact cone signals produced by the spectral composition of scene illuminant (using a variety of cues available in the scene), it is not mathematically guaranteed that we can fully (or even approximately) recover the cone signals of the surface viewed under neutral illumination. This simplification of the colour constancy problem is termed "RGB heuristics" [67] and highlights why perfect colour constancy is computationally difficult.

1.5.3 Framework to implement colour constancy

Both the difficulty in estimating the tristimulus values of illumination accurately, and the fallacy of "RGB heuristics", makes colour constancy challenging. Nevertheless, the existence of behavioural colour constancy indicates that our visual system overcomes such difficulties at least to some extent. To model and understand how our brain does it, we need a framework of how to implement colour constancy once the illuminant colour is estimated.

In a milestone study, Brainard and his colleague explored what sort of framework can best predict the observer's settings in asymmetric matching experiments [68], or more precisely the least complex framework that can account for observers' behaviour. In the study, they tested four different frameworks, three existing ones and their novel proposal. Firstly, the most complex framework was the affine transformation, which first uses a 3×3

matrix to transform cone signals, and then adds gains (3×1 matrix) to the transformed cone signals. Thus, it has nine parameters in multiplicative gain (elements in 3×3 matrix) and three parameters in additive gain (elements in 3×1 matrix), which results in a total of 12 parameters. Second, they used a linear transformation, which is a special case of affine transformation, where the additive gains are set to zero, and therefore uses nine parameters. These two models can be thought of as an imitation of colour opponent process as interactions across cone classes are considered. A third simpler framework is a diagonal transformation, where each class of cone signal is scaled separately by a single scalar. Physiologically, this is analogous to von Kries adaptation [69] (detailed later in this chapter). They fitted these three models using standard regression to minimise the squared error of the fit to the asymmetric matching data they obtained in the experiment.

The proposed framework was the equivalent illuminant model. In this framework, the illuminant estimated by observers was called the "equivalent illuminant", and for any given equivalent illuminant, the model can predict the asymmetric match for any test surface. For the form of transformation used in that matching process, the standard linear-model-based colorimetric techniques, which rely on only three parameters, was used. Based on measured asymmetric matching data, they searched for the equivalent illuminant that gives the best fit to the measured datasets. Note that the model assumes that observers performance error comes purely from the error in estimating an illumination colour. It should also be noted that the equivalent illumination framework is a special case of a linear model. Therefore, it can perform at best as well as a linear model. Nevertheless, they found that this model was almost as good as a linear model that uses 9 parameters. Since the emergence of this framework, it has been applied in several studies that use more complex stimuli and it has been shown that it can indeed account for empirical data [70], [71]. A full description of the equivalent illuminant model and other models are available in the appendix in [68].

1.5.4 Influence of higher-level process

As argued above, various studies have made an attempt to understand human colour constancy by trying to find frameworks that work based on bottom-up (cone or opponent

signals) available to the visual system. However, such models cannot be the full description of human colour constancy, as even when you have equated low-level signals, such as cone responses, the degree of colour constancy can vary. A complete colour constancy model should account for those higher level influences. For example, there are several studies that report that colour constancy works differently in three-dimensional scenes compared to in two-dimensional scenes such as a still image on a computer or a picture hanging on a wall [72]–[79]. Similarly, object movements can potentially provide cues to solve colour constancy [80], [81]. Moreover, memory of colour or knowledge of the object also aids colour constancy as we can refer to our memory to know the "true" surface colour regardless of illumination [82]–[86].

1.5.5 Physiological substrates of human colour constancy

Compared to the large number of perceptual studies, there is less research on the neural mechanisms that support human colour constancy. However, there is a general consensus that colour constancy is not implemented by a specific type of cell or areas in cortex, but rather the whole visual system together works to suppress the influence of illumination.

There is a rich amount of empirical evidence that retinal adaptation plays a vital role in achieving colour constancy. The most well known mechanism is von Kries adaptation [69], which scales the strength of each cone signal in inverse proportion to the level of stimulation. Importantly, this scaling is performed separately for each cone class, and acts spatially locally. For example, if the scene is illuminated by reddish illumination, L cone excitation will be high on average over time. Therefore, if you stay in this illumination for a prolonged time, your L-cone sensitivity will decrease, and eventually reddish illumination will appear less reddish than when you entered the illumination. Also, there is empirical evidence that human colour constancy can indeed act locally [87]. In contrast to von Kries adaptation, the Ives transformation takes account of the whole scene more globally [88]. In practice, assuming a relatively long time-constant of adaptation, this can be readily implemented by means of eye movements. Assuming the visual system samples the scene uniformly from one location to another, a part of the retina that is used for judging colour (e.g. fovea) would adapt to the mean cone

signals across the whole scene, which would be useful to implement colour constancy as the illumination usually affects the scene globally.

There are other potential sites that contribute to colour constancy computations in the retina. For example, physiological recording from horizontal cells in monkey retina revealed that adaptation of those cells is cone specific and can act spatially locally which therefore follows the rule of von Kries' adaptation [89].

Once the cone signals get out of our retina, signals reach the lateral geniculate nucleus (LGN). There is evidence that neurons in the LGN recorded from anaesthetised macaques are found to adjust their sensitivity after being exposed to adaptation light that has a simple centre-surround configuration, which could shift the perceptual white point [90].

Dichoptic presentation, which projects different images to right and left eyes, allows us to psychophysically separate the retinal contribution to perception as it is believed that the signals from the two eyes are not combined until at least V1. Using such a technique it has been shown that the level of colour constancy decreased with dichoptic presentation but importantly still largely remained, suggesting that there is strong support from higher-order mechanisms on colour constancy in addition to retinal contributions [91].

One way to see whether a specific cell codes colour constancy is to look at whether it maintains a constant signal despite a change of illumination. With this view, several studies have identified such colour constant cells in area V4 of either rhesus or cynomolgus monkeys [92]–[95]. Other studies more directly reinforced such view by showing the failure of colour constancy in lesion studies of rhesus monkeys [96], [97]. However, influences from other cortical areas have been also reported. A wider range of cortical areas, V1, V2, V3 and V4, seem to be involved in the implementation of the colour constancy process [91] revealed by functional magnetic resonance imaging (fMRI) measurement from human participants. Another study also directly suggested the contribution from V1 [98]. However, all of this evidence is found for simple visual stimuli, either coloured lights that have a simple spatial configuration (either centre-surrounded concentric circles or squares) or Mondrian stimuli. Therefore, whether these claims can be applied to more complex stimuli, where objects have three-dimensional shape and are illuminated from every direction, remain largely unclear.

In summary, studies that have identified physiological substrates for human colour constancy are still limited, and we do not know enough yet to identify the numerous mechanisms that support colour constancy. Nevertheless, it seems to be reasonable to say that colour constancy is a not a process that happens at a specific level. Rather, it starts from retina, and is further reinforced during processes in LGN, V1, and higher visual cortices. The availability and sophistication of brain imaging techniques such as fMRI has advanced, and several studies have started finding colour constancy representations of cells as early as V1 [99], [100]. It is expected that more evidence will emerge in future years.

1.5.6 Experimental task

In the past decades, various methods to quantify the degree of colour constancy have been developed. This subsection briefly summarises them.

Asymmetric matching

One common way to quantify colour constancy is called asymmetric matching [29], [64], [101]–[109]. The term matching normally refers to a set-up where observers match one attribute of the reference stimulus (e.g. orientation) to the target stimulus. For asymmetric matching in colour constancy studies, participants do this task for reference and target stimuli that are placed under different illuminants. It is also important to note that there are two distinct criteria that participants can rely on when making their judgement. In the study by Arend and Reeves (1986) [29], observers were instructed to match stimuli so that two surfaces (i) appear as if they were cut from the same piece of paper or (ii) match in terms of hue and saturation. These criteria are called surface-colour-match (or paper match) and appearance-match (or hue and saturation match). Interestingly, the degree of colour constancy was substantially higher when participants were asked to judge based on a surface-colour match criterion than a hue and saturation match criterion. Thus, it is important to keep this difference of criterion in mind when designing an experiment, and precise instruction to the participants is essential. This matter is discussed in more detail elsewhere [110].

Achromatic setting

Achromatic setting has been also commonly used in past studies. In this task, observers are required to adjust the chromaticity of a target stimulus so it elicits an achromatic percept under a specific illumination [55], [75], [76], [111]–[118]. One advantage of achromatic setting is that the task is easy for naive participants to understand. It also takes a relatively short time to make a judgement. However, the point producing "achromatic" comes from observer's internal criterion unlike asymmetric matching where an external reference is given. Therefore, it might be hard to make a comparison across observers. It is thus important to couple this method with a relative measurement to quantify the degree of colour constancy. For example, past studies typically compared achromatic points across different illuminant conditions within an observer, which is expected to at least partially cancel out such criterion differences. The way of quantifying colour constancy is detailed in the next subsection.

Conceptually, measurement of the achromatic point is a powerful technique to study colour constancy because the observers' subjective achromatic point corresponds to what they think is the chromaticity of illumination. Thus, the result gives us not only information about a chromaticity change of a specific surface (as would be the case for asymmetric matching with a white surface), it gives us an important indication of the observer's estimation of the illuminant.

Also, for this achromatic setting, aforementioned criteria differences about surface-colour-match and appearance-match can be applied. Thus, it is again important to give a careful instruction so that participants can clearly understand whether they are asked to adjust the test patch so that it simply appears achromatic or asked to make it appear as an achromatic surface placed under the illumination.

Binary classification - red vs. green and blue vs. yellow

Instead of directly asking participants to set the test patch to achromatic, Smithson and Zaidi (2004) [87] used a binary classification task in which participants were simply asked whether the presented surface appeared (i) reddish or greenish and (ii) yellowish or bluish. If participants repeat this task for a large number of trials, estimates of unique hue loci

(neither red nor green and neither blue nor yellow) can be drawn, and intersection of those two curves will give an achromatic point that appears neither red, nor green, nor yellow nor blue. The advantage of this method is that we are able to compare how the perceptual chromatic loci change depending on illumination as opposed to the single point, which is often the only measurement obtained in other methods.

Colour naming and hue scaling

Several studies have used colour naming as a measure of colour constancy [106], [119]–[122]. Colour naming is a natural task in the sense that people use colour terms in a daily life. Names may be chosen from a fixed set of colour terms [123], [124], or participants may freely name their subjective appearance. However, the lack of chromatic resolution could be an issue. The number of colour terms commonly used is limited, whereas matching procedures allow participants to adjust the colour of the test patch continuously. For example, the term blue does not indicate a specific chromaticity or luminance. Rather it includes some range of blues, and the range may vary from one individual to another.

Discrimination between an illuminant-colour change and a surface-colour change

Assigning an observed colour change to either an illuminant change or a surface-colour change is sometimes used as a measure of colour constancy [31], [125]–[128]. In contrast to other methods stated above, this method is a performance measure where the experimenter can objectively score accuracy (correct vs incorrect) on each trial, and trials are repeated many times to give an indication of how colour constancy held. It does not provide a specific chromaticity or luminance of the estimated illuminant or perceived surface colour, but better discrimination ability directly indicates better colour constancy performance since it shows that participants correctly identify the source of the change as coming from the surface or the lighting. Importantly, it is a powerful tool to quantify the degree of colour constancy especially when the classic methods above would not be possible to apply, for example when the stimuli become more complex than a flat and spatially uniformly coloured patch. To implement such a method, real surfaces and illuminations would be difficult to use as fine control of colours can be difficult especially for surface colour

changes. Thus, early studies used flat surfaces to simulate two situations and they were presented on a computer-controlled monitor. However, the difficulty associated with such simple stimuli is that it may not look as though the material colour or illuminant colour is actually changing. As a result, there is a risk that observers are forced to discriminate simply between situations A and B regardless of whether they appear material change or illumination change. The advance in computer graphics techniques has allowed us to create much more realistic stimuli [129], which might reduce this problem. Problems associated with the use of simplified stimuli are discussed in more detail in the next section.

1.5.7 Quantification of colour constancy - constancy indices

To understand the extent to which colour constancy works under a specific circumstance, it is useful to quantify the degree of colour constancy.

To give a concrete example, we assume that an observer made an asymmetric match, under the conditions shown in panel (a) in Figure 1.9. The left-hand and the right-hand stimuli are presented under 6500 K and 4000 K illuminations, respectively. The observer's task was to adjust the chromaticity of the test patch under the 4000K light so that it appeared the same as the target patch under 6500K. Thus, we obtain the chromaticity of the test patch as a result. Next, we introduce the commonly used constancy index (CI) and Brunswik ratio (BR). They are defined by equations 1.7 and 1.8, where a, b and c indicate the chromatic distances defined in panel (b) in Figure 1.9. The orange circle symbol indicates the chromaticity of the test patch that an observer found to be matched to the target patch in colour. The gray diamond symbol indicates the chromaticity of the target patch. The orange diamond indicates the chromaticity of the target surface when placed under the 4000K, and thus this indicates the ideal match point. For the calculation of this chromaticity, it would be necessary to have information about the surface spectral reflectance of the target patch or to make a simplified assumption about illuminant transformations.

$$CI = 1 - \frac{b}{a} \tag{1.7}$$

$$BR = \frac{c}{a} \tag{1.8}$$

Here, a value of zero indicates no constancy, and higher value indicates better constancy for both CI and BR. For both indices, if the observer's setting exactly falls on the ideal match, CI = 1, and BR = 1. On the other hand, if the setting is on the chromaticity of the surface under 6500K (gray diamond), both indices become zero. Note that maximum values for CI is 1.0 while BR can be more than 1.0.

Finally, it is worth noting that these indices can be easily applied to the data obtained by achromatic settings, simply by replacing the chromaticities of diamond symbols with the chromaticities of illuminants themselves (6500 K and 4000 K, in this case).

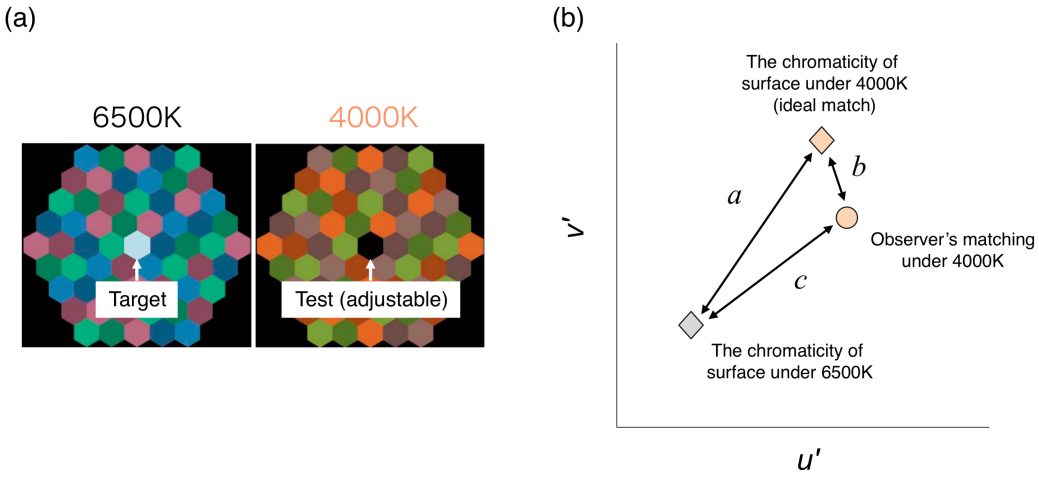


Figure 1.9: (a) A hypothetical experiment to give an example of how to calculate a constancy index. The observer's task is an asymmetric matching task where she/he is asked to adjust the colour of the test patch on the right so that its appearance matches with the colour of the target patch on the left. (b) The definition of the colour constancy index. The orange circle indicates the observer's settings and the gray and orange diamond symbols indicate the chromaticities of the surface when placed under illumination with colour temperatures of 6500 K and 4000 K, respectively. See equations 1.7 and 1.8 in the main text for the definition of indices.

Alternative indices may be defined in a different way, but the main concept here is to measure how much the perceptual appearance of test patches changed under the illuminant change relative to the physical change of illuminant chromaticity. The choice of colour space is also important, and generally uniform colour spaces are preferred to accurately reflect the perceptual differences.

1.5.8 Limitations of past colour constancy studies

Colour constancy has been remarkably resistant to analysis. For example, colour constancy may seem to be explained by retinal adaptation, but removing the effect of retinal adaptation (for example, using brief stimulus presentation) does not completely eliminate colour constancy. This suggests that its mechanisms do not seem to rely on specific neural or computational process, but instead rely on processes at various stages of the colour vision system from retina to cortex. Nevertheless, past colour constancy studies have been successful in identifying potential cues for human observers to implement colour constancy as discussed above, which significantly advanced our understanding.

However, a limitation of past studies is that they have mostly used simplified stimuli, in which objects were typically two-dimensional, matte and uniformly illuminated by a single light source [29], [55] (so-called "Mondrian" stimuli as shown in Figure 1.10). Such stimuli have the advantage of allowing careful experimental control, but they lack some of the important properties of surfaces and illuminants that may offer cues for colour constancy. In the past two decades, researchers have started addressing these questions using more realistic objects in three-dimensional setups [73]–[75], [118], [127], [130] (e.g. Figure 1.11). Moreover, recent advances in computer graphics techniques have enabled experimenters to investigate various features of material perception [131]–[133]. In addition to research on gloss and lightness perception [134], [135], an increasing number of studies on colour perception have been conducted [78], [79], [129], [136]–[141]. However, colour constancy with such computer-rendered, realistic stimuli is still little explored.

The next subsection describes concrete examples of simplification regarding surface properties and illuminant properties, and how adding complexity can change the chromatic signals that reach our eyes from objects' surfaces.

Simplification of surface properties

One simplified surface property is the specularity of objects. The matte objects used extensively in past studies exhibit only diffuse reflection whose spectral content is the product of the spectral energy distribution of the illumination and of the surface spectral reflectance of the object (Figure 1.12 (a) and (b)). Conventional colour constancy

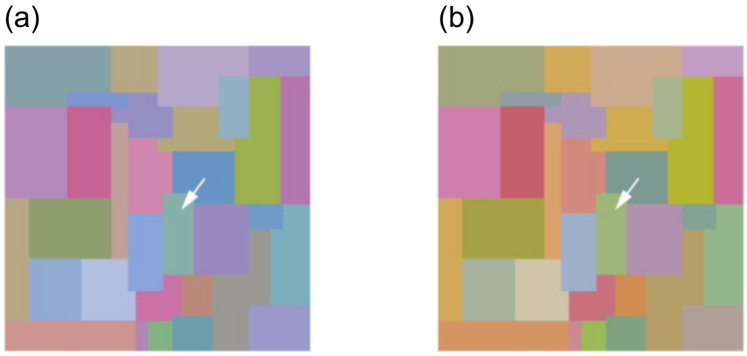


Figure 1.10: An example of the simplified "Mondrian" stimuli typically used in colour constancy research [2]. In the left panel, the flat surfaces are uniformly illuminated by a bluish illuminant (colour temperature, 10000K). In the right panel, the same surfaces are illuminated by an orangish illuminant (colour temperature, 4000K). The white arrows indicate a target patch and a test patch which observers were able to adjust its chromaticity and luminance.

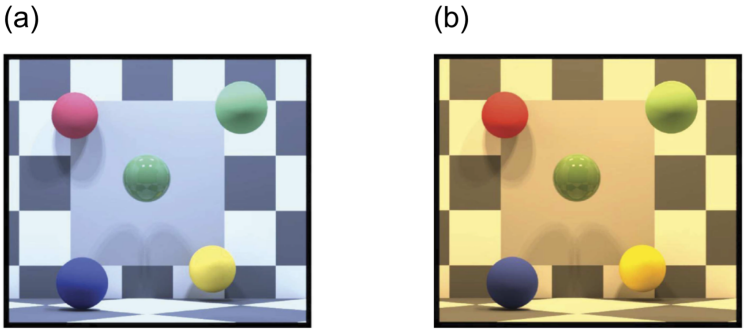


Figure 1.11: An example of a study that employed a computer-rendered 3-D scene [78]. In panel (a), the objects are three-dimensionally illuminated by a high-colour-temperature illuminant (15000K). In the panel (b), the same objects are illuminated by a low-colour-temperature illuminant (3500K).

algorithms, such as ones discussed above, consider only this diffuse reflection. By contrast, glossy objects additionally exhibit a specular reflection that carries direct information about the illuminant spectrum, and can therefore provide information about the lights incident on the object's surface (Figure 1.12 (c) and (d)). Thus, colour constancy for glossy objects may be supported by additional mechanisms that are not available from simplified stimuli.

In the following simulations we use computer graphics to see how chromaticity and luminance change as we add specular reflection and three-dimensional geometry to stimuli. To highlights the effect of changing surface properties, here we use point light sources which would be the simplest example of illumination that emits light uniformly in every direction.

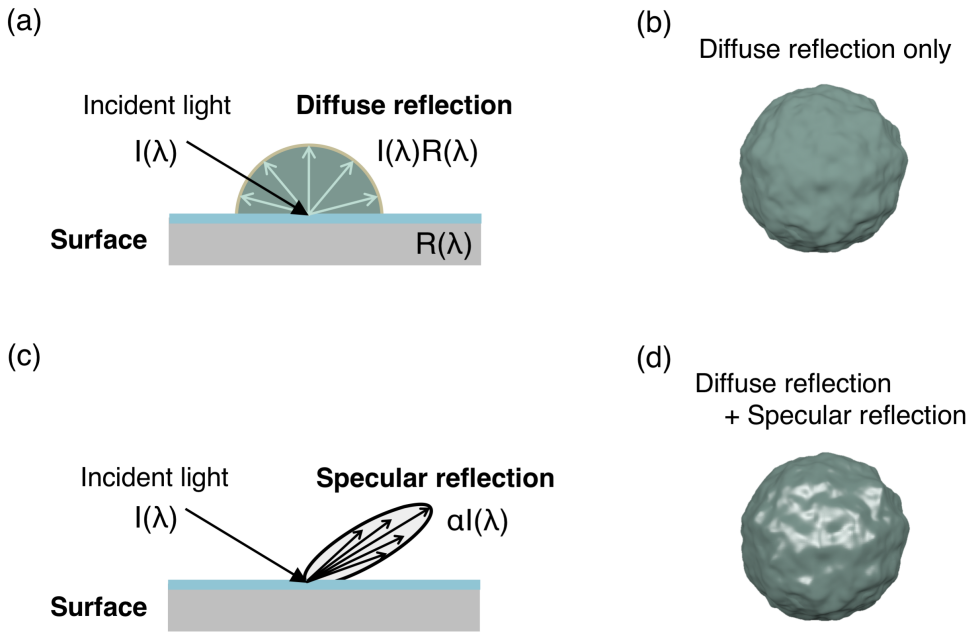


Figure 1.12: Two types of reflection (diffuse and specular) that occur when light hits an object. (a) Diffuse reflection whose spectral content is the product of the spectral composition of the illuminant and the spectral reflectance of the surface. (b) An example of an object that only exhibits diffuse reflection. (c) Specular reflection that is a direct reflection of the incident light. Thus, its spectral content matches that of the illuminant. (d) An example of an object that exhibits both diffuse reflection and specular reflection.

To render an object in a scene, we need to take several steps. First, we define the scene geometry (locations of a viewpoint or camera, an object and an illumination) using modeling software, Blender 2.78 (Blender Foundation). The scene contains only a single surface illuminated by three point sources of light, each with the same spectral content. Figure 1.13 shows the spectral energy distribution of two possible illuminants ((a) Sunlight $I_1(\lambda)$ and (b) Skylight $I_2(\lambda)$) and (c) the surface spectral reflectance that we assign to objects. We used two types of shape: smooth and bumpy objects.

According to the Ward reflectance model [142], for glossy objects, the light reflected back from a point on the object's surface can be expressed as a weighted sum of two types of reflection: a diffuse reflection in which the illuminant spectrum is modified by the surface reflectance, and a specular reflection in which (for non-metallic surfaces) the scene illuminant is returned spectrally unaltered.

Thus, at any point on the object's surface, incident light to our eye $E(\lambda,x,y)$, reflected from the point (x,y) is given by equation 1.9 and 1.10 where γ denotes the normalisation

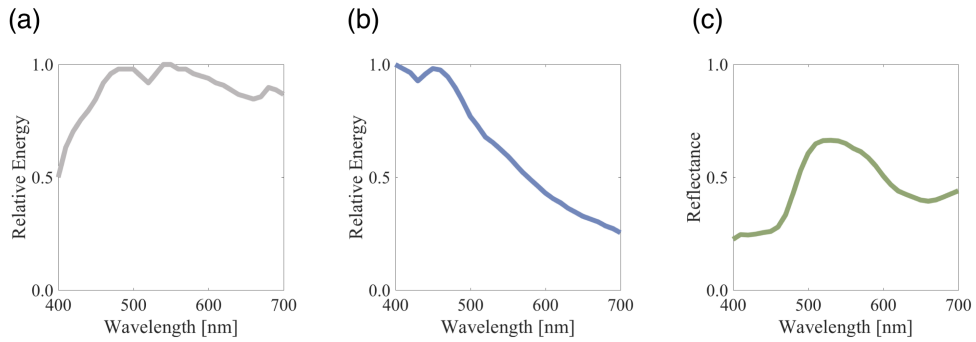


Figure 1.13: The spectral composition of (a) sunlight and (b) skylight. (c) The surface spectral reflectance that was applied to the object.

factor to preserve the total energy of light I(λ).

$$E(\lambda, x, y) = \gamma\{R(\lambda)I(\lambda, x, y) + \alpha I(\lambda, x, y)\} \tag{1.9}$$

$$\gamma = \frac{\int_\lambda I(\lambda, x, y)d\lambda}{\int_\lambda \{R(\lambda)I(\lambda, x, y) + \alpha I(\lambda, x, y)\}d\lambda} \tag{1.10}$$

Here, the independent variable λ refers to the wavelength, and x and y are the coordinate on the object's surface projected onto the image plane. R(λ) indicates the spectral reflectance of the object and I(λ,x,y) corresponds to an illuminant. The coefficient α signifies the level of the specular reflection. Thus, the first term R(λ)I(λ,x,y) corresponds to the diffuse reflection and the second term αI(λ,x,y) indicates the specular reflection. γ is merely a normalisation factor to preserve the total amount of energy in I(λ) as shown in equation 1.10.

Figure 1.14 shows the rendered images (the left column) and the 3-D and 2-D colour distributions of the objects (the centre and the right columns). First, for the sake of comparison, panel (a) shows a two-dimensional surface. The surfaces A and B are illuminated by sunlight and skylight, respectively. Importantly, each of these surfaces has only one colour and no variation over the surface in either chromaticity or luminance. These are indeed the type of stimuli that most past studies have employed.

Panels (b) and (c) allow us to see the effects of the specular reflection (across panels) and also the effects of changing illuminant colour (within each panel). In the chromatic distributions (right column), we also mark the chromaticities of the diffuse components

by plus symbols and of the specular components by cross symbols. The black dashed line shows the black body locus.

When an object has a three-dimensional geometry as shown in (b), the object's surface has luminance variation (the centre column) over the surface that is dependent on the angle of the surface and the illuminants' position. Importantly, however, the surfaces still have a single chromaticity which corresponds to the chromaticity of the diffuse component $R(\lambda)I_1(\lambda)$ or $R(\lambda)I_2(\lambda)$ (the right column).

Panel (c) shows that adding glossiness to the objects dramatically changes their colour distribution due to the contribution from the specular reflection. Their colour distribution (the centre panel) extends from the diffuse colour towards the colour of the specular highlight. We also see that the chromatic distribution (the right column) becomes a straight line which falls between the chromaticity of the diffuse component $R(\lambda)I_1(\lambda)$ or $R(\lambda)I_2(\lambda)$ and the chromaticity of the specular component $I_1(\lambda)$ or $I_2(\lambda)$. The chromaticity gets close to the illuminant chromaticity at points on the object's surface where the specular component is strong (corresponding perceptually to the shiny bits on the object's surface). However, note that in this example the lines do not fully reach the chromaticity of the illuminants themselves as the specular reflection is always mixed with the diffuse reflection unless the object has zero diffuse reflectance across all wavelengths (black object). Nevertheless, this figure demonstrates that specular reflections could offer a potential cue for our visual system to access the illuminant colour.

These are the cases for a smooth object, but there are a variety of shapes in the real world. To generalise this observation, Figure 1.15 shows the colour distribution for a bumpy surface. Interestingly, the difference in shape has little effect on the colour distribution although the appearances of the objects are very different. Therefore, from computational point of view, objects seem to contain similar amount of information about the illumination regardless of their shapes.

Simplification of illuminant properties

Another common simplification in past studies is that experimental stimuli were uniformly illuminated by single light source. Under this single-spectrum environment, the problem

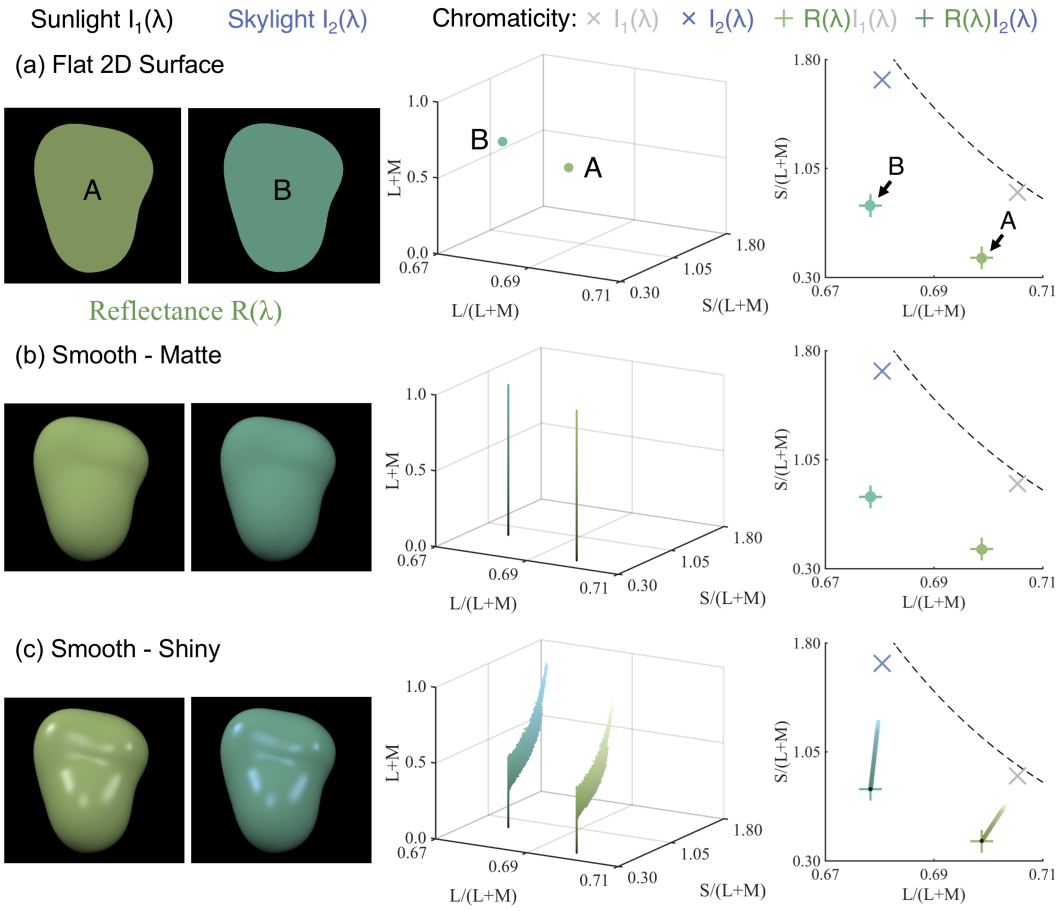


Figure 1.14: Colour distribution of rendered objects under point-light sources with one of two spectral distributions. The left column shows the rendered objects under point light sources (left, sunlight; right, skylight). The centre panel is the three-dimensional colour distribution of the objects. The right panel shows the two-dimensional chromatic distribution in MacLeod-Boynton chromaticity diagram. (a) A flat two-dimensional surface similar to those employed in past colour vision studies for the sake of comparison. The surfaces are uniform, and thus they contain only single chromaticity and luminance values. (b) A smooth matte object which comprises only a diffuse component. It has only a single chromaticity, but it exhibits a luminance variation. (c) A smooth shiny object which comprises both a diffuse and a specular component. Due to the existence of a specular reflection, colours distribute from the colour of the diffuse component to the colour of specular component (illuminant chromaticity).

of colour constancy seems simple. That is, once you manage to figure out the colour of the illumination, colour constancy can be implemented by subtracting its influence from the whole visual field under the restrictions of the "RGB heuristics". However, in more realistic situations, lighting environments can be much more complex. For example, it is possible that the scene contains multiple light sources. Even if not, lights emitted from a single light source will be reflected from other objects within a scene, and

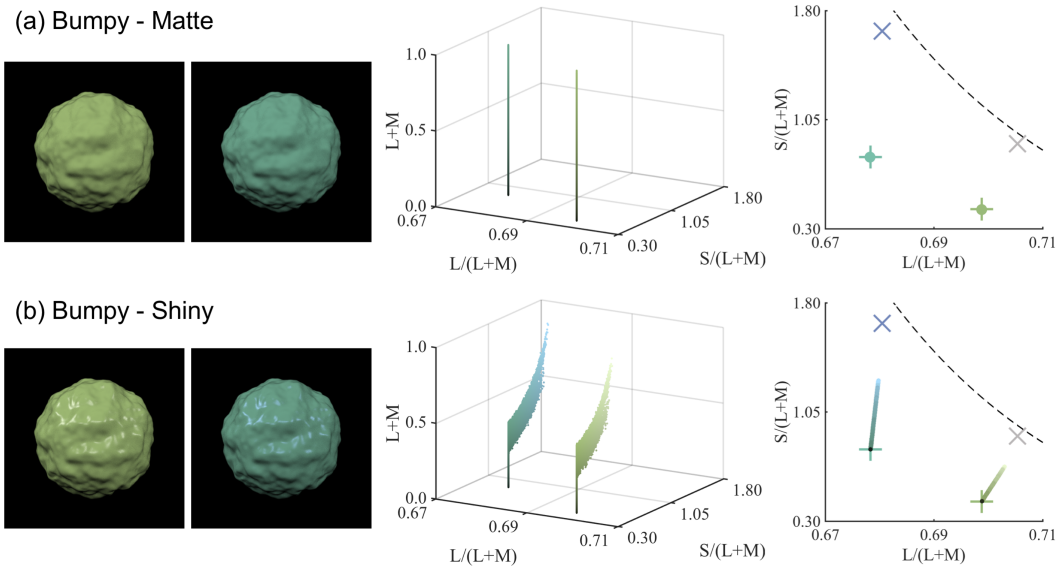


Figure 1.15: Colour distribution of rendered objects under point-light sources with one of two spectral distributions for bumpy objects. (d) A bumpy matte object that comprises only a diffuse component. (e) A bumpy shiny object that comprises both a diffuse and a specular component.

these mutual reflections will hit objects from different angles. In this sense, objects are illuminated from every direction, and the spectral content may vary from one direction to another. For complete colour constancy, position-dependent illuminant estimation should be necessary, which adds computational demands.

This notion of directional-dependent illumination has been used in computer graphics. The technique uses an environmental illumination map, which at each pixel stores incident light from a particular direction towards a single point within the scene. Thus, using the illumination map, it is possible to simulate the appearance of the object as if it were placed in the lighting environment.

Figure 1.16 shows an example of an environmental illumination map (left side, downloaded from http://dativ.at/lightprobes/) and the chromatic distribution of all pixels (right side). The chromaticities distribute broadly along the black body locus shown by the dotted black line. However, their colours also deviate from the locus. Note that at each pixel, values are specified by the three RGB channels rather than spectra, and thus we used a generic RGB to LMS conversion to calculate the chromaticity of each pixel which will be detailed later in this chapter (1.6 Calibration of CRT monitor).

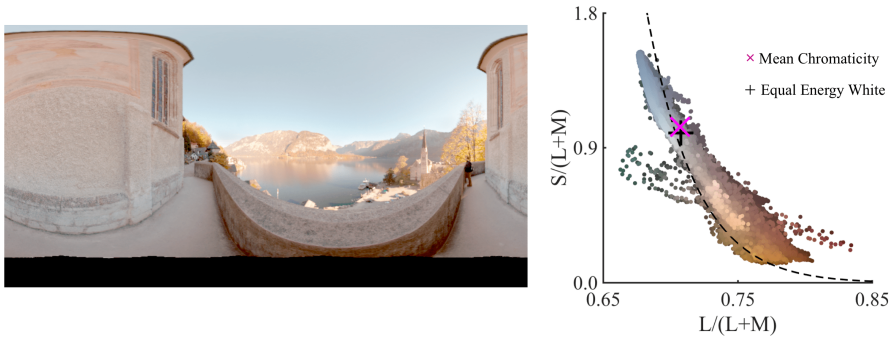


Figure 1.16: An RGB image of an environmental illumination map (left-hand panel) and its chromatic distribution (right-hand panel). The magenta cross indicates the mean chromaticity of the distribution. The black plus symbol corresponds to the chromaticity of equal energy white. The black dashed line indicates the black body locus.

Figure 1.17 shows rendered images (left column) and 3-D colour distributions (centre column) and 2-D chromatic distribution (right column) of the rendered objects placed in environmental illumination, which introduces complex chromatic variation across the object's surface.

Panel (a) shows the case for a matte object. Unlike the situation under point light sources where a matte surface has only a single chromaticity, chromaticity and luminance change dramatically over the surface. Note that this intensity variation depends on the angle of the surface. When the object becomes glossy as shown in panel (b), colours spread even more widely. The colours on the object's surface seem to preserve properties of the chromatic distribution of the environmental illumination itself, where colour spreads along black body locus (the right panel in Figure 1.16), but the whole distribution shifts towards the diffuse chromaticity shown in panel (a). Also, the object has a clear spatial reflection pattern of the surrounding environment which might allow us to understand the environment's geometry. The effect of shape (Figure 1.18) again seems to be small in terms of colour distribution, but importantly the geometry of the specular highlights is now distorted on the object's surface.

Now, what are the consequences of the presence of these chromatic and luminance variations over the surfaces introduced by environmental illumination in understanding human colour perception? When the surface contains only a single colour (as in the majority of past colour constancy studies), it makes sense to ask participants to judge

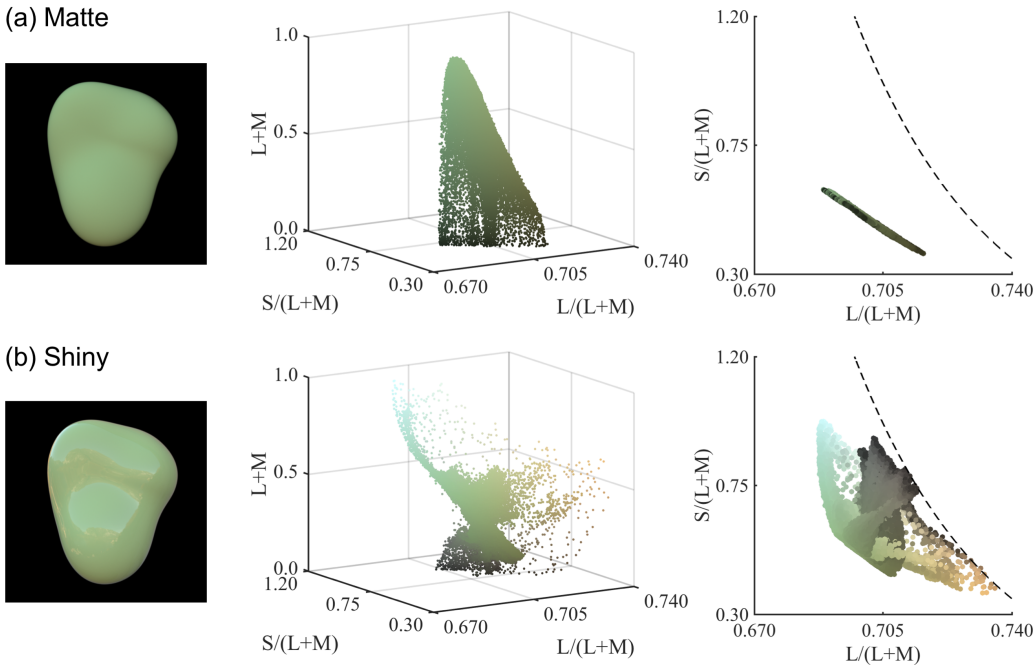


Figure 1.17: The effects of environmental illumination on the colour distributions of rendered objects. The left column shows the rendered objects under environmental illumination. The centre panel is the three-dimensional colour distribution of the object. The right panel shows the two-dimensional chromatic distribution. (a) A smooth matte object which comprises only a diffuse component contains a variety of chromaticities (unlike under single-spectrum illumination) due to the chromatic variation in environmental illumination as well as luminance variation across the surface. (b) A smooth shiny object which comprises both a diffuse and a specular component. The colours distribute much more widely compared to the matte surface.

the surface colour of an object, or the colour of illumination. However, since the light reflected back from the object now contains chromatic variation, it is possible for observers to interpret the surface as multi-coloured as opposed to uniformly-coloured. In the same way, the estimation of the "colour of illumination" might not make sense as there are millions of illuminations included in the environmental illumination maps. These possibilities make it hard to employ a classic single-colour judgement, such as achromatic setting or asymmetric matching, which typically asks participants to judge only a single colour of surface or illumination. To get around these matters, we never asked observers about the perceived colour of either the illumination or a surface, and instead used performance-based colour constancy measurements. Specifically, we employed discrimination of the surface-colour change versus illuminant-colour change (chapter 2), discrimination of surface colour of objects (chapter 3) and an object-selection task

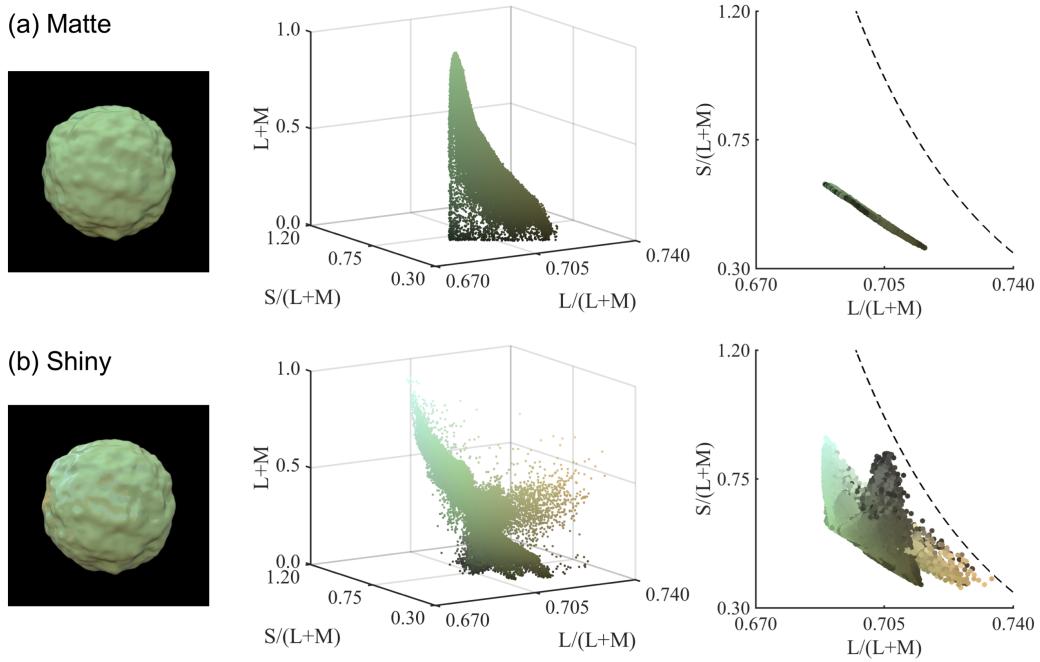


Figure 1.18: The effects of environmental illumination on the colour distribution of rendered object. (a) A bumpy matte object that comprises a diffuse component. (b) A bumpy shiny object composed of both a diffuse and a specular component.

across different lighting environments (chapter 4).

Moreover, the extent to which past colour constancy frameworks such as a diagonal model can be applied to complex stimuli is an open question. A recent study showed that when an object exhibits non-uniform colour over the surface, observers make an asymmetric match differently depending on the location of the test patch [143]. This result challenges the idea that a framework whose parameter is fixed at certain values by global image statistics can reasonably predict human observer's perception of surface colour. In chapter 4, we investigated to what extent models based on past single-point estimation of the illuminant can predict human observers' behaviours.

1.6 Calibration of a CRT monitor

This section describes the procedure of calibration conducted for the experimental CRT monitor (NEC, FP2141SB, 21 inches, 1600 × 1200 pixels) used in the experiments in Chapters 2, 3 and 4. The purpose of the calibration was to present experimental stimuli so that they have the desired chromaticity and luminance values. More specifically, to

present a colour which has specific L/(L+M), S/(L+M), and luminance [cd/m^2], what RGB values do we need to input to the monitor?

This requires several steps. First of all, we fixed the monitor position and adjusted the monitor brightness setting to the minimum value (zero) to make the black level as low as possible and adjusted the contrast setting to maximum value (100%). Then, we measured the spectrum of each phosphor (R, G, and B) with a SpectroCAL MKII spectroradiometer (Cambridge Research Systems). Based on these measurements, we calculated cone excitations from each phosphor using Stockman and Sharpe cone fundamentals [18] and created a conversion matrix between LMS and RGB coordinates. Let [L_R,M_R,S_R], [L_G,M_G,S_G], [L_B,M_B,S_B] be the cone responses to red, green and blue phosphors, respectively. The matrix transformation from RGB coordinates to LMS coordinates is given by equation 1.11.

$$C_1 = \begin{bmatrix} L_R & L_G & L_B \\ M_R & M_G & M_B \\ S_R & S_G & S_B \end{bmatrix} \tag{1.11}$$

To convert from LMS to RGB, the required transform is given by taking the inverse of C_1.

$$C_2 = C_1^{-1} \tag{1.12}$$

Figure 1.19 (a) shows their spectral shapes and luminance values. The panel (b) indicates the monitor gamut, which we see does not fully cover the chromaticity diagram.

Using the matrix C_2 given in equation 1.12, we obtained required R, G, and B values from desired L, M and S values. At this point, it is assumed that RGB values are linearly related to the monitor output. However, this does not hold for most monitors including our experimental monitor. Instead, the luminance value increases exponentially as the input value increases, and this non-linear property of the monitor needs to be corrected for. This procedure is called gamma correction. For this, we measured the relationship between monitor input and monitor output using a ColourCAL MK II colourimeter (Cambridge Research Systems) via software called ViSaGe Desktop (Cambridge Research Systems). Figure 1.19 (c) shows the inverse of the measured function which can be

used to convert the desired output to the required input values. This conversion was performed separately for each of R, G, and B channels. In experiments, input signals were sent to the CRT monitor via the ViSaGe MkII (Cambridge Research Systems), which allows 14-bit intensity (from 0 to 16383) resolution for each phosphor. We conducted this calibration procedure immediately before each experiment to minimise the effect of spectral and intensity changes over time.

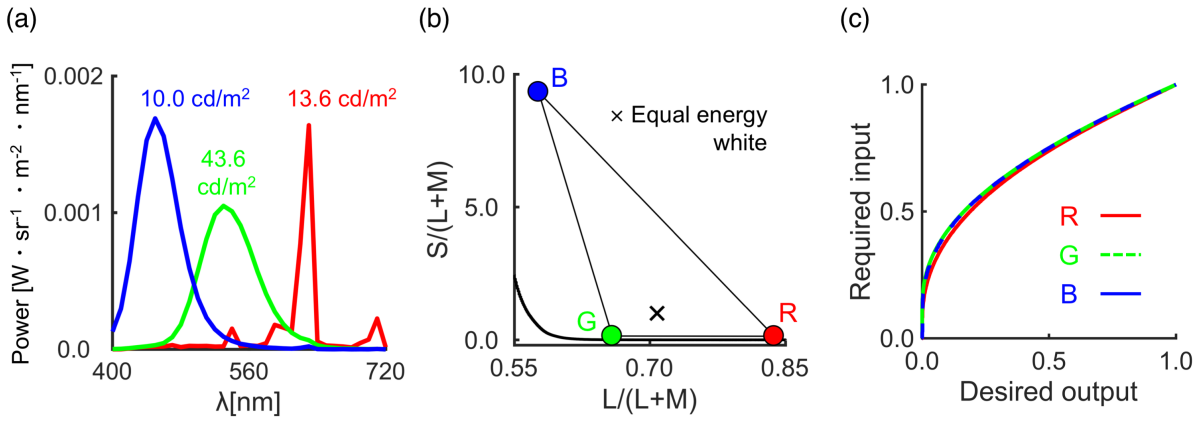


Figure 1.19: Measurements of the CRT monitor we used for experiments. (a) Spectrum of each RGB phosphor. Coloured numbers indicate luminance values for corresponding spectra. (b) Gamut of the CRT display used for experiments in the MB chromaticity diagram. (c) Gamma correction function for each phosphor. Note that we recalibrated the monitor before each experiment. Figures shown here are for the monitor condition immediately before we conducted the experiment in Chapter 4.

1.7 Overview of each chapter

Until this point, we have discussed (i) the basis of human colour vision, (ii) the nature of colour constancy, (iii) past colour constancy algorithms, (iv) methodology to quantify colour constancy and colour constancy indices, (v) limitations in past studies, (vi) the change in chromatic signals that complexity of surface and illuminant properties introduces, (vii) the calibration of the CRT monitor.

In this section, we provide a brief overview of each chapter. This thesis comprises an introduction (current chapter), five main chapters (chapters 2 to 6), a general discussion and a conclusion (chapter 7).

Chapter 2 - Experiment 1

Experiment 1 introduces the complexity of surface properties to experimental stimuli, with a particular emphasis on the interaction between surface pattern and the specular reflection. One of the widely-known algorithms in past studies is a "brightest is white" heuristic where the brightest surface in a scene is treated as a white surface [52], and therefore assumed to provide a direct cue to the illuminant colour. However, we showed that in theory when an object's surface has a high contrast pattern and specular reflection, this assumption does not always hold. We further experimentally showed, using a discrimination task between a material colour change and an illuminant colour change, that observers' colour constancy performance does not decrease even when the brightest element of the images are not the best predictor of illuminant colour. We also showed that observers' performance significantly dropped when the stimuli were phase-scrambled. This observation suggested that it is the geometry of specular highlights that guides observers to learn the influence of illumination. This finding emphasised the importance of using a complex stimulus, and overwrote one long-standing assertion in colour constancy that human observers use the bright element to estimate illuminant chromaticity.

Associated conference presentation

Takuma Morimoto, Robert J. Lee and Hannah E. Smithson, "When the brightest is not the best: Illuminant estimation based on highlight geometry", Vision Science Society, Tampa (May, 2017) <Poster> *Elsevier/Vision Research Travel Award

Chapter 3 - Experiment 2

In chapter 3, we extended the classic chromatic discrimination experiment to glossy objects under complex lighting environments. For each trial, four objects were presented simultaneously. Three distractor objects shared a surface colour while the fourth target object had a different surface colour. The four objects were always presented under the same environmental illumination. Results suggested that our ability to discriminate surface colour strongly depends on the chromatic distribution of environmental illumination. For the environmental illumination we used, colour was distributed along the daylight locus. Thus, observers performance was worse when two target objects have

surface colours that fall on the daylight locus. Moreover, even if we manipulated the axis along which colours distributed so that it does not follow statistical regularities in natural environment, some decrease of discrimination ability was observed along the direction of greatest variance within the manipulated chromatic distribution. In addition, overall performance dropped under this atypical illumination suggesting the special status of natural lighting environments.

Associated publication

Takuma Morimoto, and Hannah E. Smithson, "Discrimination of spectral reflectance under complex environmental illumination," Journal of the Optical Society of America A, 35, 4, B244-B255 (2018)

Associated conference presentation

Takuma Morimoto and Hannah E. Smithson, "Reflectance discrimination under complex environmental illumination", International Colour Vision Society, Erlangen, Germany (August, 2017) <Poster> *ICVS Student Travel Award

Chapter 4 - Experiment 3

In the previous experiment, we tested observers' abilities to discriminate surface colours within the same environment. In Experiment 3, however, we tested the ability of human observers to identify surface colour across different environmental illuminations. We again presented four objects simultaneously and only one object had a different surface colour from the other three objects. However, the left two objects and the right two objects were placed under different lighting environments. The observer's task was also the same, to find the object that had a different surface colour, but, in order to answer the question, observers had to remove the influence of the different lighting environments. Based on the results from computational modeling of observers' behaviour, it was shown that rather than trying to estimate the influence of illumination from a single trial, spatial and temporal averaging of surrounding colours over trials predicted human observers' performance better. These results indicate that the visual system exploits the history of chromatic signals in order to learn the representative chromatic statistics of complex environmental illumination.

Associated conference presentation

Takuma Morimoto and Hannah E. Smithson, "Identifying surface colour across different lighting environments", European Conference of Visual Perception, Trieste, Italy (August, 2018) <Poster> *WD Wright Award

Chapter 5 - Field work

To date, online databases of environmental illumination are limited to RGB channels, meaning that spectral information is lost at each pixel. In this chapter, we report hyperspectral lightprobe datasets that we measured using a mirror sphere and a hyperspectral imaging system to characterise the directional spectral variation of natural environments. All datasets were collected in collaboration with Prof. Sergio Nascimento at the University of Minho, Portugal. The results showed that there is significant spectral variation depending on incident angle. For outdoor scenes, lights coming from above tend to be bluish and to have high luminance, whereas lights from below have one log unit lower luminance and their colour temperature are considerably lower. For indoor scenes, the overall light level was one log unit darker than outdoor scenes, and directional variation was much smaller compared to outdoor scenes. Finally, we argue that the use of daylight model is limited for some incident angles even in sunny outdoor scenes because of the presence of mutual reflection.

Associated publication

Takuma Morimoto, Sho Kishigami, João M.M. Linhares, Sérgio M.C. Nascimento, and Hannah E. Smithson, "Hyperspectral environmental illumination maps: Characterizing directional spectral variation in natural environments," Optics Express, 27, 22, 32277-32293 (2019) *Selected for the Editor's pick

Associated conference presentation

Takuma Morimoto, Sho Kishigami, João M.M. Linhares, Sérgio M.C. Nascimento, and Hannah E. Smithson, "Hyperspectral environmental illumination: directional spectral variation and a revisit of metamerism in the real world", International Colour Vision Society, Riga, Latvia (July 2019) <Talk> *Best Presentation Award, and ICVS Student Travel Award

Chapter 6 - Theoretical work

Human colour vision is supported by three cone classes, namely L, M and S cones. When the light hits these cones, all wavelength information is lost (principle of univariance). Therefore, two surfaces that have distinct surface spectral reflectances might appear the same under a specific illumination and this problem is known as metamerism. However, newly collected hyperspectral environmental illumination maps suggested that there is a striking directional spectral variation in natural environments. This means that depending on the tilted angle of a surface, the surface can sample different illuminations. Therefore, even when metamerism happens for one specific direction, we might be able to break the metamerism by simply tilting the surface. We performed a theoretical analysis to characterise the proportion of the metamerism that occurs under a single-spectrum environment and can be broken by utilising directional variation in natural environments. The results suggested that roughly 50% of metamerism can be solved. We also found that this observation holds for dichromats who are missing one class of cones. Data obtained from novel hyperspectral measurements of environmental illumination suggested that spectral variation in the natural environment might compensate our biological hard limit in human colour vision.

Associated conference presentation

Takuma Morimoto, João M.M. Linhares, Sérgio M.C. Nascimento, and Hannah E. Smithson, "Breaking illuminant metamerism using directional spectral variation in natural environments: dichromats might benefit more than trichromats", Optical Society of America Fall Vision Meeting, Washington D.C., U.S.A (September, 2019) <Talk>

Chapter 7 - Summary and Conclusion

Chapter 7 summarises major results from the five main chapters. We also discuss limitations of the present studies and highlight possible future directions.

2

When the brightest is NOT the best: illuminant estimation based on highlight geometry

2.1 Introduction

Colour constancy requires discounting the influence of illumination from the reflected lights reaching our eyes. However, estimating the colour of the illumination on a scene is difficult because the light reflected from an object depends both on the spectral composition of the illumination and the surface's spectral reflectance. The colour of the illumination can be in theory perfectly estimated if a scene contains a white surface with a spectrally flat surface reflectance because the surface returns the illumination spectrally unmodified. In the real-world situation, however, even if the scene contains a white surface, our visual system would not be able to know if it is white, simply because we cannot judge whether the light reflected back from the surface is spectrally modified or unmodified by the surface reflectance. Similarly, specular highlights also provide a direct reflection of the illumination, which therefore means that they are a potentially strong cue to the illumination colour, and several studies have shown the contribution of specular highlight to the colour constancy process [78], [79]. Our visual system creates a percept of glossiness from specular highlights, and the presence of this distinct perception seems to support

the notion that the visual system can identify regions of specular reflection on an object's surface. Conveniently, the specular highlight is often the brightest point in a scene. Thus, by assuming that the brightest element in a scene is either a white surface or a specular highlight, it would be possible to access the chromatic signal that is reasonably close to the illuminant colour. This heuristic is known as the "brightest is white" assumption [52], and it has been used to account for empirical data in early colour constancy studies which employed matte surfaces [54], [144]–[148]. However, we here show this principle does not always hold; for example, when a glossy object exhibits spatial intensity variation over its surface, the specular highlight may fall on a dark part of the surface, so that the centre of the specular highlight is not necessarily the brightest point. This counter example is depicted in Figure 2.1. In panel (a) the specular highlight falls on the bright part of the texture, and therefore the brightest point and the centre of the specular highlight are at the same position. However, in panel (b), the specular highlight falls on the dark part of the texture, and the brightest point does not align with the centre of specular highlight. A more detailed explanation of why the "brightest is white" assumption is not the most informative is explained in the Supplementary information section.

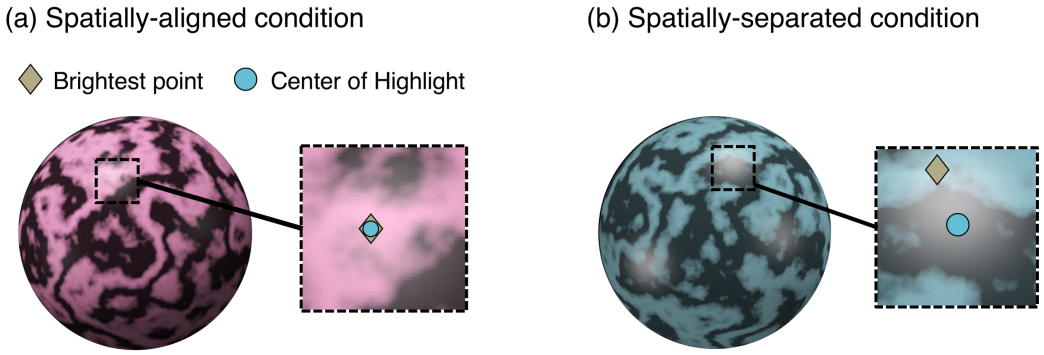


Figure 2.1: (a) The specular highlight and bright parts of the texture are spatially aligned. In this case, the brightest pixel (gold diamond) and the centre of the highlight (cyan circle) spatially overlap. (b) Specular highlight falls on a dark part of the surface. The brightest point and the centre of the highlight are spatially separated.

The present study directly tests whether the violation of the "brightest is white" heuristic impairs performance in a colour constancy task. To assess colour constancy performance, we measured discrimination between a material colour change and an

illuminant colour change which has been used previously [129], [149] to show a robust effect of specular highlights. The results in our experiment suggest that observer's performance did not worsen when the "brightest is white" assumption does not hold. However, when the spatial structure of the experimental stimuli was scrambled, observers performance significantly dropped. Thus, we suggest an alternative mechanism to infer the influence of illumination based on the geometry of the specular highlight.

2.2 Method

2.2.1 Apparatus

The experiment was conducted in a dark room. Stimuli were presented on a CRT monitor (53.3 centimetres, 1024 × 768 pixels, NEC, FP2141SB, Tokyo, Japan) controlled by a visual stimulus generator (ViSaGe MkII, Cambridge Research Systems, Rochester, UK), which allows 14-bit intensity resolution over full dynamic range for each phosphor. Gamma correction was performed by a colourimeter (ColorCAL MK II, Cambridge Research Systems, Rochester, UK) and spectral calibration was conducted with a spectroradiometer (SpectroCAL MKII, Cambridge Research Systems, Rochester, UK). Viewing distance was maintained by a chin rest positioned 92 cm from the CRT monitor.

2.2.2 Rendering

All stimuli were generated by computer graphics techniques. The geometry of each scene (locations of a viewpoint or camera, an object and three point-light sources) was defined using the 3-D modeling software, Blender 2.78b (Blender Foundation, Amsterdam). Then, rendering was conducted based on the physically-based renderer, PBRT (Physically Based Rendering Tool) [150]. The resultant multispectral images (31 channels, from 400 nm to 700 nm with 10 nm steps) were converted to LMS cone coordinates based on the Stockman & Sharpe 2° cone fundamentals [18] and then finally converted to RGB values for the presentation. We used MATLAB 2016b (MATLAB 2016b, Mathworks, Massachusetts, USA) and Rendertoolbox [151] to automate the stimulus production process.

2.2.3 Stimuli

We used smooth spheres of radius 10.0 degrees of visual angle that had a constant roughness of 0.02 defined in PBRT. Each scene contained only a single sphere illuminated by three point sources of light, each with the same spectral content. The positions of the three point sources of light were chosen for each trial from 10 possible patterns so that observers could not simply remember the locations of the specular highlights. The sphere had a single spectral surface reflectance, but across the sphere, the surface reflectance was attenuated by a wavelength-independent scalar variable α in a location-specific manner, giving rise to a spatial texture. For test illuminants, we used sunlight and skylight. The spectral composition is shown in Figure 2.2.

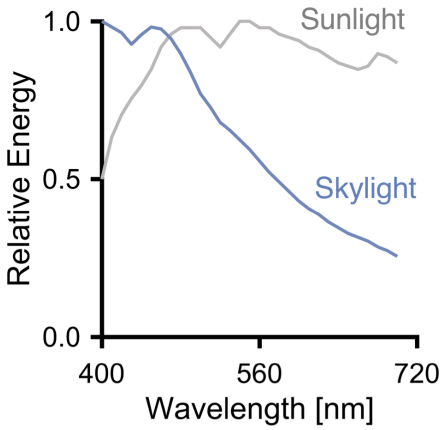


Figure 2.2: The spectral composition of two test illuminants: sunlight and skylight.

According to the Ward reflectance model [142], incident light to our eye, E(λ), reflected from a certain point on an object's surface is given by the sum of the diffuse and specular components. Thus, the spatial distribution of reflected light from our experimental stimuli can be expressed by equation 2.1.

$$E(\lambda, x, y) = \alpha(x, y)R(\lambda)I(\lambda) + \beta(x, y)I(\lambda) \tag{2.1}$$

Here, $\alpha(x,y)$ and $\beta(x,y)$ indicate coefficients due to scalar variation in the object's surface reflectance (achromatic texture) and the magnitude of the specular component, respectively. R(λ) is the surface spectral reflectance of the sphere, and I(λ) is the

illuminant spectral composition (sunlight or skylight). Thus, the first term in equation 2.1 corresponds to the diffuse reflection and the second term is the specular reflection. Note that in addition to these factors, the intensity of the diffuse component is modified by a geometrical factors of the scene (i.e. the position of the camera, and the shape of the object). The scalar values $\alpha(x,y)$ varied between 0.2 (dark) and 1.0 (bright), and thus even in the dark part of the texture, the specular component and the diffuse component were always mixed, meaning that no direct information about the illuminant colour was presented in any stimulus. The surface had one of three levels of specularity: zero (matte), low (0.02 as defined in PBRT) and mid (0.04). Note that the specularity (a rendering parameter assigned to the surface) was constant over the whole surface, but the strength of specular reflection reaching the camera (i.e. in the image of the object) spatially varied due to the 3D geometry of the light source, surfaces and camera.

2.2.4 Experimental task and creation of the animation set

The observer's task was to discriminate between a reflectance change and an illuminant change. This type of task, often known as a "operational colour constancy" task, has been employed in several past studies with matte stimuli [77], [110], [126], [152].

In each trial, observers saw a 1.5-second animation, consisting of 10 frames. The colour of the sphere gradually changed, and observers were asked to indicate whether the colour change was due to a change of illuminant colour or due to a change of surface colour. Colour changes were defined either as transitions between pairs of spectral reflectance functions or between sunlight and skylight illuminations. Figure 2.3 indicates example trials for both conditions, describing the change of chromatic distribution during animation. To test the importance of the geometry of the specular highlight, we additionally used phase scrambled images that preserved the same chromatic and luminance distribution as the original image.

To quantify the observer's performance at discriminating a reflectance change from an illuminant change, we needed to prepare a sufficient set of animations for both scenarios. First, we chose two illuminations (sunlight and skylight) as shown in Figure 2.2. Second, we chose 240 reflectances of natural objects and man-made objects[153]–[155].

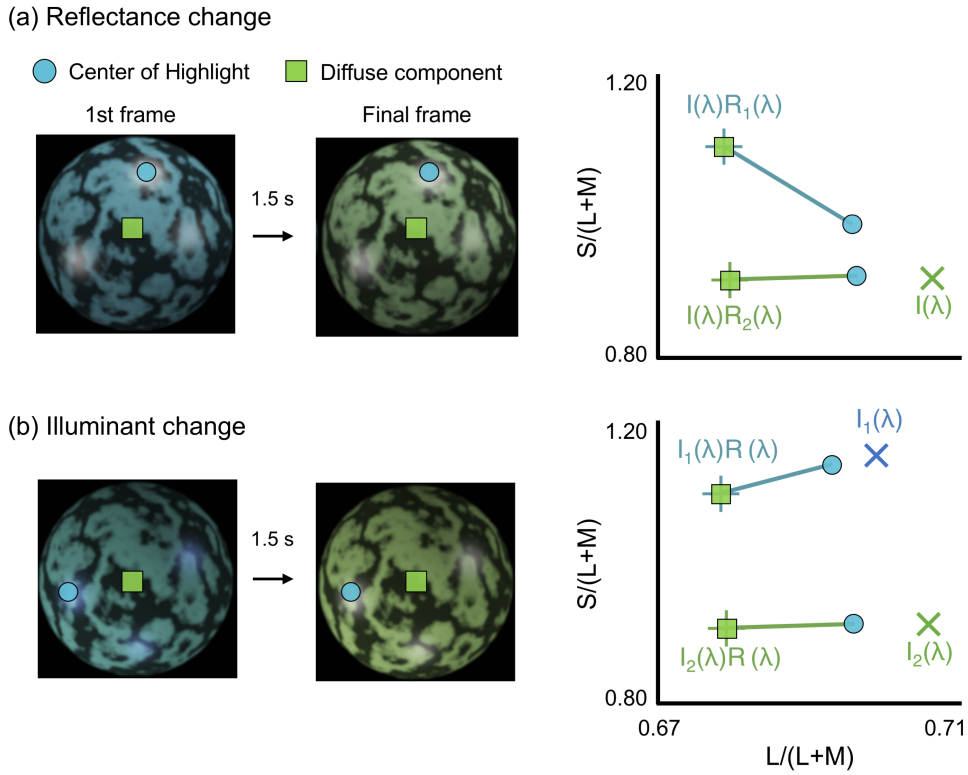


Figure 2.3: Example trials of (a) reflectance change and (b) illuminant change. The rightmost panels show the change in "locus of chromaticities" over surfaces (cyan and green lines) and chromaticities at key points (cyan circle and green square symbols) between the first and final frames. The cross and plus symbols indicate the chromaticity of illuminations and diffuse reflections, respectively.

Figure 2.4 shows the chromaticities of those surfaces under the two test illuminants. For illuminant change animations, the chromaticity of the object essentially changes along a line parallel to the line between the two illuminant chromaticities (blue and black crosses in Figure 2.4). Thus, if we randomly select reflectance pairs for the reflectance change animation, chromatic direction of change could be a sufficient cue for observers to make a judgement without relying on specular highlights. Thus, we selected 50 pairs of reflectances for the reflectance change animations and 50 reflectances for the illuminant change animations so that the colour change was predominantly in the yellow–blue or blue–yellow direction in both cases.

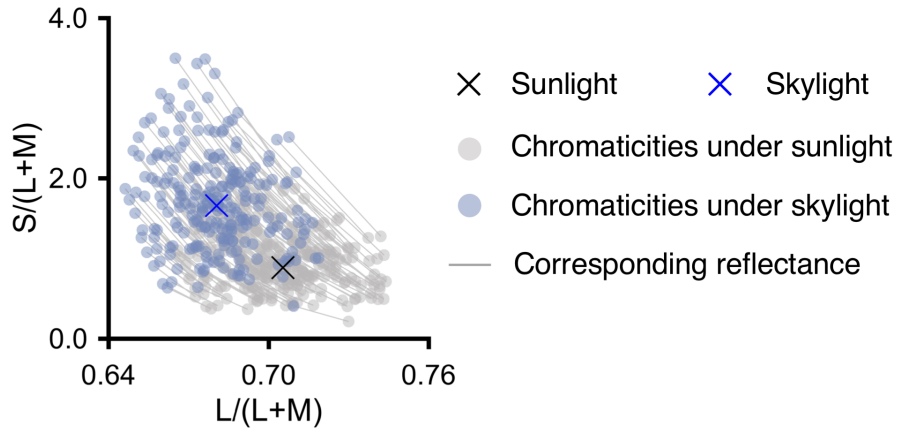


Figure 2.4: The chromaticities of 240 reflectance samples of natural objects and man-made objects when rendered under sunlight (gray circle) and skylight (blue circle). The black and blue cross symbols indicate the chromaticities of sunlight and skylight, respectively. The gray lines indicate how chromaticity of the same reflectance changes between sunlight and skylight.

2.2.5 Observers and Ethics approval

Five observers (AKH, TD, JH, SR, TM; TM was the author) completed all conditions. One additional observer (HH) completed the original sphere condition only. All observers had normal colour vision, as assessed by HRR pseudoisochromatic colour plates. HH was a naive observer. The study was approved in 2016 by the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford in agreement with the Declaration of Helsinki.

2.2.6 Procedure

Each trial consisted of a single animation and response. Feedback was provided immediately after the response. One session consisted of 500 trials (animations) composed of 100 trials (50 reflectance changes and 50 illuminant changes) for each of three specular levels (zero, 0.02, and 0.04) for the original sphere condition, and two specular levels (0.02, and 0.04) for the phase-scrambled conditions. Figure 2.5 provides a summary of conditions. For each condition, in half of trials, the centre of the specular highlight coincided with the brightest point on the object's surface (spatially-aligned). In the remaining trials, the specular highlight fell on the dark part of the texture, and thus the brightest point and the centre of the specular highlight were at different locations (spatially-separated).

There was a short break after every 125 trials. Observers completed four sessions in total (2000 trials). The order of the conditions was randomised for each observer.

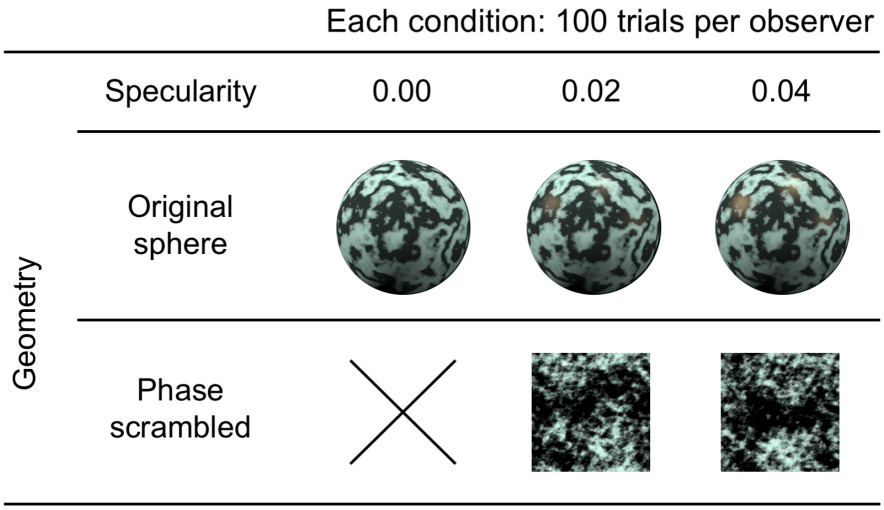


Figure 2.5: Summary of conditions. Note that specular level zero was not used for the phase scrambled condition.

2.2.7 Simulation of the "brightest is white" heuristic and other models based on chromatic statistics

In order to simulate an ideal observer who always bases her/his judgements about the material or illumination change on the chromatic change of the brightest pixel, we implemented a naive Bayes classifier based on the *fitcnb* function in the Statistics and Machine Learning toolbox in MATLAB. The basic concept of the model is to learn a probability density distribution as a function of the input parameter (i.e. chromatic change) separately for each classification category (i.e. material change and illuminant change). Once the model develops the probability function, the model can classify a new trial based on an observed value. We used only the first and final frames, because the chromatic change between those frames was the most informative. Note that chromatic change was calculated in a scaled MacLeod-Boynton (MB) chromaticity diagram, where equal energy white corresponds to the origin and each axis is independently scaled by chromatic discrimination thresholds along L/(L+M) and S/(L+M) measured prior to the experiment. First of all, in the training phase, the Bayes classifier saw all 300 trials of the

original sphere condition as seen by actual observers, and collated statistics separately for the reflectance change trials and the illuminant change trials based on the distributions of observed chromatic change of the brightest pixel. Second, in the actual simulation phase, the classifier went through all trials again, and for each trial it calculated the chromatic difference and decided whether it was more likely to belong to the reflectance change or the illuminant change distribution. No prior was given to the classifier before the training. Figure 2.6 provides a schematic illustration of how such modeling works.

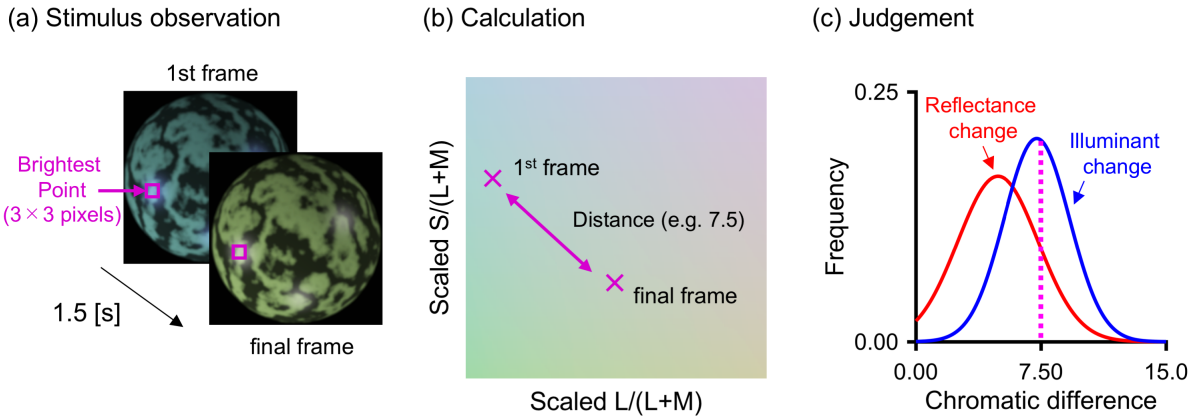


Figure 2.6: Schematic description of how the computational modeling works. (a) A model first observes an animation and identifies the spatial region of interest (the brightest point in this case). Next, (b) the chromatic difference between the 1st and final frames is calculated. (c) The model determines whether the observed chromatic difference belongs to reflectance change or illuminant change based on the learned distribution before the simulation. For this particular trial, illuminant change is selected as a response as it provides a higher probability.

We also ran simulations using the same type of Bayes model, modified to rely on other chromatic statistics: centre of specular highlight, mean chromaticity, and the luminance-weighted mean chromaticity. Regarding the model that relies on the centre of the highlight, there were three specular highlights in each animation. Thus, we constructed three separate models that focus on the most intense, the second most intense, or the third most intense specular highlight (1st, 2nd, and 3rd highlight, respectively). The mean chromaticity model simply calculates the mean chromaticity of the whole image. The weighted mean chromaticity model weights each pixel by its luminance raised to the power by of 1, 10 or 100. Note that when the exponent is infinity, it acts exactly the

same as the brightest model. In this sense, the luminance-weighted mean chromaticity is a model between the mean chromaticity (exponent zero) and the brightest models.

2.3 Results and Discussion

Figure 2.7 shows the results from human observers ((a) and (b)) and computational models ((c), (d), (e) and (f)). Panel (a) indicates the averaged discrimination index d' across all observers. The solid line indicates d' for spatially-aligned trials and the dashed line shows data for spatially-separated trials in which the specular highlight falls on dark regions of the texture. Panel (c) shows the prediction from a naive Bayes classifier that makes a discrimination based on the "brightest is white" heuristic. Panels (d) - (f) indicate models based on other chromatic statistics.

Panel (a) shows that, as predicted, at zero specularity, observers' performance was around chance level (d' = 0). However, d' increased when specularity increased to 0.02. There was also a difference between d' at low and mid specularities (0.02 and 0.04). Most importantly, when the specular highlight fell on a dark region of the texture, observers' performance improved, contrary to the "brightest is white" heuristic which predicts decline in performance (panel (c)). This interaction shows that the visual system does not merely pick the brightest element in a scene when making judgements about illumination. A major explanation of the improved performance is that the chromatic change of the specular highlight is bigger in illuminant change trial when the specular highlight is on the dark part of the texture because it is less contaminated by the diffuse component. In addition, a cognitive-level interpretation in the spatially-separated trials is that the high luminance contrast between the specular highlight and the surrounding dark region helped observers to detect the specular highlight. Importantly, the Bayes classifier based on the chromaticity at the centre of the specular highlights shown in panel (d) predicted the direction of difference between spatially-aligned and spatially-separated conditions, suggesting that human observers were relying on the specular highlight regardless of whether it was the brightest or not. The models based on mean chromaticity or luminance-weighted mean chromaticity do not seem to account for the pattern of human observers'

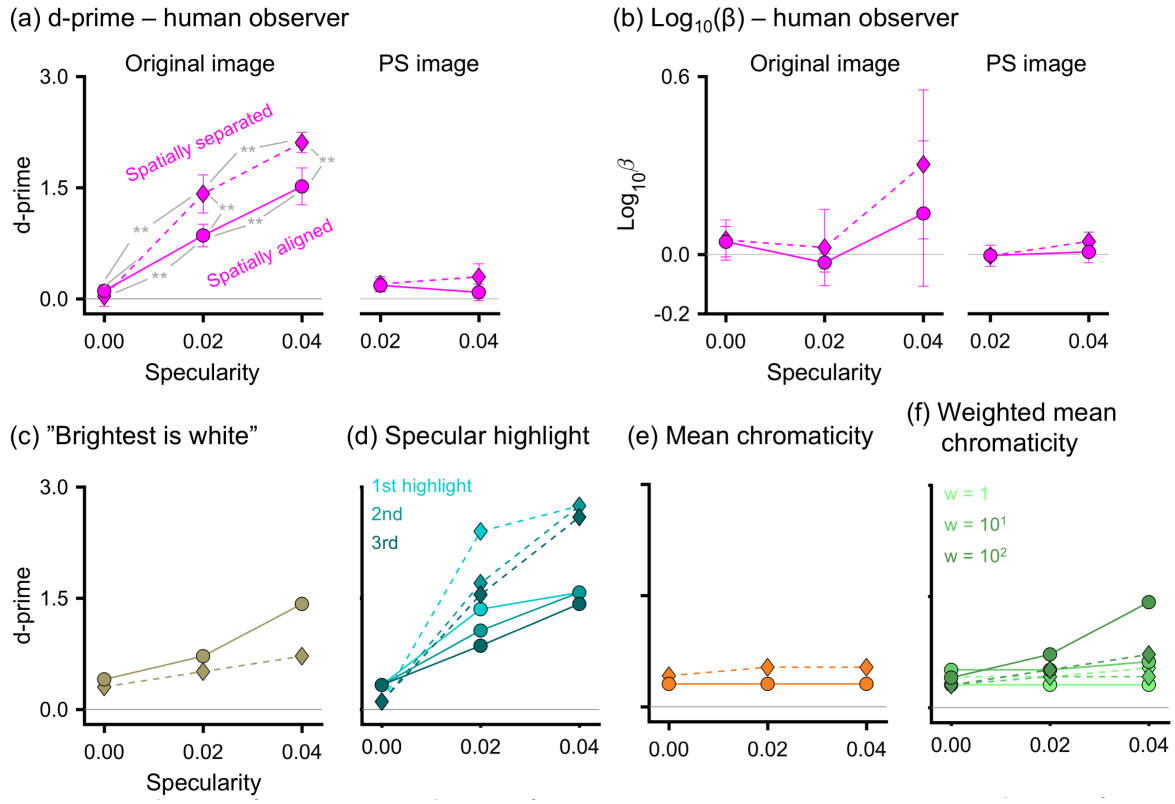


Figure 2.7: Average d' across all observers for (a) human observers. The original sphere conditions and the phase-scrambled (PS) conditions are shown side by side. The solid line shows the spatially aligned condition where specular highlight and the brightest point overlap, and the dotted line indicates the spatially separated condition where specular highlight fell on a dark part of the texture. The double stars indicate p-value less than 0.01. Panel (b) indicates the log beta values, indicating the bias towards illuminant change (positive value) and reflectance change (negative values). Error bars indicate ± 1 S.E. across observers. Panels (c) - (f) indicate the d-prime predictions from different models. In particular, the (c) "brightest is white" model predicted the difference between spatially separated and spatially aligned trials in the opposite way to that shown by human observers. In contrast, the model based on specular highlights shown in the panel (d) accounts for the pattern of human observers' performance.

performance, showing that a strategy based on global statistics is not a good descriptor. Another important result is that performance declined when the spatial structure of the specular highlight was not preserved (right panel in (a)). This suggests that the geometry of the specular highlight plays a crucial role in helping observers to discriminate the material change and the illuminant change. Panel (b) shows the log beta values to show the observers' response bias (as defined in signal detection theory), where positive and negative values indicates a response bias towards the illuminant change and the reflectance change, respectively. Note that this beta is different from one in equation (2.1). Here, we

see that the values are not much different from zero, meaning observers did not show a particular bias towards either category. For the original sphere condition (left panel in (a)), we conducted two-way repeated-measures analysis of variance (ANOVA) with spatial alignment (spatially aligned and spatially separated) and specularity (zero, 0.02 and 0.04) as within-subject factors for d'. We found a significant main effect of spatial alignment ($F(1,5) = 33.3$, $p = .00219$) and a significant main effect of specularity ($F(2,10) = 47.2$, $p < .00001$). In addition, there was a significant interaction between spatial alignment and specularity ($F(2,10) = 6.75$, $p = .0140$). Further analysis of the interaction showed that the main effect of spatial alignment at specularity zero was not significant ($F(1,15) = 0.30$, $p = .602$), whereas it was significant at specularity 0.02 and 0.04 ($F(1,15) = 17.8$, $p = .000752$; $F(1,15) = 19.6$, $p = .00049$, respectively). The main effect of specularity was significant for both spatially-separated and aligned trials ($F(2,20) = 51.7$, $p < .00001$; $F(2,20) = 23.056$, $p = .000109$, respectively). Finally, multiple comparisons based on the Ryan-Einot-Gabriel-Welch's correction showed that for the spatially-separated condition, the difference between zero and 0.02, between specularity zero and 0.04, and between specularity 0.02 and 0.04 were all significant ($t(20) = 6.654$,$p < 0.0001$; $t(20) = 9.984$, $p < 0.00001$; $t(20) = 3.33$, $p = 0.00333$, respectively). Also, for the spatially-aligned condition, the difference between specularity zero and 0.02, between zero and 0.04 and between 0.02 and 0.04 were all significant ($t(20) = 3.594$, $p = 0.00181$; $t(20) = 6.79$, $p < 0.00001$; $t(20) = 3.19$, $p = 0.00457$, respectively). For the phase scrambled image (right panel in (a)), we again conducted two-way repeated-measures ANOVAs, following the procedure outlined above, with the exception that specularity had only two levels (0.02 and 0.04). We found no significant main effect of spatial alignment ($F(1,4) = 0.893$, $p = 0.398$) nor of specularity ($F(1,4) = 0.000$, $p = 0.993$). In addition, there was no significant interaction between the two factors ($F(1,4) = 1.25$, $p = 0.326$).

The spatial summation area for the visual system's extraction of these statistics is likely to be more than one pixel in the display. Thus, when extracting the brightest pixel or the centre of the specular highlight, we averaged chromaticities across pixels with different spatial scales: one pixel, 3×3 pixels, 5×5 pixels, 10×10 pixels. However,

the spatial scale had little effect on the results, and thus we report only the results for the 3×3 pixel patches in Figure 2.7.

The present study showed that when the brightest pixel does not convey the best information about the illuminant colour, observers performance increased whereas the Bayes classifier that relies on the brightest region predicted a decrease in performance. The increase of observers performance can be explained by the fact that when the specular highlights falls on a dark part of the texture, highlights provided the better estimation of the colour of illumination as it is less contaminated by the diffuse component. This finding suggests that human observers did not perform the task by simply relying on the brightest point on the objects' surface. Instead, the Bayes classifier based on the centre of the specular highlights was able to predict the difference in observers' performance between the spatially-aligned and spatially-separated conditions. An additional key observation was found in the phase scrambled condition, where d' values for human observers substantially dropped, suggesting the importance of the geometry of specular highlights. Taking these results together, we suggest that human observers rely on the geometry of specular reflections to infer the colour of illumination, which can overwrite a widely recognised algorithm to estimate illuminant colour based on the brightest element in a scene.

Colour constancy has been a rich area of study in behavioural neuroscience. However, early studies extensively used oversimplified experimental stimuli where the object is matte and has a spatially uniform surface spectral reflectance over the surface (known as "Mondrian stimulus" [29]). This study employed surfaces that had a spatial reflectance variation, which allowed us to shed light on the observers' strategy to estimate the influence of illumination that does not merely rely on the "brightest is white" heuristic.

2.4 Conclusion

The present study showed that even when the "brightest is white" assumption is violated, observers' performances in classifying reflectance changes versus illumination changes do not decline, but can increase. Here, we suggest an alternative strategy to the "brightest

is white": Observers use the geometry of the surface and the illumination to identify highlight regions to estimate the illumination colour.

2.5 Supplementary Information

2.5.1 Rationale for why the "brightest is white" assumption fails

The specular component alone would provide a perfect sample of the illuminant chromaticity. However, equation 2.1 shows that, at any point in the image, the degree of dominance of the specular component is determined by the coefficients α and β. Figure 2.8 helps us to understand how specular dominance varies across the image.

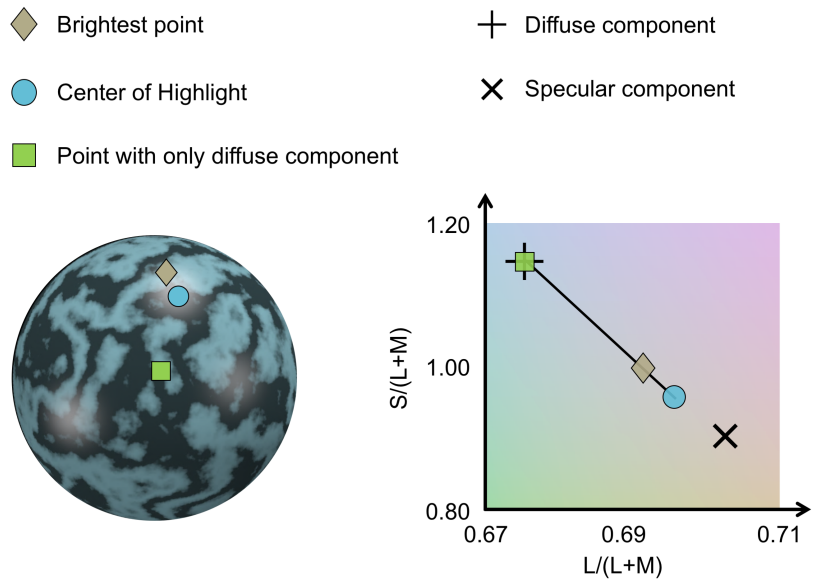


Figure 2.8: An illustration of how chromaticity distributes on an object with spatial intensity variation. See main text for details.

On the right-hand panel, the "locus of chromaticities" over of the whole object is represented by a black solid line in the MacLeod-Boynton chromaticity diagram [17]. The plus symbol is the chromaticity of the diffuse component. The black cross indicates the chromaticity of the specular component alone; thus, it is the chromaticity of the illumination that we would like to access to achieve colour constancy. Because, at any point, the chromaticity is determined simply by the balance between diffuse and specular components, the whole chromatic locus is defined by the straight line between the plus

and cross symbols. The coloured circles correspond to the chromaticities of particular points shown in the left-hand image. The reflected light from the position indicated by the blue circle provides particularly good information about the illuminant because α is small and β is maximal at the centre of the highlight. The point indicated with a brown diamond is less informative due to the higher α and smaller β. However, this is also the point selected by the "brightest element" heuristic, since the overall intensity at this location is higher than the intensity at the centre of highlight shown by the blue circle. The reflected light from the location indicated by the green square has the furthest chromaticity from the illuminant because it does not contain any specular component.

It is also worth emphasising that in this case the specular highlight provides particularly reliable information about the colour of illuminant because it lands on a dark part of the texture, and it is thus less contaminated by the diffuse component. Based on these rules, we predicted that observers would be more colour constant than would be predicted by the brightest element assumption, if they could estimate the geometry of the specular component and consistently extract an illuminant estimate from the centre of the specular highlight.

2.5.2 The nature of the experimental task

In order to understand why each model resulted in the performances shown in Figure 2.7, it is helpful to understand the nature of the task and how the models can solve it. Figure 2.3 in the main text shows how the chromaticities of key points in the image change during the 1.5-seconds animation for (a) a reflectance change and (b) an illuminant change for specific trials. For (a) a reflectance change, the chromatic distribution of the first frame and the final frame extends from their diffuse components to the same illuminant chromaticity. Thus, if we focus on the centre of the specular highlight (cyan circle symbol) during the animation, the chromatic difference we observe becomes small. On the other hand, (b) an illuminant change shows a different change in chromatic distribution as each distribution is extending towards a different illuminant chromaticity. Thus, if we pick the centre of a specular highlight, the chromatic difference should be close to the distance between the two illuminant chromaticities. In other words, if we can

extract from the animation features that have a colour signal that is dominated by the colour of illumination, we should be able to discriminate two changes. That is, when the observed chromatic change is small, we should classify the change as a reflectance change. In contrast, if we instead pick a point that has a diffuse component only, the chromatic change will be similar between the reflectance change and illuminant change trials due to the experimental design. Therefore, the key to solving this task computationally is to select a region that gives information about the illumination.

Importantly, when the specular highlight falls on the dark part of the texture (spatially-separated), the "brightest is white" model picks the pixel that is displaced from the centre of the specular highlight, and thus the model is likely to make a wrong judgement. In contrast, the centre of the specular highlight should give closer information about the illumination for spatially-separated trials than spatially-aligned trials because the dark part of the texture is less contaminated by the diffuse component.

Figure 2.9 shows the histogram of the chromatic difference for all trials if the model selects (a) the brightest pixel and (b) the centre of the most intense specular highlight. The leftmost panels show learned distributions after the training session, that are both based on 300 trials (consisting of 50 reflectance change and 50 illuminant change for each specular level). Here, we see that for (a) the brightest pixel, the peak of the distributions of the reflectance change and the illuminant change are close. By contrast, the two distributions are more separated from the perspective of (b) the centre of the specular highlight. The dotted magenta line shows the intersection of the two distributions, which importantly corresponds to the criterion line. In other words, when the observed difference is larger than the difference indicated by this line, the model should classify the trial into an illuminant change. Thus, whether the trial is classified correctly is determined depending on which side of the line the observed chromatic difference falls.

The right hand histograms break down the left distribution into each condition. First, when the specularity is 0.00, as predicted, histograms are the same for both models, and they are well overlapped. Second, for the (a) brightest model, the spatially aligned condition shows a good separation of the two distributions into the right direction

particularly when the specularity is 0.04, and thus the model should perform well. Conversely, the spatially separated condition shows worse separation.

As predicted, (b) the centre of specular highlight shows the opposite trend to the brightest model which is consistent with the pattern of d-primes shown in Figure 2.7. Especially when the specularity is 0.04, the model separates the two distributions almost perfectly for spatially-separated trials.

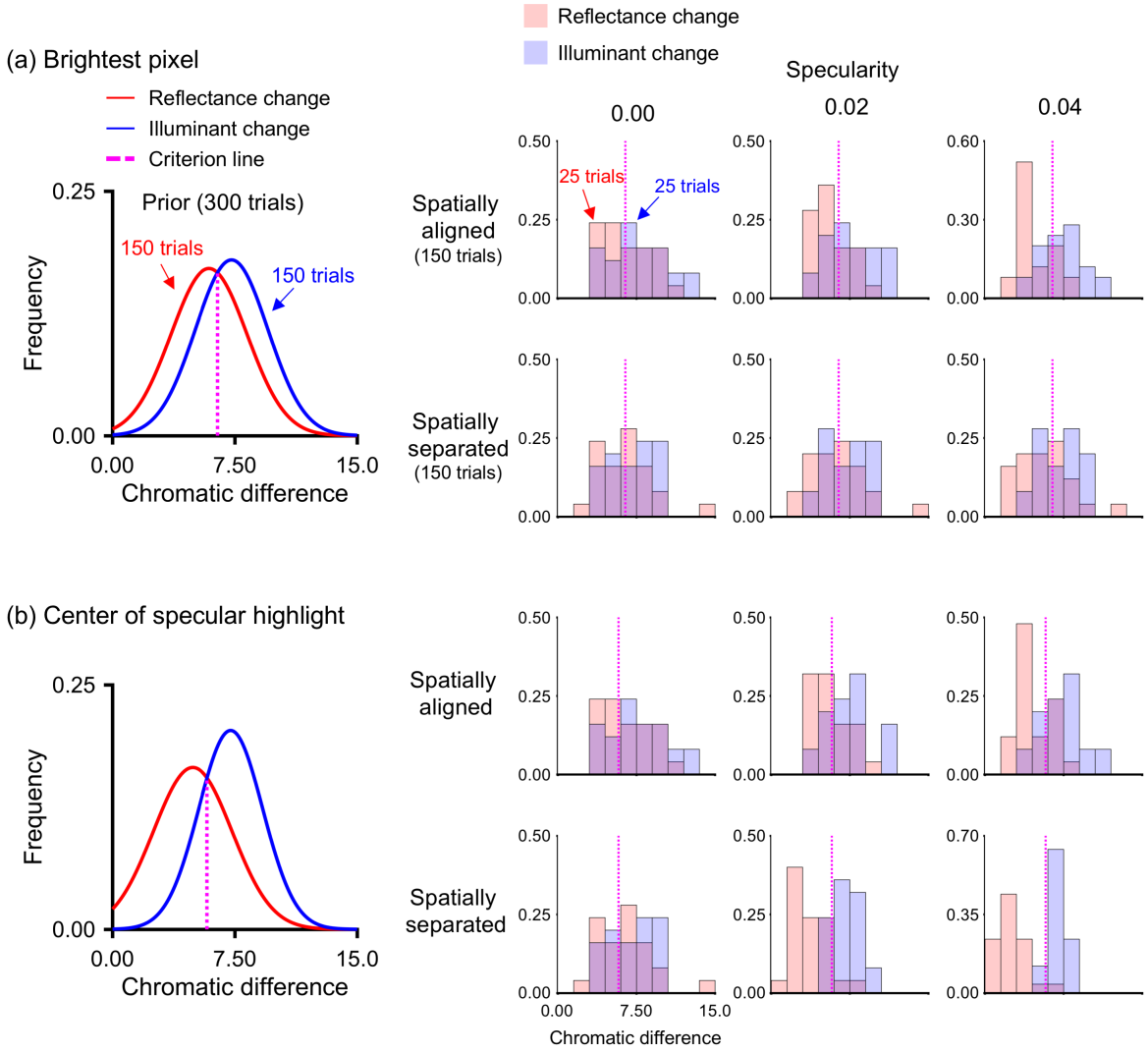


Figure 2.9: Histograms showing the distributions of the magnitude of chromatic change for (a) the brightest pixel and for (b) the centre pixel of the specular highlight for all 300 trials (leftmost histogram) and for each condition (small histograms on the right). The curves in the leftmost histogram are gaussian fits to histograms. The magenta dashed line denotes a criterion line, where smaller and larger chromatic differences than this line are classified into reflectance change and illuminant change, respectively.

3

Reflectance discrimination under complex environmental illumination

3.1 Introduction

Colour constancy is the visual ability that allows us to judge surface colours of objects under various lighting environments. To achieve colour constancy, the visual system needs to cancel the influence of illumination and retrieve an estimate of surface colour. However, such a separation is an under-constrained problem because we typically have access to only the cone signals elicited by the illumination once it has been spectrally modified by surface spectral reflectance. Considerable past research has been conducted to identify mechanisms of colour constancy, as recently summarised by Foster [31], but it is not yet fully understood. One of the limitations of past studies is the use of simplified stimuli where objects were typically two-dimensional, matte and uniformly illuminated by a single light source [29], [55]. Whilst such stimuli allow careful experimental control, they lack some of the important features of surfaces and illuminants that potentially offer cues for colour constancy. One important property of objects in this regard is their specularity. The matte objects that have been used extensively in past studies exhibit only a diffuse reflection whose spectral content is the product of the spectral energy distribution of the illumination and the surface spectral reflectance of the object. By contrast, glossy

objects additionally exhibit a specular reflection that carries direct information about the illuminant spectrum, and can therefore provide information about the lights incident on the object's surface. Thus, colour constancy for glossy objects may be supported by some additional mechanisms that are not available from matte stimuli.

Over the past two decades, colour vision researchers have attempted to address these questions by utilising more realistic objects in three-dimensional setups. Moreover, recent advances in computer graphics techniques have enabled substantial numbers of experiments to investigate various features of material perception [131]–[133]. In addition to studies on gloss and lightness perception [134], [135], [156]–[162] there is an increasing number of studies on colour perception [78], [79], [129], [136]–[140], [163].

Nevertheless, despite the significant methodological advances, even these more recent studies have typically ignored the fact that objects in the natural world are not simply illuminated by a single light source, but also receive light that has been reflected from other objects that coexist in the scene. Thus, different locations on an object's surface receive spectrally-different lights from every direction. Consequently, in the proximal image, there is spatial chromatic variation all over the object's surface. Some regions contain more information about the illumination, whereas others may be dominated by the diffuse component that gives information about surface colour. Estimates of surface reflectance therefore potentially require a local illuminant discounting mechanism. A handful of studies have considered the influence of multiple illumination regions [141], [164], or multiple illuminations that differ in spectral composition and geometry [165], [166], or even the effect of spatial variation across the surface of a glossy object that derives from scene-dependent lighting effects [143]. However, the effect of complex environmental illumination that causes abrupt changes in the light reflected from position to position on an object's surface is still under-explored. The experimental study of complex environmental illumination is now possible using a computer graphics technique that stores for a particular point in a scene the incident light from every direction in the environment [167]. Such environmental illumination typically varies in spectral composition from one direction to another. For example, some directions may contain direct sunlight, whilst others are dominated by skylight, or by light reflected from other

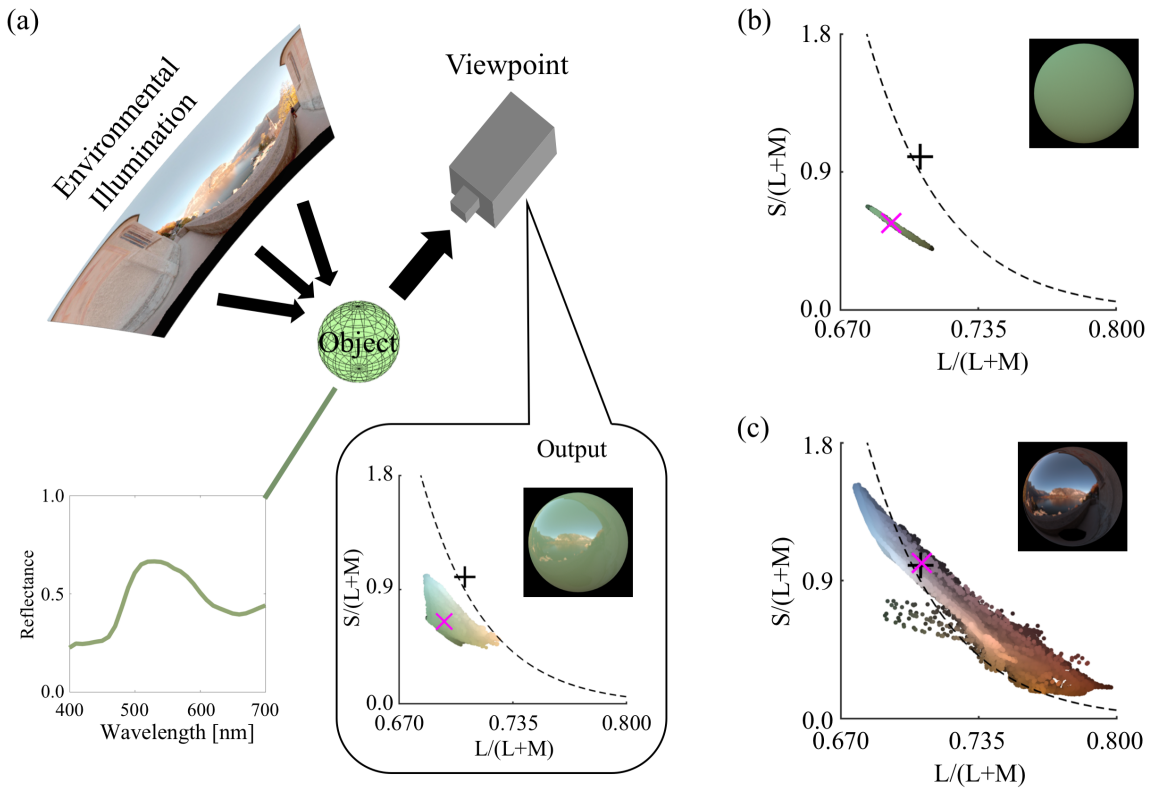


Figure 3.1: Rendering objects in natural lighting environments. (a) Schematic illustration of the rendering process using an environmental illumination. The renderer traces light from the environmental illumination to an object with a particular reflectance and, from there to the viewpoint. The sub-panel in the bottom right shows the resultant image of the object and the chromatic distribution of pixels in that image. (b) Chromatic distribution of the diffuse component. (c) Chromatic distribution of the specular component. (a-c) Chromatic distributions are represented in the MacLeod-Boynton chromaticity diagram [17]. Black plus symbols indicate equal energy white. Magenta cross symbols indicate the mean chromaticity. The black dashed line is the black body locus.

surfaces in the scene. Importantly, environmental illumination includes not only directly emitted light from a light source but also mutual reflections from surfaces that coexist in the environment. When we render a test object at the target location within the environment, the technique allows us to simulate the effect of all incident light that hits the object's surface. For computer graphics, environmental illumination maps are typically stored as an unwrapped two-dimensional image in which each pixel contains incident light from a specific direction. Figure 3.1 (a) shows an example of such a map and the other components of the rendering process, along with the resultant chromatic distribution of pixels that comprise the rendered glossy object. Unlike a flat, matte surface under

uniform illumination, the image of the surface includes a wide range of chromaticities.

As mentioned earlier, the chromaticities in the image of the object arise from combination of two different reflections. Light from any point on the object's surface can be expressed as a linear weighted sum of the diffuse component and the specular component. Therefore, at each point, the chromaticity falls between the chromaticity of the diffuse component and that of the specular component. Figure 3.1 (b) and (c) shows the chromatic distributions of the diffuse and specular components, respectively. A matte object exhibits only diffuse reflection (Figure 3.1b), and although chromaticity varies across the surface, the variance is not large. Conversely, the specular component (Figure 3.1c) reflects the chromaticity distribution of the surrounding environment which spreads widely, mainly along the black-body locus in this case. For comparison, the colour distribution of a glossy object with a single spectral reflectance viewed under simpler illumination such as a single point light source is discussed in detail in General Introduction (chapter 1).

Environmental illumination is complex in that it introduces spatial chromatic variation on an object's surface. Some studies have investigated lightness or gloss perception and estimation of lighting direction under environmental illuminations [168]–[170]. Fleming et al. [171] tested the ability to estimate specularity (from matte to glossy) and roughness (from crisp to blurred highlights), and showed that humans judge these properties well, as long as the spatial structure of environmental illumination under which the test object was placed is natural (as opposed to phase-scrambled). With regards to colour vision, Doerschner et al. [172] investigated the ability to judge the surface colour of a matte surface using the method of achromatic setting, and found that the visual system can eliminate the influence of environmental illumination from a matte surface. However, we still know relatively little about our ability to perceive the surface colour of glossy objects under environmental illumination.

The present study specifically aimed to evaluate our ability to discriminate surface colours (surface spectral reflectance) of objects with various properties under complex lighting environments. Such reflectance discrimination could be considered as an extension of colour discrimination in the presence of chromatic noise [173] because a glossy object could be represented as the diffuse component, which contains the reflectance information,

masked by the specular component. The reflectance discrimination task also connects to the literature on illuminant discrimination [174], [175]. With objects of spatially-uniform surface spectral reflectance, such as the objects that we use here, one potential strategy would be to use the mean colour across the whole surface. Alternatively, performance may be supported by more specific mechanisms that separate the colour that stems from the surface spectral reflectance from the widely spread chromaticities carried in the specular component.

Experiment 1 aimed to explore the potential factors that could influence our ability to discriminate the surface spectral reflectance of matte and glossy objects (either spheres or bumpy spheres) under complex environmental illumination. Experiment 2 was designed specifically to investigate whether our visual system exploits the statistical chromatic regularity in the natural world when making a judgement about an object's colour. In both experiments, we used carefully controlled computer-generated stimuli, and employed a reflectance-discrimination procedure to identify how much reflectance change was needed to successfully select the stimulus with a different spectral reflectance.

3.2 Method

3.2.1 Apparatus

All experiments were computer-controlled and conducted in a dark room. Stimuli were presented on a CRT monitor (NEC, FP2141SB, 21 inches, 1600 $\times$ 1200 pixels) controlled with ViSaGe MkII (Cambridge Research Systems), which allows 14-bit intensity resolution for each phosphor. Gamma correction was performed by the ColourCAL MK II colourimeter (Cambridge Research Systems) and spectral calibration was conducted with the SpectroCAL MKII spectroradiometer (Cambridge Research Systems). The viewing distance was maintained by a chin rest positioned 92 cm from the CRT monitor. Observers were asked to view the stimuli binocularly.

3.2.2 Stimuli

Rendering

All stimuli were generated by computer graphics techniques. The geometry of each scene (locations of the viewpoint or camera, an object and an illumination map) was defined using the 3-D modeling software, Blender (Blender Foundation). Then, rendering was conducted using the physically-based renderer, Mitsuba [176]. We here used the Mitsuba renderer (instead of PBRT as in previous chapter) because the use of environmental illumination is not supported in PBRT. The resulting multispectral images (31 channels, from 400 nm to 700 nm with 10 nm steps) of the rendered objects were converted to LMS cone coordinates based on Stockman & Sharpe 2° cone fundamentals [18] and then finally converted to RGB values for display on the calibrated CRT. We used Rendertoolbox [151] to automate the production of the multispectral images and MATLAB (MathWorks) to convert the images to RGB images.

Environmental Illumination

We used two environmental illuminations from a publicly available database [177], "Distant Evening Sun (Hallstatt)" and "Overcast day at Techgate Donaucity". Figure 3.2 shows images of the environmental illumination maps (top row) and the colour distribution of all pixels. Environments 1 (a) and 2 (b) were used for Experiment 1. Environment 1 (a) and its chromatic inversion (c) were used for Experiment 2. We see that the chromaticity distributes along the black-body locus for environments 1 and 2. These environments were selected to have a different mean chromaticity; the mean of environment 1 is close to equal energy white while the mean of environment 2 is displaced from equal energy white towards blue-green. In Experiment 2 we used a chromatically-inverted environmental illumination to test the effect of the direction of chromatic variation. More detailed rationale is provided in the introduction to Experiment 2. The environmental illuminations were originally 1024×512 images with three channels (RGB), but were promoted to multi-spectral images within the rendering process by Mitsuba using a method by Smits [178]. This method employs seven basis functions of wavelengths. Based on given RGB values, weightings for each basis function were determined and they were

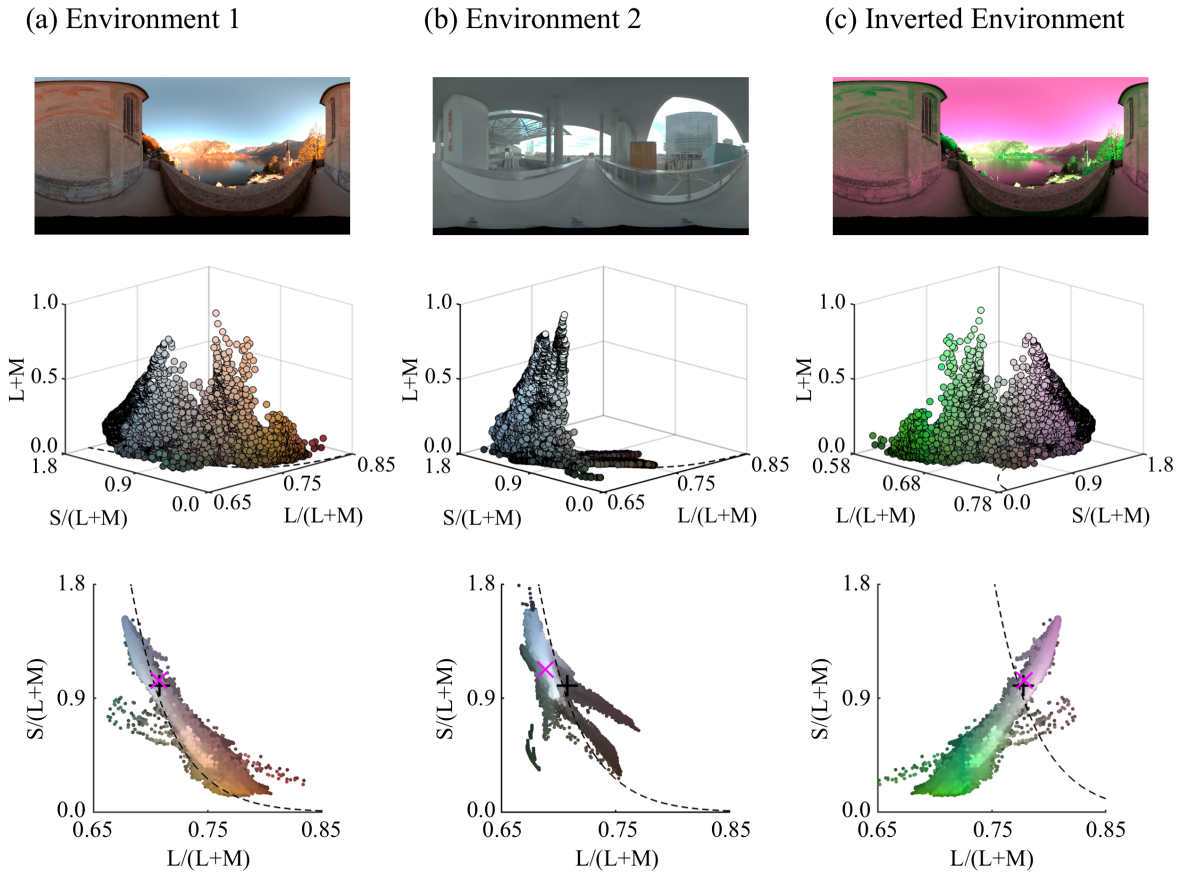


Figure 3.2: Chromatic properties of the three lighting environments used in the experiments. (a) Environmental illumination 1 ("Distant Evening Sun (Hallstatt)"), used in Experiments 1 and 2. (b) Environmental illumination 2 ("Overcast day at Techgate Donaucity"), used in Experiment 1. (c) Chromatically-inverted Environmental illumination 1, used in Experiment 2. (a-c) The top panel shows a 2-D projected image of the 3-D environment map. The middle and bottom panels show the 3-D and 2-D colour distributions of the environmental illuminations. The magenta cross and black plus symbols indicate the mean chromaticity of the distribution and the chromaticity of equal energy white, respectively. The black dashed line is the black body locus.

linearly summed. Weights were determined subject to the constraint that each was greater than zero (to create physically plausible spectra).

Object shape and specularity

In Experiment 1, two types of object shape were tested (sphere and bumpy). In Experiment 2, we used only the bumpy objects. The different object shapes modified the spatial pattern of specular reflection on the object's surface. The sphere provided a spatially clear reflection of the surrounding environment whereas the bumpy object provided a

distorted reflection. We used two levels of specularity for both experiments: a completely matte surface (control) and a glossy surface with a specular reflectance of 0.20 across all wavelengths as defined within the Mitsuba renderer[176] that was set to use the Ward reflectance model [142] that linearly sums diffuse component and specular component. Note that the level of specularity 0.20 may not be directly comparable with chapter 2 as it employed a different renderer (PBRT).

Surface spectral reflectance

Measuring thresholds of reflectance discrimination required continuous and systematic control of the surface's spectral reflectance. Thus, we selected eight reflectance functions as shown in Figure 3.3 (a) from a database of natural objects [153]–[155] and, by combining a spectrally flat (equal energy) reflectance with each of the eight reflectances in differing proportions, we controlled the reflectance along eight colour directions. Reflectance functions were specified between 400 nm and 700 nm in 10 nm steps. In Figure 3.3 (a), each reflectance is normalised by its highest value for the sake of visibility, but in the actual experiment all reflectances were scaled to have equal luminance under an equal energy white illuminant. The coloured circles in panel (b) show the chromaticities of the eight reflectances when rendered under equal energy white. The plus symbol is the chromaticity of the spectrally flat reflectance under equal energy white. Panels (c)-(f) show the effects of environmental illuminations on the mean chromaticity of objects for all conditions in Experiment 1. Semitransparent symbols show the data points re-plotted from panel (b) for the purpose of comparison. We see that environment 1 only minimally distorts the arrangement of eight chromaticities whilst environment 2 vertically expands the chromatic circle and shifts the overall position towards higher S/(L+M) and lower L/(L+M). The effect of shape on mean chromaticity is small, but the bumpy shape has a systematically slightly lower S/(L+M), because the bumpy object has attached shadows that block some of the light coming from above (i.e. blue sky in this environment map). For glossy objects, the change in reflectance has less influence on mean chromaticity. We independently rendered matte objects and glossy objects, and scaled the corresponding matte and glossy images to keep the total energy in the

environmental illumination constant for both objects. Consequently, when more light is reflected in the specular component, less is available for the diffuse component so the surface spectral reflectance has less influence on the proximal image.

Observers and Ethical approval

Three observers (JH, SR, and TM; TM is first author of the study) participated in Experiment 1. For Experiment 2, JH and TM were again recruited as well as two additional observers (AKH, TD). All observers were aged from 23-26 and had corrected visual acuity and normal colour vision as assessed by the Ishihara pseudoisochromatic plates. All observers had previous experience in colour constancy experiments and were not naive to the purpose of the study. This study was approved in 2016 by the Medical Sciences Inter-Divisional Research Ethics Committee at the University of Oxford, in agreement with the Declaration of Helsinki.

3.2.3 Procedure

Before each block, observers adapted to random-dot 20 Hz temporal chromatic noise for two minutes (Figure 3.4, (a)). The noise consisted of the chromaticities of the eight reflectances under equal energy white with equal probability. Thresholds were then measured using a maximum likelihood adaptive staircase method based on the procedure implemented by the Palamedes toolbox [179]. The staircase dynamically estimates threshold based on the history of responses during the experiment and selects the subsequent stimulus magnitude to be close to the estimated threshold to allow efficient convergence of the staircase. During each trial, four objects were simultaneously presented for two seconds as shown in Figure 3.4 (b). The four objects were always presented from different camera angles in order to prevent observers comparing the colour of specific points across objects. We prepared 36 viewpoints (from 0° to 350° with 10° steps) for each rendered image. On each trial, the viewpoint for each object was randomly assigned. Three distractor objects always had a flat reflectance and the other target object had a reflectance that was biased towards one of the eight hue directions defined by the eight reflectances. The observer's task was to indicate the object that had a different

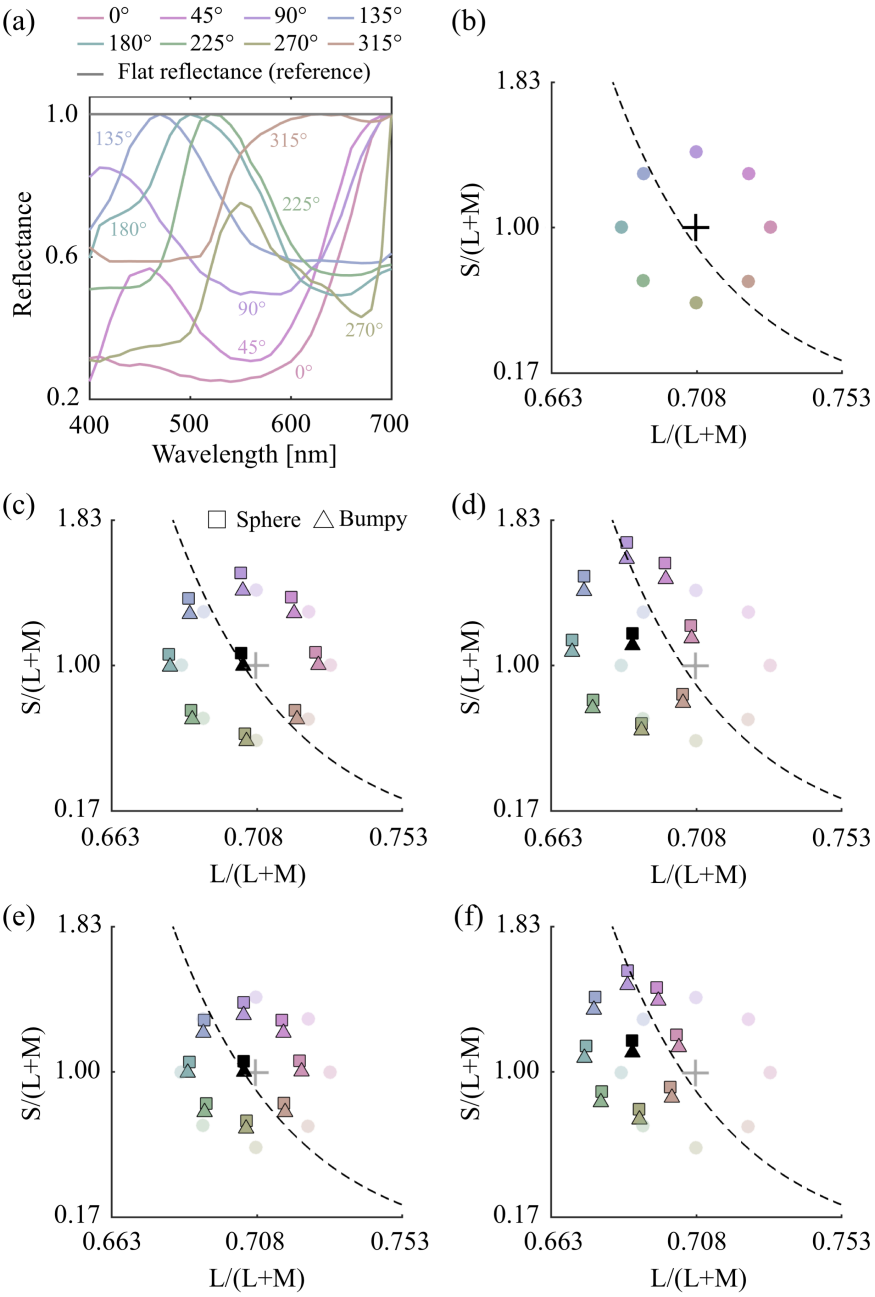


Figure 3.3: Chromatic properties of the surface spectral reflectance functions used in the experiments. (a) Eight surface reflectances used to measure reflectance discrimination thresholds from the spectrally flat reference reflectance. Each reflectance is independently normalised by its maximum value for the sake of visibility. Note that in the actual experiments they were normalised so that all reflectance functions would produce stimuli of equal luminance when rendered under equal energy white. (b) The coloured circles show the chromaticities of the eight reflectances under equal energy white. The plus symbol shows the chromaticity of the flat reflectance under equal energy white. The black dashed line is the black body locus. (c)-(f) Effects of environmental illumination on the mean chromaticity of rendered objects for conditions in Experiment 1. The square and triangle symbols show sphere and bumpy conditions, respectively. The semitransparent symbols are re-plotted from (b), for comparison purposes. (c) Matte objects under Environment 1, (d) Matte objects under Environment 2, (e) Glossy objects under Environment 1, (f) Glossy objects under Environment 2.

reflectance. Observers were instructed to find the object with a different surface colour. There was no fixation point, and observers were instructed to move their eyes to look at each object. The participant's response determined how the surface reflectance of the target object was updated for the next trial: when the response was incorrect, the update was towards the target reflectance, and when the response was correct, the update was towards the flat reflectance (radially in MB chromaticity diagram). There was a one-second inter-trial interval filled with the same temporal chromatic noise used at the start of the block. Thresholds for the eight hue-directions were measured in parallel in the same block (eight interleaved staircases). The presentation order of the eight hues was randomised. The staircases were judged to have converged when the standard deviation of stimulus magnitude of the last 10 trials was smaller than 2% of the prepared stimulus range. This was a conservative criterion as the standard deviation of the thresholds was generally larger. In addition, trajectories of all staircases were manually checked by the experimenter. Each staircase typically needed 20-30 trials to converge. When one staircase reached threshold, that staircase dropped out, and trials continued until all eight staircases converged.

One session consisted of eight blocks (conditions) in Experiment 1 and four blocks in Experiment 2. Specific conditions are detailed in each experimental section. Observers performed five sessions in total for each experiment. All procedures including the observer's task were identical for both experiments.

3.3 Experiment 1

Experiment 1 was designed to explore the effects of environmental illumination, specularity and shape of objects on thresholds of reflectance discrimination.

3.3.1 Conditions

We used eight conditions, consisting of the combination of two environmental illuminations (environment 1, and environment 2), two specular levels (glossy, and matte as a control) and two shapes (sphere, and bumpy). Figure 3.5 shows example objects with a spectrally flat reflectance rendered under settings for all eight conditions.

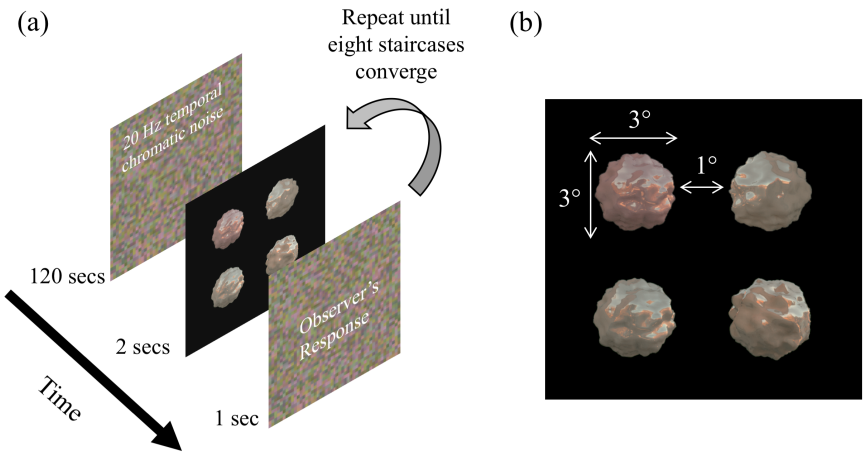


Figure 3.4: Schematic illustration of the procedure. (a) After an initial adaptation to chromatic noise, a series of 4AFC trials was presented, until eight interleaved staircases converged. See the main text for details. (b) An example of stimulus presentation on a single trial. Four objects were simultaneously presented for two seconds, and the observer's task was to select the one with a different spectral reflectance. All objects in a trial had the same level of specularity (either matte or glossy) and the same 3-D shape (either sphere or bumpy). The viewpoint from which objects were rendered was different for the four objects presented, so the distractors were not identical to one another.

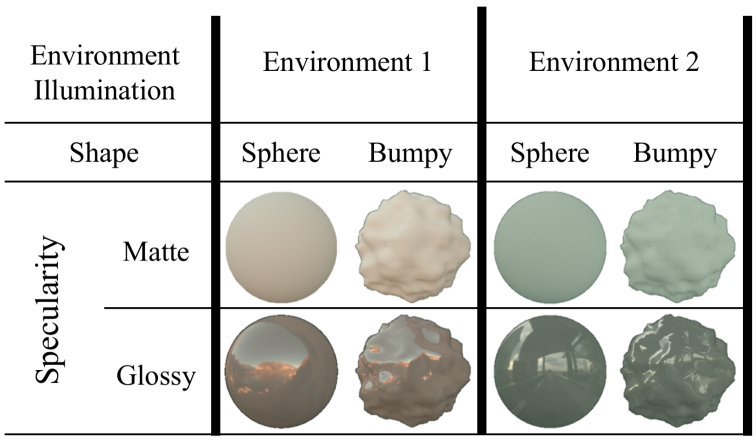


Figure 3.5: Summary of conditions in Experiment 1 with example objects that have spectrally flat reflectance functions.

3.3.2 Results and Discussion

Figure 3.6 (a) shows thresholds of reflectance discrimination for all observers. The data were averaged across five repetitions. Each point indicates the chromaticity of the threshold reflectance under equal energy white. This reflectance-based plot is not the only way to represent results, but it allows us to compare the results directly across conditions; i.e. overlap of the plots means that same physical reflectance was chosen as a threshold. Data are plotted in a scaled MacLeod-Boynton (MB) chromaticity diagram, where equal energy white corresponds to the origin and each axis is independently scaled by chromatic discrimination thresholds along L/(L+M) and S/(L+M) measured prior to the experiment. This threshold was measured from one observer (TM) using a uniform colour disk. This was not intended to measure precise thresholds, but instead aimed to scale the axes so that colour variation in both directions was given approximately equal consideration in subsequent measurements. The left column shows the results for environment 1 and the right column shows the results for environment 2. In the figures, the cyan circle symbols represent the sphere condition, and the magenta diamond symbols represent the bumpy condition. The dashed and solid lines are the matte and glossy conditions, respectively. The black dashed line is the black-body locus. The red dashed line is the axis that corresponds to the maximum variance of chromaticity in the environment map.

Three notable results emerge from environment 1. First, the glossy condition shows a slightly larger discrimination ellipse than the matte condition. Second, the effect of shape is small. Third, and importantly, the discrimination ellipse is tilted along the direction of maximum variability of colours in the environmental illumination (red dashed line). These trends were consistent across all observers. For the results in environment 2, we see generally similar trends to environment 1, but the overall ellipses were noticeably smaller. One of the differences between two environments is that chromaticity does not distribute around equal energy white in environment 2, which might help observers to separate the colours that belong to the illumination and those that belong to the surface.

These results are presented from the perspective of reflectance. However, the actual mean chromaticities of rendered objects under each environment are shifted from these

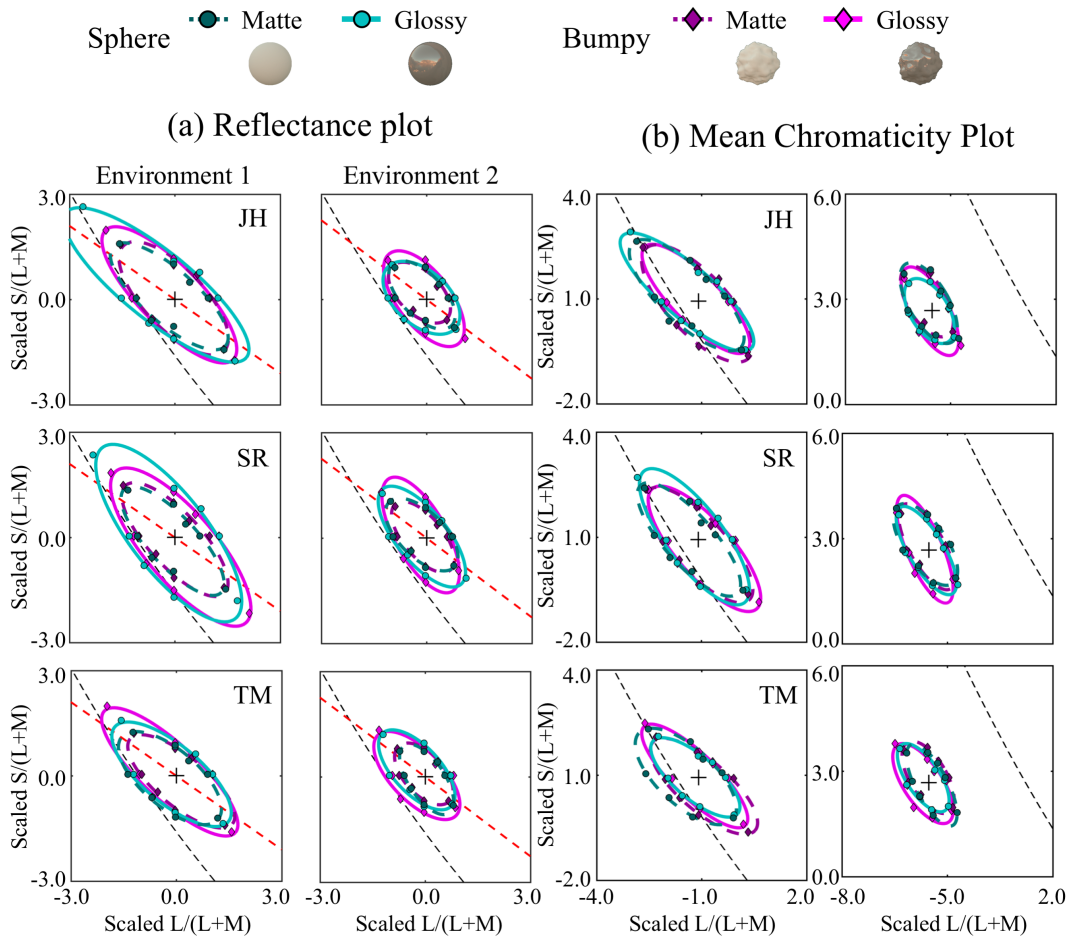


Figure 3.6: Results from Experiment 1. (a) Reflectance discrimination thresholds plotted in a reflectance-based plot, where each data point is represented by the chromaticity of the reflectance function at threshold, viewed under equal energy white. (b) Reflectance discrimination thresholds plotted in a mean-chromaticity-based plot, where each data point is represented by the mean chromaticity of the object at threshold. (a & b) Different rows show data from different observers. Different columns show different environmental illuminations. The black dashed lines show the black body locus. The red dashed lines in (a) show the axis that exhibits the maximum variation in chromaticity of the environmental illumination. Data are plotted in a scaled MB chromaticity diagram, where equal energy white corresponds to the origin and each axis is independently scaled by chromatic discrimination thresholds along L/(L+M) and S/(L+M) that were measured prior to the experiment.

points, as shown in Figure 3.3 (c)-(f). Thus, Figure 3.6 (b) provides another representation where each point indicates the mean chromaticity of objects with the threshold reflectance. Strictly speaking, mean chromaticity changes depending on camera angles presented, but effects of viewpoint were averaged.

A simple discrimination model based on mean chromaticity, which assumes that observers are able to make a discrimination when the mean chromaticity of the target object becomes far enough from the mean chromaticity of distractor objects, would predict circular discrimination ellipses. However, we see that plotting thresholds in such a way does not eliminate the elongation of the ellipse or the way it is tuned in colour space. Consequently, the analysis of chromatic thresholds confirms that we have worse discrimination in a direction that corresponds to the major axis of variation of colours in the environment. A plausible account would say that spatial colour variation of the specular component is greater in that direction and selectively masks co-aligned colour differences [173] that arise from reflectance changes.

The enclosed area and eccentricity of the discrimination ellipses provide summary measures of discrimination performance. Figure 3.7 allows us to compare the area and the eccentricity for all conditions in terms of the reflectance plot (panels (a), (c)) and the mean chromaticity plot (panels (b), (d)). Error bars indicate ± 1 S.E. across observers. From the area plot (panel (a)), we see that discrimination is consistently worse in environment 1. Plotting in terms of mean chromaticity (panel (b)) rather than reflectance reduces the differences between matte and glossy conditions. The eccentricities of the ellipses are nearly equal across conditions (panels (c) and (d)) meaning that the ratio between major and minor axes remained constant.

A three-way repeated-measures analysis of variance (ANOVA) was performed with environmental illumination (environment 1 and environment 2), specularity (matte and glossy), and shape of the object (sphere and bumpy) as within-subject factors for the area of ellipses in the reflectance plot. We found a significant main effect of the environmental illumination ($F(1,2) = 48.8$, $p = .0199$); however, the main effects of specularity and of shape were not significant ($F(1,2) = 17.9$, $p = .0516$; $F(1,2) = 0.58$, $p = .526$, respectively). Also, no significant interactions were observed (environmental illumination $\times$ specularity,

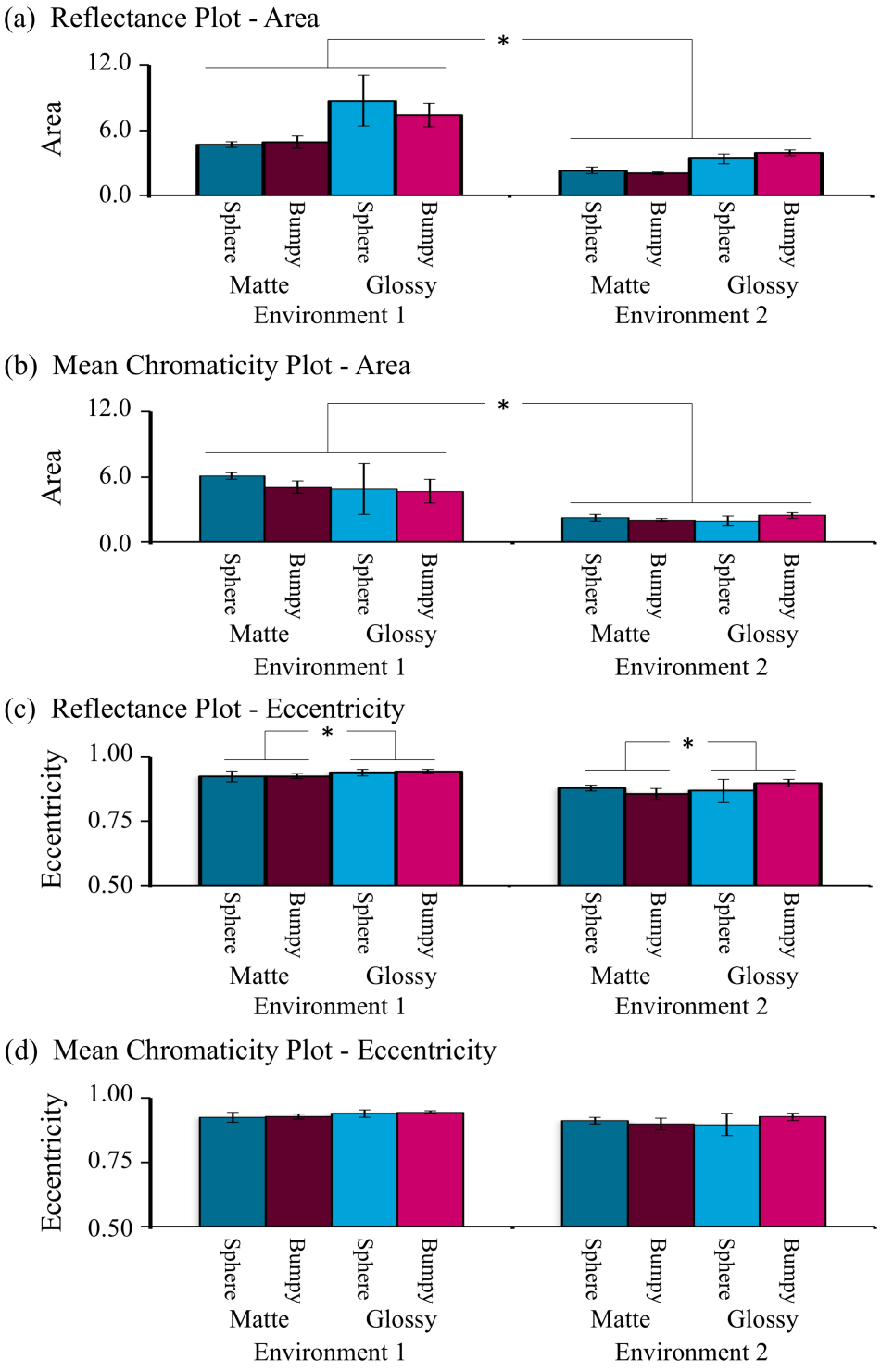


Figure 3.7: Summary of reflectance discrimination performance across conditions of Experiment 1. (a) Mean area of ellipses measured in the reflectance-based plot. (b) Mean area of ellipses measured in the mean-chromaticity-based plot. (c) Mean eccentricity of ellipses measured in the reflectance-based plot. (d) Mean eccentricity of ellipses measured in the mean-chromaticity-based plot. Error bars indicate ±1 S.E. across all observers. The stars indicate p-value less than 0.05.

$F(1,2) = 3.46$, $p = .204$; specularity $\times$ shape, $F(1,2) = 0.29$, $p = .644$; environmental illumination $\times$ shape, $F(1,2) = 2.82$, $p = .235$; environmental illumination $\times$ specularity $\times$ shape, $F(1,2) = 2.20$, $p = .276$). In the same way, a three-way repeated-measures ANOVA was performed for the averaged area of ellipses for the mean chromaticity plot. We again found a significant main effect of the environmental illumination ($F(1,2) = 66.3$, $p = .0148$); however, the main effects of the specularity and shape of objects were not significant ($F(1,2) = 1.2$, $p = .388$; $F(1,2) = 2.1$, $p = .284$, respectively). Again, no significant interactions were observed (environmental illumination $\times$ specularity, $F(1,2) = 1.69$, $p = .323$; specularity $\times$ shape, $F(1,2) = 5.35$, $p = .147$; environmental illumination $\times$ shape, $F(1,2) = 6.93$, $p = .119$; environmental illumination $\times$ specularity $\times$ shape, $F(1,2) = 0.01$, $p = .929$). For the eccentricity measure, a three-way repeated-measures ANOVA showed the following results. In terms of the reflectance plot (c), we found a significant main effect of specularity ($F(1,2) = 31.7$, $p = .0301$); however, the main effect of environmental illumination and shape were not significant ($F(1,2) = 13.7$, $p = .0659$; $F(1,2) = 7.81$, $p = .108$, respectively). The interaction between the three factors was not significant (environmental illumination $\times$ specularity, $F(1,2) = 0.001$, $p = .978$; specularity $\times$ shape, $F(1,2) = 2.49$, $p = .255$; environmental illumination $\times$ shape, $F(1,2) = 0.001$, $p = .978$; environmental illumination $\times$ specularity $\times$ shape, $F(1,2) = 1.96$, $p = .296$). In terms of the mean chromaticity plot (d), we found no significant main effect of environmental illumination, specularity or shape ($F(1,2) = 2.68$, $p = .243$; $F(1,2) = 10.7$, $p = .0821$; $F(1,2) = 3.39$, $p = .207$, respectively). No significant interaction was observed (environmental illumination $\times$ specularity, $F(1,2) = 0.28$, $p = .650$; specularity $\times$ shape, $F(1,2) = 1.09$, $p = .406$; environmental illumination $\times$ shape, $F(1,2) = 0.03$, $p = .878$; environmental illumination $\times$ specularity $\times$ shape, $F(1,2) = 2.29$, $p = .269$).

Therefore, the environment in which objects exist affects our performance of reflectance discrimination while properties of objects (shape and specularity) have little impact.

One of the interesting effects we found in Experiment 1 was that discrimination ellipses were tuned in a certain way. It is known that the chromaticities of lights available in natural environments typically distribute along black-body (or daylight) locus [180]. If this is the case, it is likely that we would have hue-dependent discrimination

ability in general. Therefore, it would be useful to determine whether the observed elongated ellipse is due to the specific environments under which stimuli were rendered or if it is due to the variation in natural environments that we experience in daily life. In other words, it is possible that sensitivity along the black-body locus is reduced because we adapt to such a chromatic distribution in daily life. Such an effect of long-term adaptation to environmental stimuli has been reported in past studies of colour appearance [181]–[184], and it may be possible that a similar effect is found for threshold-based measurements. If that is the case, we should observe this tuning effect with whatever scene we employ in a rendering process. To address this question, we decided to invert the chromatic distribution of environment 1 to change the direction of chromatic variation as shown in Figure 3.2 (c), and conducted Experiment 2 using the same procedures as Experiment 1. There is no unique way to invert the chromatic distribution of environmental illumination, but we chose to first convert RGB values to LMS values, and then to reflect the chromaticities in a line parallel to the S/(L+M) axis of MB space, where the inversion axis intersected the L/(L+M) value of equal energy white. colours that were out of gamut as a result of this inversion were mapped to the gamut boundary (a manipulation known as "clipping"). Finally, the reverse transformation was performed to return MB coordinates to RGB values for rendering and display.

3.4 Experiment 2

3.4.1 Conditions

In Experiment 2, there were four conditions: factorial combination of two environmental illuminations (environment 1 and the inverted environment) and two specular levels (matte, and glossy). In this experiment, we used only bumpy stimuli.

3.4.2 Results and Discussion

Figure 3.8 shows the results for Experiment 2. We see that inverting the chromatic distribution of the environment also inverts the discrimination ellipse. Therefore, the tuned ellipses observed in Experiment 1 were influenced by the environment in which

objects were rendered rather than solely by the natural environment in which we live. Also, we found that under the atypical lighting environment, discrimination ellipses became less elongated, and larger. Figure 3.9 compares the area and eccentricity of ellipses in both a reflectance-based plot (a, c) and a mean-chromaticity-based plot (b, d).

A two-way repeated-measures ANOVA was performed with environmental illumination (environment 1 and inverted environment), and specularity (matte and glossy) as within-subject factors separately for each panel in Figure 3.9. For the area measure, the analysis produced the following results. In terms of the reflectance plot (a), we found a significant main effect of the environmental illumination ($F(1,3) = 168.9$, $p < .001$) and of specularity ($F(1,3) = 10.6$, $p = .0473$). The interactions between the two factors were not significant. In terms of the mean chromaticity plot (b), we found a significant main effect of the environmental illumination ($F(1,3) = 358.2$, $p < .001$); however, the main effect of specularity was not significant ($F(1,3) = 3.7$, $p = .150$), and again, no significant interaction was observed. For the eccentricity measure, the analysis produced the following results. In terms of the reflectance plot (c), we found a significant main effect of the environmental illumination ($F(1,3) = 19.8$, $p = .0211$); however, the main effect of specularity was not significant ($F(1,3) = 3.06$, $p = .179$). The interaction between the two factors was not significant. In terms of the mean chromaticity plot (d), we found a significant main effect of the environmental illumination ($F(1,3) = 80.3$, $p = .00293$) and of the specularity ($F(1,3) = 10.7$, $p = .0467$). Again, no significant interaction was observed.

The data from Experiment 2 suggest that discrimination of spectral reflectances, would not work well when the chromatic distribution of illumination does not follow the way in which colours distribute in natural environments, i.e. along the black-body locus. This in turn might imply that our visual system can exploit the statistical regularity of chromaticities in environmental illumination to separate which colour-variations across an object's surface are likely to arise from the illumination conditions.

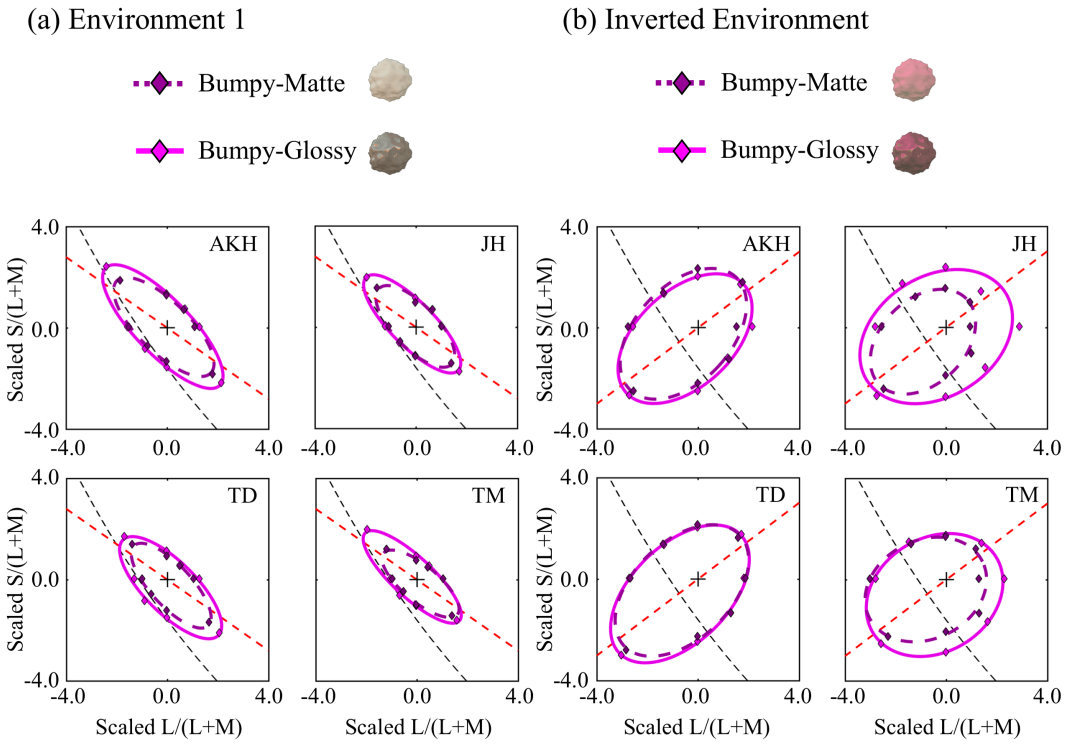


Figure 3.8: Results from Experiment 2, with reflectance discrimination thresholds plotted in a reflectance-based plot (analogous to Figure 3.6a). (a) Data obtained with stimuli rendered under Environment 1. (b) Data obtained with stimuli rendered under the chromatically-inverted version of Environment 1. The axis scaling is the same as in Fig. 3.6.

3.5 General Discussion

3.5.1 Overview

The purpose of the present study was to measure thresholds of reflectance discrimination for glossy and matte objects under various environmental illuminations. For matte objects, environmental variation in incident illumination expands the gamut of chromaticities contained in the diffuse component of the proximal image of the object (Figure 3.1b). For glossy objects, in addition to this variation in the diffuse component, the full variation of chromaticities in the environmental illumination is carried, unmodified, in the specular component (Figure 3.1c).

For a perfectly colour-constant observer, the type of illumination should not affect thresholds for reflectance discrimination. However, the results of Experiment 1 (Figures 3.6 and 3.7) indicated that reflectance discrimination is actually not constant across

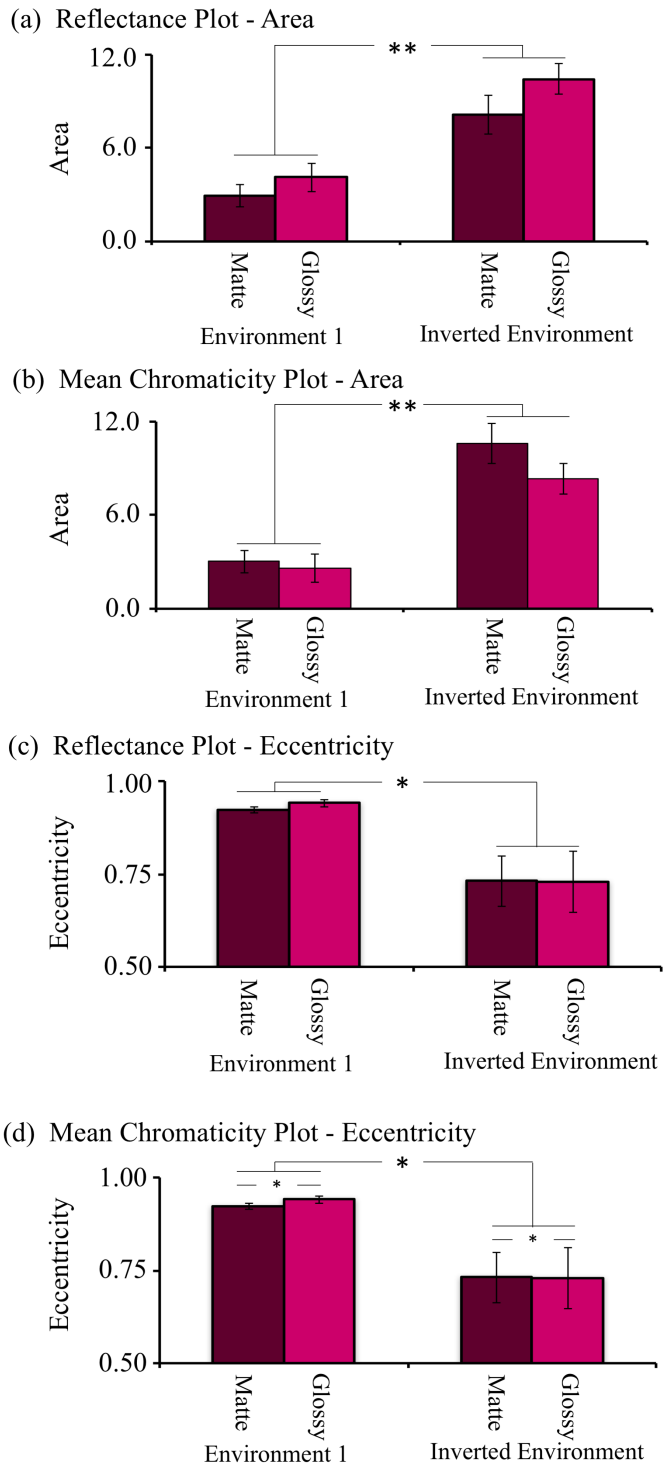


Figure 3.9: Summary of reflectance discrimination performance across conditions of Experiment 2. (a) Mean area of ellipses measured in the reflectance-based plot. (b) Mean area of ellipses measured in the mean-chromaticity-based plot. (c) Mean eccentricity of ellipses measured in the reflectance-based plot. (d) Mean eccentricity of ellipses measured in the mean-chromaticity-based plot. Error bars indicate ±1 S.E. across all observers. The single stars indicate p-value less than 0.05. The double stars indicate p-value less than 0.01.

environmental illuminations, suggesting that our ability to perceive surface colour depends on the lighting environment in which objects are placed. It was also evident that discrimination ellipses were elongated along the direction of maximum variance of chromaticities in the environmental illumination map. The results of Experiment 2 (Figures 3.8 and 3.9), where we used a chromatically-inverted environmental illumination map, showed that this tuning effect seems to be dominated by the chromatic distribution of the environment used for rendering rather than a statistical invariance of natural environments. The chromatically-inverted environment additionally gave rise to larger discrimination ellipses than the typical environment. These findings were consistent across all observers.

It is interesting that worse chromatic discrimination along the daylight locus is also found in so-called MacAdam ellipses [185], which are a set of discrimination ellipses drawn at different locations in colour space. Importantly, MacAdam ellipses were measured using simple chromatic stimuli suggesting that the tuning effect we found in this study may not be specific to objects placed under environmental illumination. Thus, our results could be simply explained by the non-uniform sensitivity of the colour vision system. However, the result from experiment 2 suggests that the orientation of the ellipse is dominated by the chromatic distribution of environmental illumination used for rendering suggesting that any tuning effect is at least partially beyond the explanation provided by MacAdam ellipses. However, the eccentricities of ellipses were lower for atypical environmental illumination, meaning that the worse discrimination along the daylight locus still exists even in the results in experiment 2. Further investigation is desired to isolate these multiple effects that shape discrimination ellipses, but it seems to be reasonable to say that both the chromatic discrimination within the illuminant and underlying variation in sensitivity contribute to discrimination of surface colours under environmental illumination.

3.5.2 Available cues

The 4AFC procedure for reflectance discrimination is a relative judgement that does not necessarily require estimating surface or illuminant properties. Instead, global statistics, such as mean chromaticity, may have allowed observers to select the odd-one-out. It is

possible therefore to consider the task as one of colour discrimination in the presence of a chromatic noise mask, where the noise arises from the spatial variation in incident illumination. For matte objects, which contain only a diffuse image component, the spatial chromatic noise is low contrast and low spatial frequency; for glossy objects, which additionally include a specular image component, the spatial chromatic noise is high contrast and contains variation at many spatial frequencies including sharp edges. There are many studies on colour discrimination with chromatic noise [173], [186]. They essentially show that adding chromatic noise elongates the discrimination ellipse along the direction in which the chromaticity of the noise extends. This is consistent with the obliquely-tuned discrimination ellipses we measure: For different illumination environments, the discrimination ellipse reliably aligns with the axis of maximum chromatic variation. One limitation of our study was that stimuli were rendered from a single viewpoint and therefore presented without binocular disparity. Since diffuse and specular components differ in their imaging geometry, it is possible that disparity information would have helped observers parse the image and judge surface colour. With similar rendered objects under simple illumination, motion parallax does not improve constancy [149]. However, depth information may be particularly important in complex lighting environments.

3.5.3 Matte vs glossy objects

For glossy objects the chromatic variation across the image of the object is much greater than for matte objects. Yet observers' abilities to discriminate spectral reflectances were not significantly different for the two types of object (no main effect of specularity for the ANOVA in Experiment 1 or Experiment 2). The small, consistent trend observed both in Experiment 1 and Experiment 2 towards poorer discrimination performance for glossy objects when analysed in terms of reflectance was eliminated or reversed when discrimination ellipses were plotted with respect to mean chromaticity. This relative improvement for glossy objects when plotted in chromaticity space is consistent with the stimulus properties: When more light is reflected in the specular component, less is available for the diffuse component so a given surface spectral reflectance has less

influence on chromaticities in the proximal image. However, the overall similarity of performance with matte and glossy objects is somewhat counter-intuitive.

As noted above, glossy objects exhibit much greater variation in chromaticity across their surface. Thus, previous measurements of colour discrimination in noise [39], would predict higher thresholds for reflectance discrimination of glossy objects. This raises the possibility that, unlike random chromatic noise, specular reflections may contain some regularities specific to the illumination, which may help the visual system to eliminate the influence of specular reflection effectively. These regularities may be in the chromatic or spatial properties of the noise, and we discuss each possibility below.

3.5.4 Chromatically typical and atypical lighting environments

In addition to the finding that discrimination ellipses align to the major axis of chromatic variation in the environmental illumination used for rendering, our second major result was that discrimination abilities became worse under the atypical environmental illumination in Experiment 2. This decline in performance was accompanied by a reduction in the eccentricity of the ellipse. Had discrimination thresholds depended only on the chromatic distribution of environmental illumination used for rendering, there would have been no change in the overall area of the ellipses between the typical and chromatically-inverted environments, and no change in the eccentricity of the discrimination ellipses. One explanation would be that, irrespective of the particular environment used for rendering, chromatic change parallel to the black-body locus is likely to be assigned to the illumination (and not to a surface reflectance) because chromaticities of lights in natural environments distribute along the black-body locus [180], [187]. This is consistent with the idea of using priors such as statistical regularities in natural environments to solve the problem of colour constancy [54], [115].

The results of Experiment 2 raise the question of whether the visual system understands the chromatic properties of specular reflection of incident illumination. Further investigation is needed to address this question, but it is possible that the visual system internalises chromatic regularities in environmental illumination to separate diffuse and specular components of the proximal image of an object. Indeed, several studies have

reported the importance of the black-body (or daylight) locus for colour vision [174]. The chromatic tuning of natural environmental illumination is heavily determined by the chromatic tuning of natural illuminants. Conversely, for object colours, we know that the spectral reflectances of natural objects are less constrained than the spectral content of natural illuminants. However, from theoretical arguments, distributions of spectral reflectances rendered under equal energy white are expected to show a negative correlation in L/M-opponent and S-opponent signals [188]. We conducted our own analysis and confirmed that the chromaticity of 4824 surface spectral reflectances of natural objects [153]–[155] under equal energy white light distributes broadly along the black-body locus. Therefore, it seems unlikely that colour changes along the black-body locus could be strongly diagnostic of the difference between reflectance-based or illuminant-based variation in the proximal image. Interestingly, chromaticities in #theDress, which is an "illusion" that gives rise to observer-dependent colour appearance, also spread along the black-body locus [37], [49] and this may explain why observers have difficulties in separating surface and illuminant information in this image. The asymmetric effects of chromatic change along particular directions in colour space would be an interesting topic to pursue further.

3.5.5 Spatial signatures of lighting and reflectance changes

The reflectance discrimination paradigm used here has parallels with the paradigm of illuminant discrimination [174], [175]. The stimuli for reflectance discrimination offer a single, spatially uniform, surface spectral reflectance, with complex spatial variation in incident illumination; the stimuli for illuminant discrimination offer complex spatial variation in reflectance, with uniform illumination. Both reflectance discrimination and illuminant discrimination seem to be poor along the black-body or daylight locus. An important link between the two tasks is that they may both depend in part on discrimination of mean chromaticity (or low spatial frequency chromatic signals), which raises the possibility that the chromatic tuning effects are specific to spatial scale.

It has been claimed in many studies that the problem of colour constancy is equivalent to a problem of illuminant estimation [1]. That is, once the visual system recognises the

illumination colour, colour constancy can be readily implemented by globally subtracting the illuminant colour from the whole field of the view. This idea has an attractive simplicity, and it may also be true under constrained situations in which objects are matte and a scene has only a single light source. However, the spatial variation in chromaticity and luminance that derives from the specular component in the image of a glossy object seems to challenge simple estimates of illuminant colour because local illuminant cancellation is required [143], which seems to be unrealistic from the perspective of computational cost. Instead, what the visual system needs is a way to extract the component that attaches to the surface property of spectral reflectance.

When the object is a smooth sphere the specular reflection preserves the spatial structure of the surrounding environment, as shown in Figure 3.1(a). Thus, one might expect an advantage for smooth versus bumpy spheres, since smooth spheres might allow the visual system to more readily understand how the specular components distribute on the surface of an object. However, in the present study, when objects were presented in a dark void, there was no significant difference in thresholds between the sphere and bumpy conditions. This suggests that although the spatial patterning of specular reflections provides important information about the shape of objects [189], spatial variations that arise from surface geometry do not have a strong effect on the ability to judge surface colours. In a follow-up study we intend to present rendered objects within the full rendered environment to assess whether contextual information helps with judgements of surface colour.

The study of material perception, especially considering colour-related effects, has started only recently. Such work is expected to give new insights to understand the colour that is produced by a complex interaction between illumination and surface properties. At the same time, the field of colour constancy needs to move from simplified stimuli towards more complex stimuli that more fully reflect the properties of real objects and their lighting environments. This work represents a preliminary step and it will be important to extend the investigation to different surface properties, shapes and lighting environments so as to assess the generalizability of the findings.

4

Reflectance identification across complex environmental illuminations

4.1 Introduction

Colour is useful for identifying objects in daily life. Even though the lighting environment can change dramatically from one place or time to another, we feel that we have little problem in selecting an object by its colour under various lighting environments. However, the change of scene illumination can substantially alter the lights reflected from objects to our eyes. Therefore, for objects' surface colours to be useful in identification, it is essential that our visual system discounts the influence of illumination. Colour constancy refers to the visual ability to support stable perception of surface colour despite a change of lighting environment. The problem of colour constancy lies in the fundamental difficulty that the reflected light from the object is contaminated by the colour of the illumination at the time the light arrives at our retinas. One typical way to conceptualise colour constancy is that our visual systems estimate the influence of scene illumination in a given scene, and subtract its influence from the whole visual field. Many past studies have been designed with this view in mind, and a rich body of work tried to understand potential cues in the retinal image to implement colour constancy (for review see [30], [31], [190]).

However, these past studies have been limited for the following two reasons. First,

experimental stimuli were often over-simplified in surface properties. Most past colour constancy research used simplified scenes where surfaces are flat and matte as represented by so-called "Mondrian stimuli" [29]. Matte stimuli exhibit only diffuse reflection which carries the product of the spectral composition of illumination and the surface spectral reflectance of the object. The objects in a real world environment, in contrast, often additionally elicit the percept of gloss which is generated by specular reflection, which is a direct reflection of the incident illumination. Since specular reflection provides direct information about the illumination, the visual system might be able to use it to infer the influence of illumination falling on the object's surface.

In addition to this simplification of surface properties, past studies have typically ignored the fact that the illumination in a natural environment holds directional dependency. That is, objects should receive light from every direction, and the spectral contents of those lights might vary from one incident angle to another. For example, when an object is placed outside, the upper part of the object may be directly illuminated by sunlight, but the lower part of the object is more likely to be lit by inter-reflections from the ground and other objects in the scene. Most past studies employed a single-spectrum lighting environment, and there are only a limited number of studies of colour constancy that considered the effects of multiple illumination [141], [165], [166], [191]. There is an increasing number of studies aimed at understanding the mechanisms underlying various perceptual attributes such as gloss, colour and shape under natural lighting environments [168]–[172], [192], [193], but we know relatively little about the extent to which colour constancy holds under complex lighting environments.

Directional-dependent illumination can be simulated with a computer graphics technique called environmental illumination. It uses a directional illumination map defined for a particular environment, where each pixel stores information about the incident light from a particular direction towards a single point in the scene. Such maps allow us to generate the image of an object as it were placed at that location in the environment. Typical environmental illuminations have around 1024×2048 pixels sampling illuminations from every direction, with which we are able to simulate the interaction between directional-dependent illuminations and surface properties.

The simplest view would be that the visual system may need to deal with more than two million illuminations (1024×2048) to achieve colour constancy. However, not all of those illuminations have a direct impact on the reflected light from the objects for the following two reasons. First, in natural environments, it is known that the spatial variation in illumination is relatively smooth [167], [194] meaning that neighbouring pixels are often not independent. Second, when illumination hits a matte surface, diffuse reflections spatially average the influence of incident light. Thus, at the point reflected lights reach our eye, the variation of illumination should decrease. Yet, the influence of illumination spatially varies over an object's surface, and thus it is curious to know under such complex lighting environments how well the traditional correction of a single white point accounts for human colour constancy. As argued by [132] it would be computationally too demanding for the visual system to inversely infer a complete representation of the illumination map since the visual inputs are the combination of scene geometry (i.e. object position, illumination location and view point) and properties of surface and illumination. This is particularly true when an object exhibits specular reflection, generating a fine reflection pattern of the surrounding environment on its surface, preserving the spatial variation of surrounding illuminations. The extent to which we are able to create stable representations of surface colours across complex lighting environments might reveal strategies our visual system uses to implement colour constancy in the highly complex lighting environments in which we live. The present study therefore tests the extent to which single-point illuminant correction accounts for colour constancy under complex lighting environments. We emphasise this attempt is not intended to rule out more complex colour constancy models, but it allows us to see how well simplistic strategy explain the observers' behaviour.

In addition to the consideration of the complexity of surface and illuminant properties, it is also important to select an appropriate methodology to measure colour constancy. Environmental illuminations introduce spatial chromatic variation over an object's surface. Thus, traditional appearance-based measurements of colour constancy such as achromatic settings [111]–[113], [195] or asymmetric matching [101]–[103], [105] might be difficult for observers to judge as the appearance of objects might also vary spatially. In addition,

it is also desirable to choose tasks that can provide a direct measure of our everyday behaviour. Past appearance-based accounts may not explain performance in everyday life. As the degree of colour constancy can vary significantly depending on the experimental task, as argued in [31], comprehensive understanding of human colour constancy would require a naturalistic way of investigating its mechanisms. One common way of using colour in everyday life is to select an object, for example picking a ripe fruit from a pile of fruits in a supermarket. This idea is implemented in several past experiments where observers were asked to select an object under different illuminations, in which observers implicitly need to compensate for the influence of illumination [138]–[140], [196]–[200]. This selection-based colour constancy provides a direct measurement of colour constancy performance and is thus useful for us to understand the human visual systems' behaviour in everyday life. However, these past studies also employed matte stimuli and single-spectrum environments, and thus our current understanding of human colour constancy is still limited to a only a subset of objects and illumination we encounter in a daily life.

These observations encouraged us to investigate how much human colour constancy holds under real-world lighting environment using selection-based colour constancy task. To tackle the question, we ran two psychophysical experiments. In both experiments, we simultaneously presented four computer-rendered objects. The screen was vertically split, and two objects on the left side were rendered under one lighting environment and the other two were rendered under a different environment. Three objects shared the identical surface colour, while the forth, target object had a different colour. The observers' task was "odd-one out", to find the target object. Experiment 1 was designed to explore how well observers perform the colour constancy task under complex illuminations and to identify potential factors that affect observers' performances. We tested whether specularity and the surrounding context affects the observers' performances to identify surface colours across lighting environments. Objects were either matte or glossy, and presented either with surrounding context or in a dark void. For each task, camera angle was randomly selected so that a different part of the surrounding environment was visible to the observer on each trial. The results suggested that the presence of specular reflections had little effect, but the presence of surrounding information

increased performance. One interesting question to ask is what strategy observers uses to implement a stable identification of surface colour in a complex lighting environment. More specifically, we asked to what extent observers' performances can be predicted by applying traditional simple illuminant estimation algorithms, such as mean chromaticity, brightest pixel, or luminance-weighted mean chromaticity. Importantly, we found that a model that integrates the temporal average of the surrounding context over trials predicted observer's performance better than an alternative model that tries to infer the influence of lighting environment from a single trial.

In experiment 2, we tested this hypothesis more directly using environmental illuminations that have high anisotropy, where informative cues to the influence of illumination are not available from the context in a single trial, and thus observers were implicitly forced to integrate chromatic information over trials. To implement such a situation, we employed only two viewing angles. One of which was a congruent viewing angle that provides reliable information about the illumination from a spatial average of the surrounding context in a single trial, and an incongruent viewing angle that was chosen so that a spatial average provides wrong information about the lighting hitting the object's surface. It was confirmed that observer's performance does not decrease for such conditions where the spatial average of the surrounding context does not give us representative information about the whole environmental illumination. Thus, the results were again in agreement with the temporal integration hypothesis.

However, the extent to which the candidate models we tested (based on white point correction) accounted for human observers' behaviour was limited. This implies that the visual system might rely on more sophisticated strategies to compensate for the influence of illumination.

4.2 General method

4.2.1 Apparatus

All experiments were conducted in a dark room. We used a CRT monitor (53.34 centimetres, 1024 × 768 pixels, FP2141SB, NEC, Tokyo, Japan) for the presentation of

experimental stimuli, which were displayed through a visual stimulus generator (ViSaGe MkII, Cambridge Research Systems, Rochester, UK) that allows 14-bit intensity resolution for each RGB phosphor. The gamma correction and spectral calibration were performed via a colorimeter (ColorCAL MK II, Cambridge Research Systems, Rochester, UK) and via a spectroradiometer (SpectroCAL MKII, Cambridge Research Systems, Rochester, UK), respectively. Observers were asked to view the stimuli binocularly. The viewing distance was maintained throughout the experiments at 92 cm supported by a chin rest.

4.2.2 Experimental stimuli

Rendering

All experimental stimuli were rendered using computer graphics techniques. The scene geometry, including the viewpoint and the position of objects, was defined through the 3-D modeling software Blender (Blender 2.78, Blender Foundation, Amsterdam, Netherlands). Physics-based rendering software Mitsuba [176] subsequently generated images of the scene. The rendering was performed multispectrally, from 400 nm to 700 nm in 10 nm steps. The renderer was set to use Ward reflectance model [142] in which diffuse reflection and specular reflection are linearly summed. We converted the resultant multispectral images to RGB images to be displayed on the calibrated CRT monitor to give equivalent cone excitations (defined via the Stockman and Sharpe cone fundamentals [18]). We used Rendertoolbox4 [151] and MATLAB (MATLAB 2016b, Mathworks, Massachusetts, USA) to automate all rendering processes.

Environmental illumination

Four environmental illuminations were chosen from a publicly available database (http://dativ.at/lightprobes/). "Distant Evening Sun (Hallstatt)" (environment 1) and "Overcast day at Techgate Donaucity" (environment 2) were selected for Experiment 1. We used "Direct Evening Sun (Schoenbrunn Vienna)"(environment 3) and "Building Site Interior (Metro Vienna)"(environment 4) for Experiment 2. Figure 4.1 shows images of the environmental illumination maps (top row) and the colour distribution of all pixels in MacLeod-Boynton chromaticity diagram [17]. We used the Stockman and Sharpe

cone fundamentals for the calculation of cone excitations. For Experiment 1, we used environments 1 and 2, which were selected to have a different mean chromaticity. These maps were also employed in our previous study [193]. Environmental illuminations that have high directional asymmetry (detailed in Experiment 2) were selected for Experiment 2. All environmental illuminations were originally in RGB format, but they were promoted internally by Mitsuba [201] to 31 channels from 400 nm to 700 nm in 10 steps.

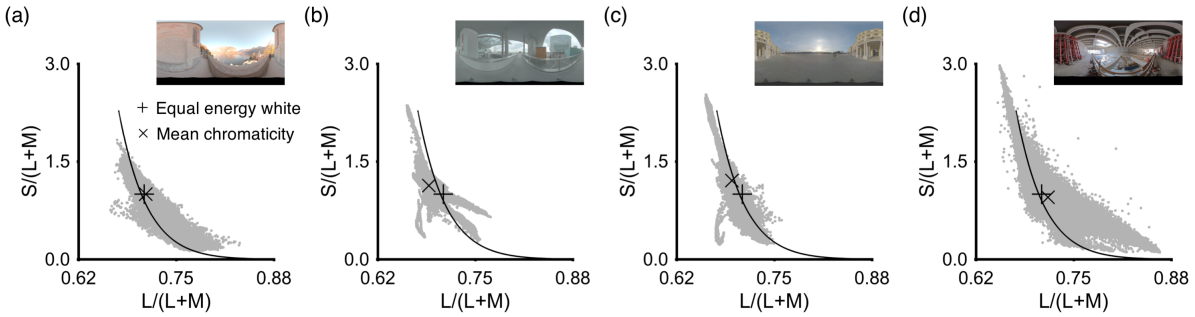


Figure 4.1: Images of four environmental illuminations and the chromatic distribution of their pixels in the MacLeod-Boynton chromaticity diagram. In the plots, the black curve shows the black-body locus. The black plus and the black cross symbols indicate the chromaticity of equal energy white and the mean chromaticity across all pixels, respectively. (a) "Distant Evening Sun (Hallstatt)" and (b) "Overcast day at Techgate Donaucity" were used for Experiment 1. (c) "Direct Evening Sun (Schoenbrunn Vienna)" and (d) "Building Site Interior (Metro Vienna)" were used for Experiment 2.

Object shape and specularity

We used a bumpy sphere in Experiment 1 and a smooth object in Experiment 2, respectively. The bumpy objects produced a distorted pattern of specular reflection on its surface while the smooth object preserves a clearer reflection pattern of the surrounding environment. We used two levels of specularities for both experiment 1 and 2, a completely matte surface and a glossy surface with a specular reflectance 0.20 across all wavelengths defined within the Mitsuba renderer.

Surface spectral reflectance

The present study employed a selection-based colour constancy task (detailed in Procedure), where observers were asked to identify objects under different lighting environments by their colour. The choice of surface spectral reflectance functions is critical because it

directly affects the task difficulty. If we randomly sampled reflectances, objects might end up sharing too similar reflectances for our perceptual system to discriminate, and consequently the task will be too difficult for observers. This perceptual discriminability is hard to define especially for the complex experimental stimuli in this study. Thus, we used findings from our previous study [193] where we measured discrimination thresholds for 8 different hue directions from equal energy white point under a single environmental illumination. The result showed that our ability to discriminate surface colour under environmental illumination was worse along the chromatic direction in which the environment holds most variation which was along daylight or black-body locus (see Figure 4.1). As shown in Figure 4.2 (a) and (b), for Experiment 1, we chose nine reflectance functions for each environment that have chromaticities 2.5 times threshold (black cross symbols) from equal energy white. For experiment 2, we used new environments in which we did not measure thresholds in the previous study. Thus, we used average reflectances from environments 1 and 2. Note that these reflectances were originally selected from a database of natural objects [153], [154], and they were mixed with a flat reflectance to manipulate the distance from equal energy white. We set these selected reflectances to bumpy objects and performed rendering under each environmental illumination. All experimental stimuli are shown in Figure 4.3.

4.2.3 Observers and ethics

Seven observers (AKH, JH, LY, MS, SR, TD and TM; TM is first author of the study) participated in both experiments. The observers were aged from 20-32 and had corrected visual acuity and normal colour vision as assessed by the Ishihara pseudoisochromatic plates. All observers except for TM were unaware of the purpose of the experiment. This study was approved in 2016 by the Medical Sciences Inter-Divisional Research Ethics Committee at the University of Oxford, and was conducted in agreement with the Declaration of Helsinki.

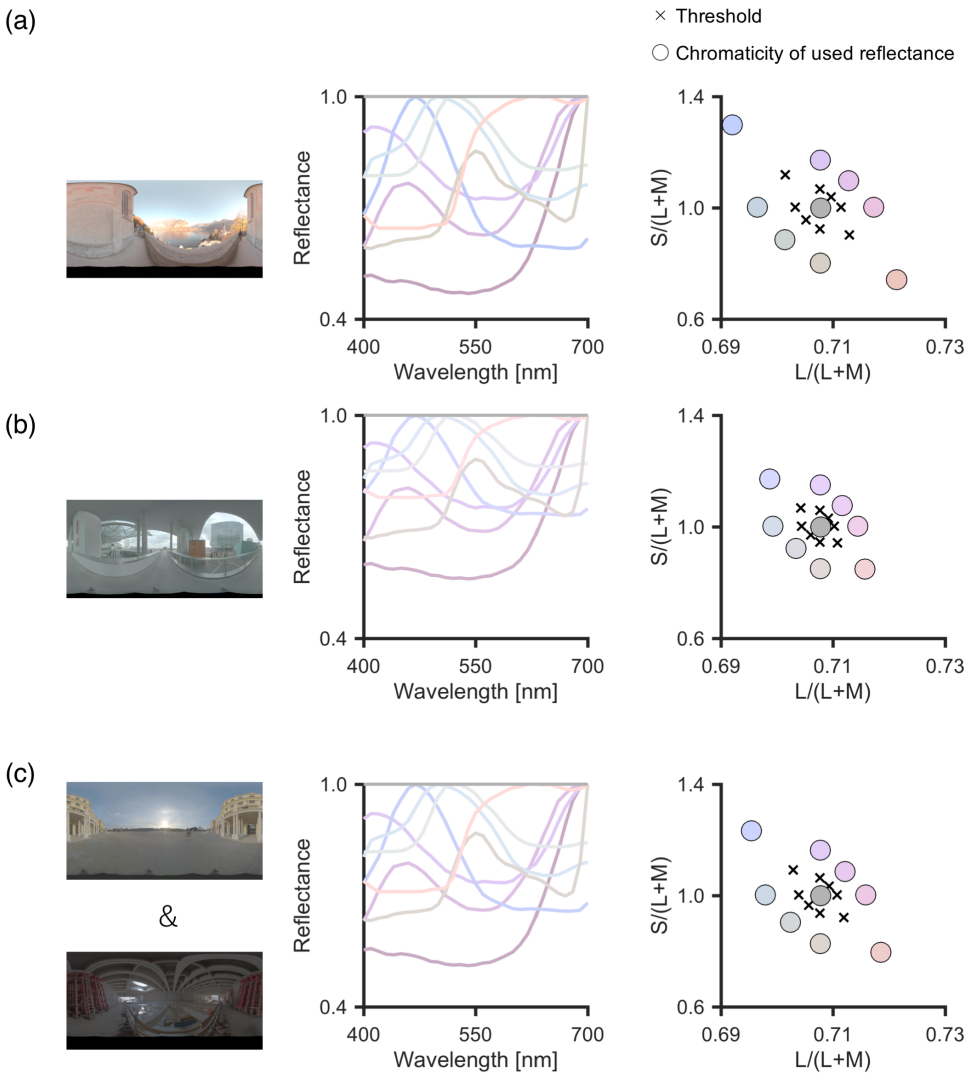


Figure 4.2: Environmental illumination (left column), selected surface spectral reflectances (centre column), and chromatic distribution of reflectances when rendered under equal energy white (right column).

4.2.4 Procedure

Before each block, observers adapted to random-dot 20 Hz temporal chromatic noise with a mean chromaticity and luminance of the nine reflectances rendered under equal energy white, which we used in our previous study (see panel (a) in Figure 3 [193]). Then, observers pressed a button to begin the experiment. For each trial, four objects were presented for 3 seconds as shown in Figure 4.4. The left two objects were rendered under one environment, and the right two objects were rendered under another environment. So that observers could easily notice that left and right regions were under different

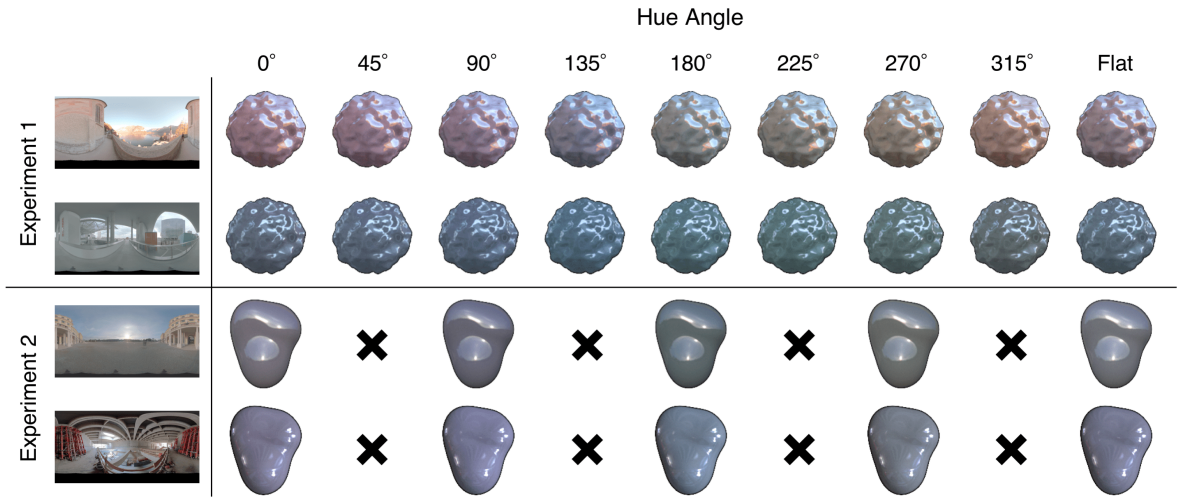


Figure 4.3: All stimuli used for experiments 1 and 2. A bumpy sphere was used for experiment 1 while the smooth object was used for experiment 2. Nine hue angles were used for experiment 1, and five hues were used for experiment 2.

lighting environments, the left and right environments were separated by a vertical gray bar that had the same mean luminance as the whole screen in each trial. For each trial, two reflectances were chosen from the 9 possible reflectances in Figure 4.2. One was assigned to the target object and the other reflectance was assigned to the three distractor objects. The observers' task was to find the target object. The key idea of this experimental design follows. Imagine that you are an observer and a trial shown in Figure 4.4 has just begun. The first thing you notice is that objects 1 and 2 have different surface colours, and that therefore one of them must be the target object. This task to identify the correct side is related to chromatic discrimination. Thus, if an observer selects a distractor object placed on the wrong side, it tells us that the observer simply could not discriminate the surface colours within the environment. After choosing a correct side, in order to answer which of object 1 or 2 is the target object, you need to infer whether object 1 or objects 2 would match objects 3 and 4 when presented under the same environment. This second part of the task corresponds to the observer's colour constancy ability as it requires compensation of the influence of the illuminant to correctly identify surface colours across environments. The example in Figure 4.4 is a relatively difficult trial in which the target has a chromaticity at 135 degrees in MacLeod-Boynton space and the distractor is at 45 degrees, but the task can be easier

(or sometimes even harder) depending on the combination of selected reflectances (see Figure 4.3). This experimental design is analogous to [198], [200]. One block continued until the full combination of 9 reflectances (${}_9C_2$) was tested. Assigning one reflectance to the target object and the other to distractor objects and vice versa were treated as different. As was a target object being placed under the left environment and the right environment. Consequently, one block consisted of 144 trials (${}_9C_2 \times 2 \times 2$).

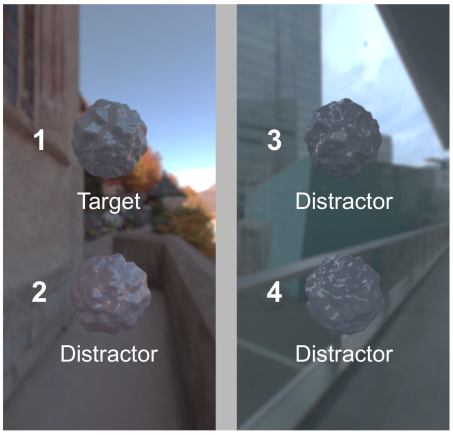


Figure 4.4: Stimulus presentation for one trial. Left and right object pairs are presented under different environments. Three objects shared a surface reflectance (distractor objects) while the fourth, target object had a different surface reflectance. The observer's task was to identify the target object.

4.3 Experiment 1

4.3.1 Conditions

We used four conditions, consisting of two specular levels (glossy, and matte) and the presence or absence of surrounding context (context and no-context). For the no-context condition, rendered objects were presented against a black background (i.e. R = G = B = 0.0). Figure 4.4 shows the case of a glossy-context condition. Each session had four blocks, which corresponds to the four different conditions. Four sessions were completed in total. One block had 144 trials as mentioned in Procedure, and therefore for Experiment 1, each observer completed 2304 trials in total.

Selection of viewing angle

We selected 20 viewing angles for each of environment 1 and environment 2 in the following steps. First, the scene geometry (the location of viewpoint and objects) was set as shown in the panel (a) in Figure 4.5, and then we rendered two objects placed under environment 1 and 2 from 36 viewing angles (0 to 350 degrees with 10 degree steps). Depending on the viewing angle, different parts of the environmental illumination were visible to observers (panel (b)). The visible region from a single viewpoint extended 76.4 $\times$ 41.5 degrees in elevation and azimuth angles, respectively. Raw rendered images were multispectral, but they were converted to RGB values for the experimental display. Then, all pixels were scaled to the 5th percentile pixel value across all pixels and channels (red, green, and blue). This is because environmental illuminations have a high dynamic range, and thus normalisation by the most intense pixel made the images too dark for human observers to see. Consequently, the top 5 % of the pixels were saturated, and they were all rounded to 1.0. At this point, there were still some viewing angles in which the rendered objects were too dark (assessed by the mean luminance across the object's surfaces). Thus, we eliminated the viewing angles that gave the 16 lowest mean luminances to make sure that all objects were sufficiently bright. For the experiment, we adopted the remaining 20 angles as shown in panels (c) and (d). As a result of these selection processes, the mean luminance across the two objects in the presented image was always more than 10.0 cd/m^2 in both the matte and glossy conditions. Also, we made sure that any saturated pixels occurred only in the surrounding context region rather than a part of the object region of the image.

4.3.2 Results - human observers

Figure 4.6 shows the percentage correct value across all trials for each observer in the different conditions. Each panel indicates a different observer and each bar shows different conditions. The mean across seven observers is shown in the bottom rightmost panel. For the analysis, we included only trials where observers chose correctly the side that has the target object. Thus, these values are conditional probabilities for trials where observers were able to perform the colour discrimination task. Consequently, the chance

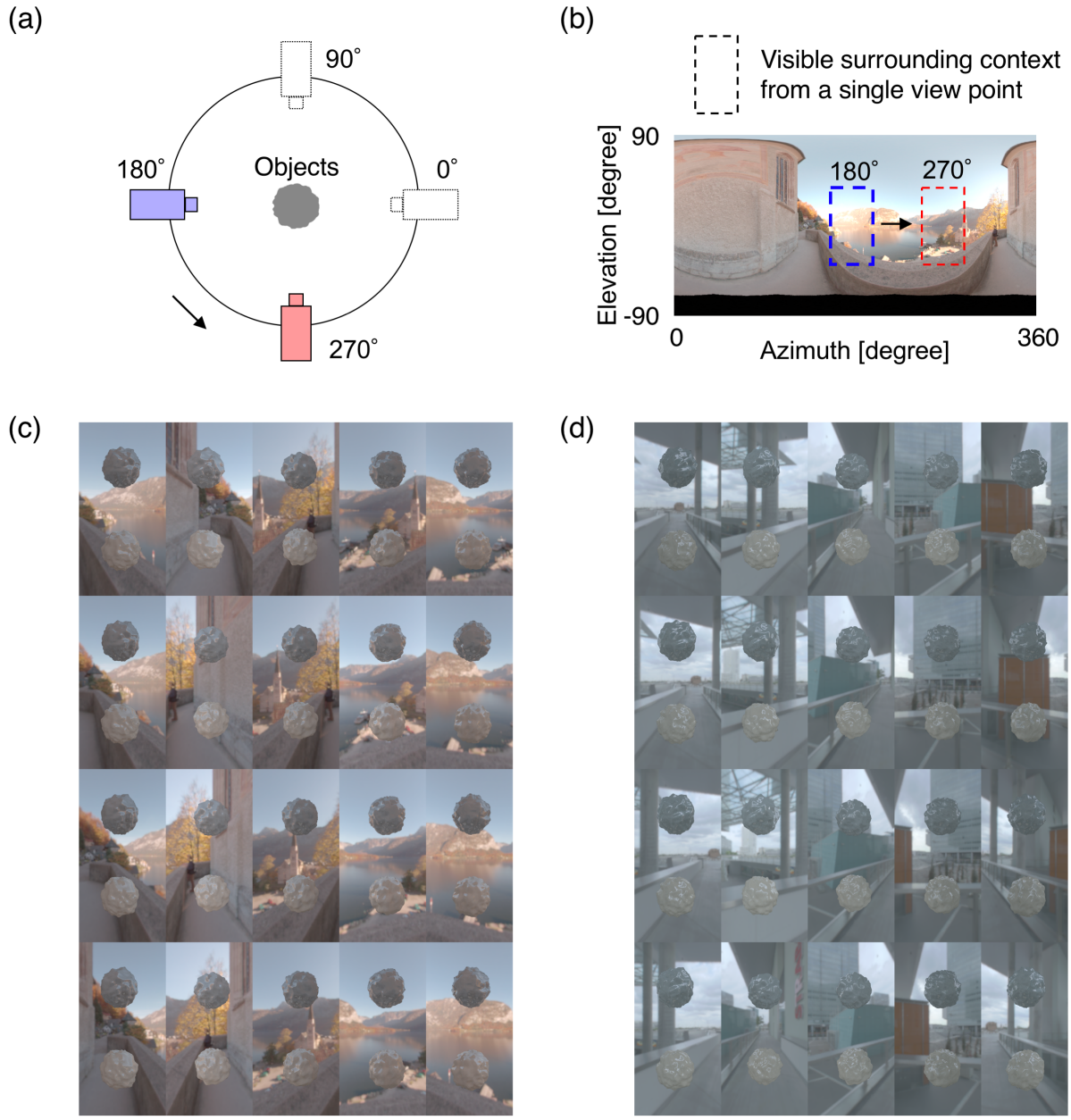


Figure 4.5: The selection of viewing angles in Experiment 1. (a) The scene geometry. (b) The visible region of environmental illumination from a single view point. (c), (d) The 20 selected viewing angles for environments 1 and 2, respectively.

level is 50 % instead of 25 %. In other words, the reason why observers did not reach 100 % here originates from the failure of colour constancy rather than the limits of chromatic discrimination. Observers were able to judge the correct side for 95.6 % of all trials, so excluded trials were only a small proportion of datasets.

A two way repeated measures ANOVA for mean across observers (rightmost column)

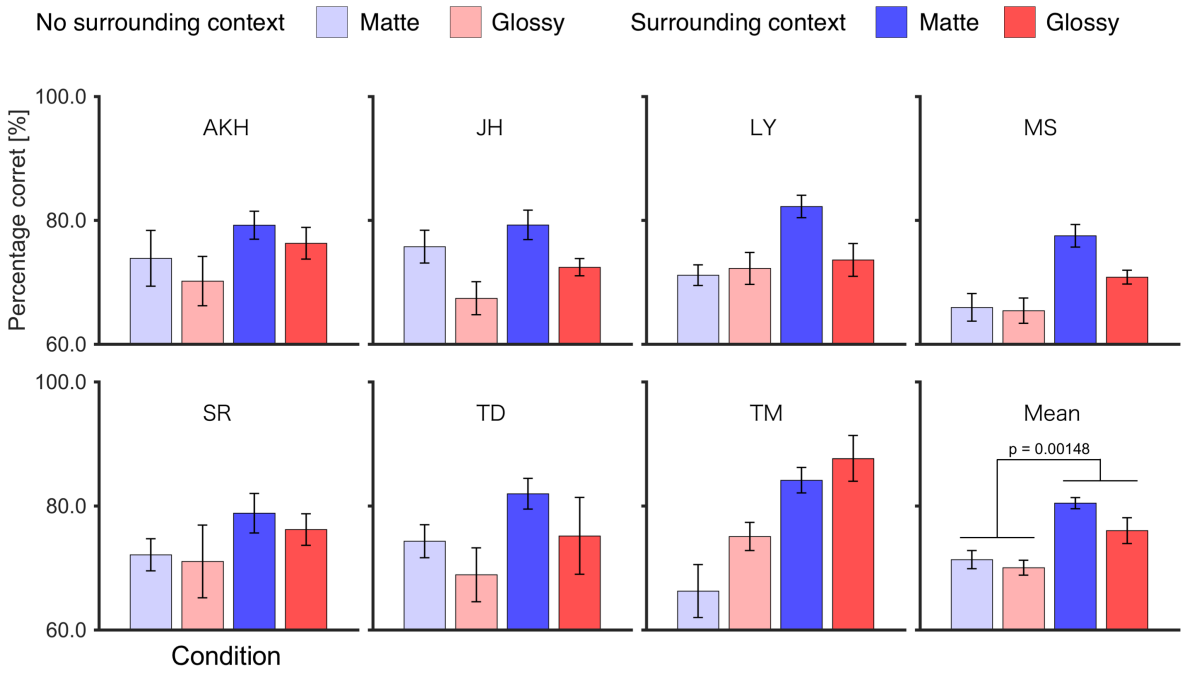


Figure 4.6: Performance across all trials in Experiment 1. Each panel corresponds to each observer, and the bottom rightmost column shows the average across seven observers. Each bar indicate percentage correct value for a different condition. Error bars for individual observers indicate *SE* across four sessions. Error bars for the mean data indicate *SE* across seven observers.

revealed that there is a significant main effect of the presence of surrounding context ($F(1,6) = 30.5$, $p = 0.00148$) while the main effect of specular reflection (glossy or matte) was not significant ($F(1,6) = 2.95$, $p = 0.137$). No significant interaction was observed ($F(1,6) = 4.16$, $p = 0.875$). Therefore, adding the surrounding context helped human observers perform the task. This is presumably because observers were able to more easily infer the influence of the surrounding lighting environment on the colour of objects, which helped them to match surface colour across environmental illuminations.

Figure 4.7 shows performance for individual trials. Each cell indicates the percentage correct value for a particular reflectance pair. Rows show one reflectance and columns show the other reflectance. Note that irrespective of whether the reflectance was target or distractor the data are included in the same cell, and thus each cell gives the average of a maximum 8 repetitions. The same reflectances were not presented in any trial, and thus there is no data along diagonal line (shown by cyan colouring). Each panel shows the averaged values across seven observers. As shown in the right side of the figure

(indicated with a black arrow), the lower left and upper right triangle of each matrix shows trials in which the target object was presented under environment 1 or under environment 2, respectively. In this figure, cells close to the diagonal indicate reflectance pairs that have similar hue angles, and thus they were difficult trials for observers. As the location of cells moves away from the diagonal, the hue angle difference of the reflectance pair becomes bigger, and thus represents easier trials.

First, for the "no surrounding context" condition (labelled as No SC), the region away from the diagonal indicates relatively high performance. In contrast, cells nearby the diagonal has low performance, and in some cases, performance is lower than chance level (50 %) as shown by orange colouring (from 40 to 50 % and from 30 to 40 %). We performed t-test with no correction for each orange cell separately to check if they are significantly different from the chance level. However, none of orange cells showed significant difference from 50 % (significance level 0.05), and thus they can be treated as a chance level performance. By adding surrounding context, overall performance increased and those orange cells mostly disappear. The effect of specular reflection seems to be small or almost absent, for both context and no-context conditions. Overall, these plots show observers performance is strongly dependent on the combination of reflectances, and adding surrounding context enhanced performance.

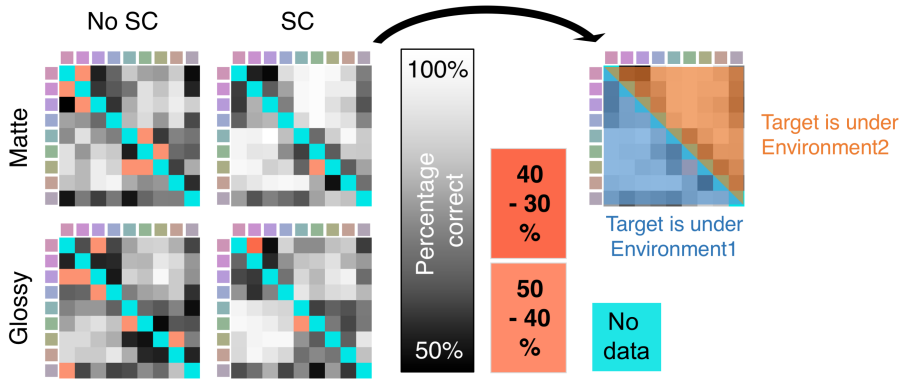


Figure 4.7: Performance for individual pairs of target and distractor reflectances. Each cell indicates the average percentage correct value across seven observers for a certain reflectance pair by shade defined by the colour bar. The four panels corresponds to different condition. The orange cells show the percentage correct values from 40% to 50 % and from 30% to 40 %. As shown in the inserted figure (shown by the black arrow), the lower left and upper right triangle of each matrix shows trials in which the target object was presented under environment 1 or under environment 2, respectively.

These results may be unsurprising as using context information has been claimed as an essential factor in colour constancy. However, in this study we would like to address a deeper question: what signals in surrounding contexts did observers use? Unlike a traditional single-spectrum environment, which has only a single colour of illumination, environmental illumination offers multiple spectra. What cues does our visual system use in such a complex environment to infer the influence of illumination? We tackled this question by developing a computational observer model. Specifically, we asked the extent to which past statistical model can account for the results as explained in the following section.

4.3.3 Implementation of computational observer model

Single trial model

As argued in the previous subsection, the results in Experiment 1 revealed that human observers performed better when the surrounding context is present. However, we would like to address which cues in the surrounding context observers relied on to make judgements. In order to address this question, we constructed a computational model that performed the experimental task in the way shown in Figure 4.8. The model goes through all trials and tries to find the odd-object (target) for each trial, just like human observers were asked to do. The goal here was to find a computational observer model that accounts for human observers' patterns of success and failure.

First of all, the model separates original image to object and surround image regions. In the model, it is assumed that the estimation of objects' surface colours and the illumination colour are performed separately. To estimate the surface colour of each of the four objects, the model used the spatial mean of three cone signals in the object region. We considered three candidate models for illuminant estimation identified in previous studies: (i) mean chromaticity [53], (ii) brightest pixel [52], and (iii) luminance-weighted mean chromaticity [145]. When estimating illuminant and surface colours, gaussian noise was added to the true estimated value based on the idea that this type of estimation may not be perfect in our visual system. The standard deviation of the gaussian noise was set to the thresholds obtained in our previous experiment for each

of L/(L+M) and S/(L+M) axes. Here, it is worth noting that the mean chromaticity estimate is not influenced by the luminance value of each pixel, whilst the brightest pixel estimates solely relies on the pixel that had the highest intensity across the image. For the luminance-weighted mean chromaticity, we averaged all chromaticites weighted by their luminance. The weightings were luminance values raised to the power of 1.0, 3.0 or 5.0. The power 1.0 indicates a simple luminance-weighted mean chromaticity which is equivalent to take a mean of cone signals. The power zero is equivalent to the mean chromaticity. The power infinity is equivalent to the brightest pixel. For this reason, the weighted-mean-chromaticity family of models acted as models between the mean chromaticity model and the brightest pixel model.

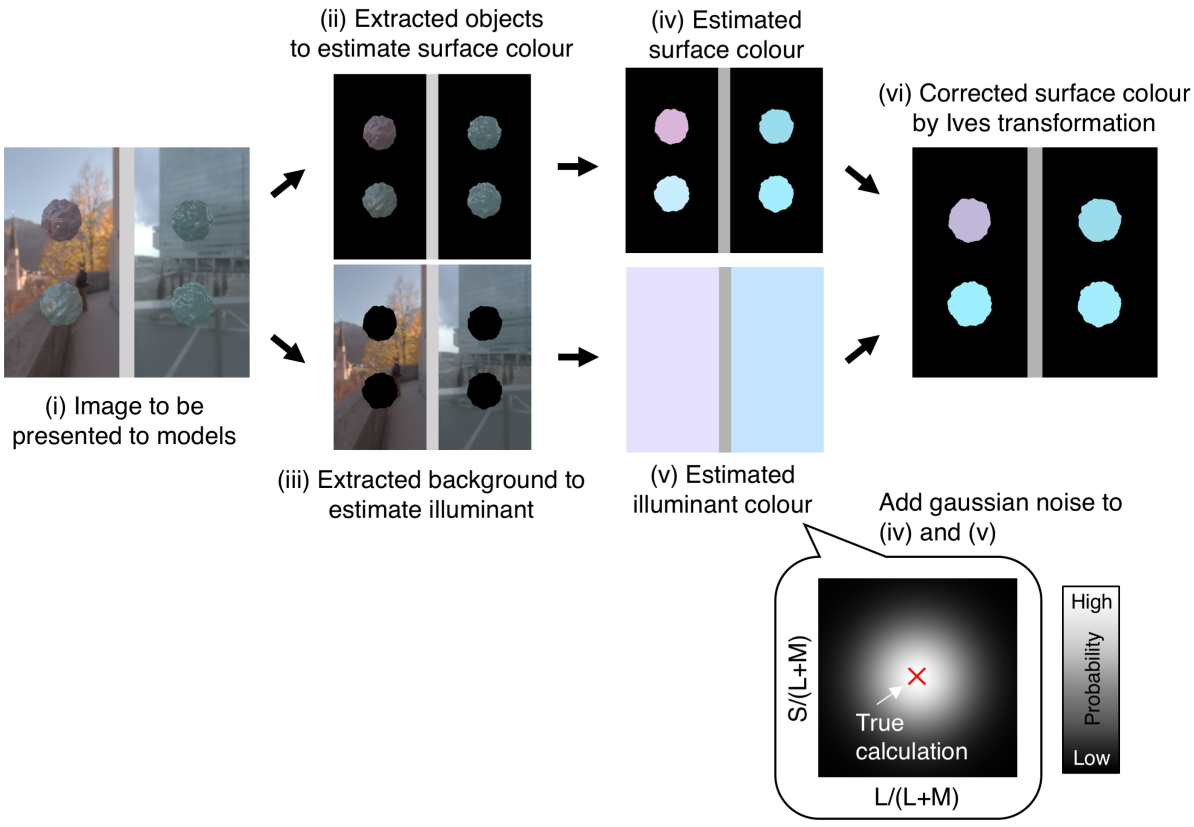


Figure 4.8: A schematic description of the model used in the present study. First, using (i) the presented image, the model segments it to (ii) objects and to (iii) background. Then, (iv) the surface colour of each object is estimated based on mean cone signals over each surface and (v) the illumination colour is estimated based on mean chromaticity, mean cone signals or brightest pixel. For the estimation of surface colour and illuminant colour, we added gaussian noise to the estimated values. Finally, using the Ives transformation, surface colour was corrected and the odd-one out was determined by choosing the outlying one in the chromaticity diagram.

After estimating surface and illuminant colours, the model subtracted the influence of illuminant colour from the estimated surface colours using Ives transformation [88]. This correction was always applied to the surface colours of the two objects that are placed under the environment that has the target object (as shown in (vi) in Figure 4.8). We multiplied the cone signals of these objects by the ratio between cone signals of estimated illuminants as shown in Equation 4.1. When observers make a judgement, they first identify the side that has a target object (by chromatic discrimination within the illumination), and then judge if the upper or lower object matches the objects placed under the opposite illumination. Our transformation reflects this judgement process. Nothing was done for objects placed under the opposite environment. Here, L', M', S' denotes the transformed cone signals. L_t, M_t and S_t indicate the cones signals of the estimated illuminant of the environment that includes the target object. L_o, M_o and S_o indicate the cones signals of the estimated illuminant of the environment that does not include the target object. This equation transforms the target illuminant to the other illuminant, which allows us to compare the surface colours within the same illuminant space.

$$\begin{pmatrix} L' \\ M' \\ S' \end{pmatrix} = \begin{pmatrix} L_o/L_t & 0 & 0 \\ 0 & M_o/M_t & 0 \\ 0 & 0 & S_o/S_t \end{pmatrix} \begin{pmatrix} L \\ M \\ S \end{pmatrix} \tag{4.1}$$

Finally, the model selected the object that had the most different chromaticity from the other three objects in a scaled MacLeod-Boynton chromaticity diagram in which the horizontal and vertical axes were divided by thresholds of chromatic distribution in our previous study, and thereby had roughly similar perceptual scales. Note that for conditions in which the surrounding context is not presented, the models did not perform an illuminant correction and instead tried to find the target object from the estimated surface colour without illuminant correction (as shown in panel iv).

Figure 4.9 (a) shows the overall performance of each candidate model. The leftmost column shows the performance of human observers for comparison. The second left column shows the null-model that always tries to select the target object without any attempt to subtract the influence of the illumination. Therefore, we can consider this model as a baseline showing how models perform without any correction for the influence

of illumination. The null model shows performance around 60 % which is higher than the chance level of 50%. This is presumably because the model can correctly identify the target objects for some trials where the chromatic difference between target object and distractor objects is sufficiently bigger than the effect of the difference in environmental illumination. Notice that the effect of surrounding context is absent for this null model, such that predicted performance is the same with and without surrounding context.

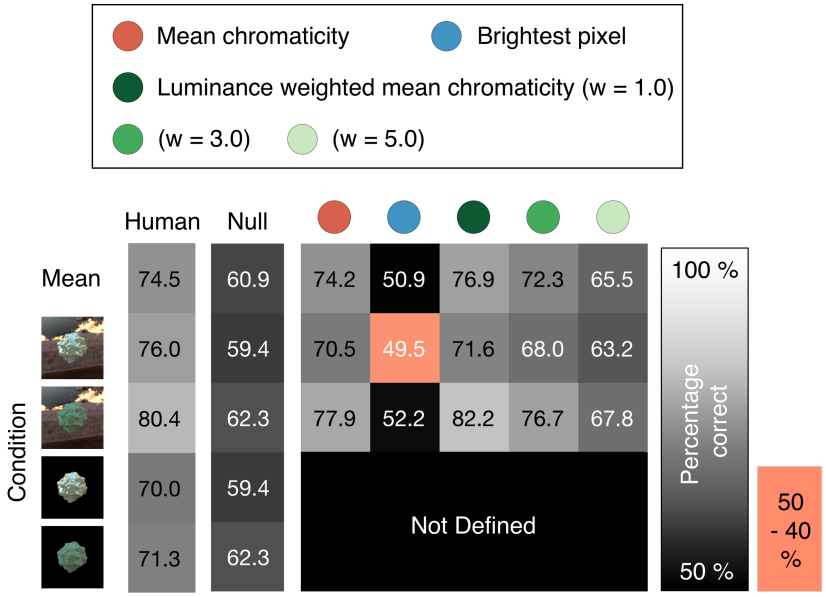


Figure 4.9: Average human performance across all trials compared to model performance. The number in each cell and the shading indicates the percentage correct value. Rows indicate different conditions and columns indicate different models. The orange cell indicates the performance lower than chance level. "Not defined" cells indicate conditions in which the model could not make an illumination correction because there was no background in those conditions, and therefore performance matches the null model.

Next, we consider the effect of subtracting illumination for overall performance. We found consistent improvement by adding surrounding context for all models as observed in human observers' responses, except for the brightest pixel model where correcting the influence of illumination worsened the performance. The degree of performance improvement due to the correction depends on the type of model. In terms of percentage correct, the model based on luminance-weighted mean chromaticity (w = 1) shows the best accuracy, but the mean chromaticity model also showed almost equivalent performance. The performances of models based on luminance weighted mean chromaticity raised by

the power of 3.0 or 5.0 were worse compared to the power of 1.0, and it is expected that this will eventually converge to the prediction of the model based on the brightest pixels.

Interestingly, all models consistently showed better performance for the matte condition than the glossy condition including the null model. This suggests that under this segmentation-based model, the observers' worse performance in the glossy condition can be explained by any of the tested low level features.

Figure 4.10 allows us to see the patterns of performance for all reflectance pair combinations for all tested models. For the sake of comparison, the upper leftmost and the upper second left plot are the average across the seven human observers for no-surrounding context and surrounding context condition, respectively. The bottom rightmost panel shows that the null model can indeed give correct answers when the hue differences are sufficiently large (i.e. far from the diagonal), but otherwise its performance was around chance level. Again, the brightest pixel model does not seem to match the pattern of human observers' performance. In contrast, mean chromaticity or weighted-mean chromaticity models showed reasonable performance and somewhat similar patterns to the human observers' average. Note that the order of trials was randomised for each observer, so the model was also ran seven times following each order, and the results shown here are average across those seven times. There was a slight variation in results every time we ran due to the presence of noise when estimating the illuminant colour and surface colour.

These usual comparisons of percentage correct give us a rough idea of how well the models perform to solve the task and of how similar their predictions are to human observers' performance. However, a quantitative assessment of the similarity between each model performance and human observer performance would be desirable. Therefore, it would be helpful to find an index that quantifies how well the model predicts the pattern of human behaviour. One might imagine that correlation using performance in individual cells will do for this purpose. However, note that only 8 trials go in each cell, therefore the percentage correct value can take only one of 8 values. For this reason, we instead considered responses for all individual trials and thus full correspondence between the model prediction and human observers' performance was available. We can therefore

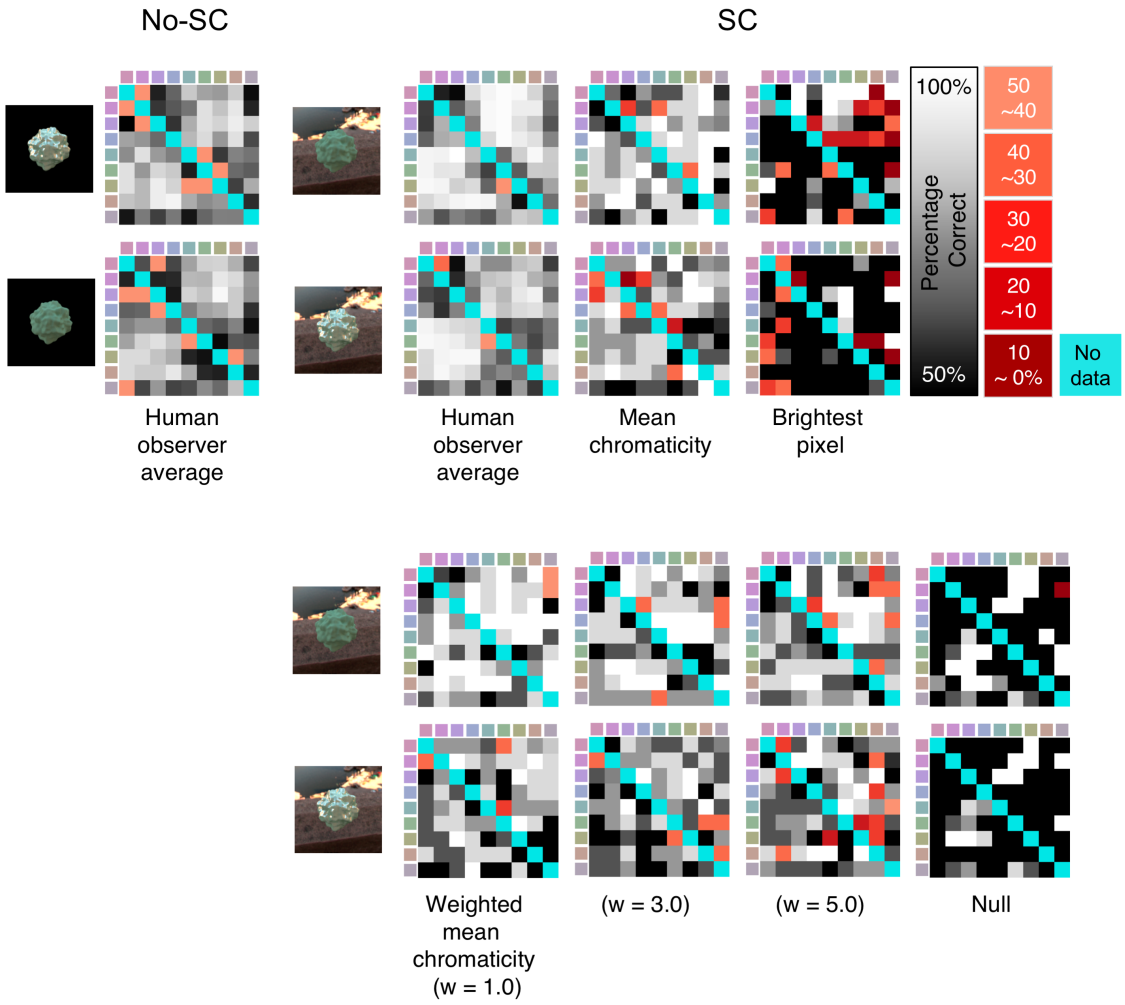


Figure 4.10: Performance across all individual reflectance pairs. The lightness of each cell indicates the percentage correct value. Rows indicate different conditions. The leftmost column shows the average across all human observers for no surrounding condition for comparison purpose. Note that here only surrounding context condition is shown for models because all models will produce the same performance as null model for no surrounding context condition.

use an aggregate metric of classification performance, such as d-prime. The d-prime metric is usually calculated against the correct answer (physical ground-truth) to quantify how well humans or models performed a specific binary task. However, here we calculated it against human observers' responses. For all 2,304 trials that each human observer completed, we classified responses from the computational models into 4 categories. We used only those trials in which both human observer and the model chose the correct side. Using the labels in Figure 4.4, suppose an observer correctly chose the object 1. If a model chose 1 for this trial, this model response was classified as a *hit*. If a model

chose object 2, that was classified as a *miss*. Now, instead suppose an observer chose the object 2 (distractor). If the model chose 1 in this case, the model response was classified as a *false alarm* whereas if the model chose 2, it was a *correct rejection*.

Based on this scoring, we calculated d-prime for all conditions independently for each observer as shown in Figure 4.11. Higher d-prime values indicate that the model is imitating human observers' response patterns well, and lower values indicate that the model performed more differently from human observers. A d-prime of zero indicates no association between the model responses and human responses. Again the null model gives us a rough idea of how well we can predict observers' performances without any correction for illumination. Rather surprisingly none of the tested models showed a particularly higher d-prime than the null model. The highest mean d-prime across conditions was the luminance weighted mean chromaticity (w=1), but the value was almost as high as the null model. The brightest pixel model in particular showed negative d-prime, suggesting that it systematically selected a different response from the human observers. A good model should predict human observers' success and failures, which leads a high d-prime value. Consequently, these results suggest that these models, which try to estimate the influence of illumination from single trial, do not provide a good account of human observer's performance.

Temporal integration model

There are two features in human observers' responses that the single-trial model could not capture. First, regarding the non-context condition, human observers showed much higher performance than the null model as shown in Figure 4.9. This would suggest that humans were using additional cues to perform the task even when the surrounding context was not presented. The simplest alternative would be a model that tries to estimate the influence of illumination from the object surfaces, as the objects, especially glossy ones, should include some information about the illumination falling on them. Thus, we additionally tested the same five models that correct the influence of illumination based on the (i) mean chromaticity, (ii) brightest pixel, and (iii) luminance-weighted mean chromaticity over objects' surface when there is no surrounding context. However, we found that such

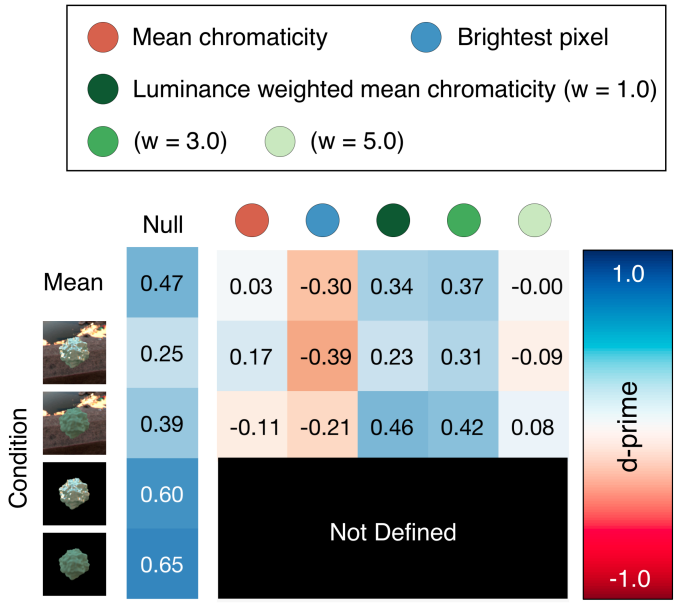


Figure 4.11: A comparison of human and model performance. The number in each cell indicates d-prime calculated against human observer's performance. Rows indicate different conditions and columns indicate different models. Note that these models are all single-trial models that do not care about the history of trials when estimating illuminant colour.

models did not give higher d-prime values than the null model. Thus, this analysis did not support the hypothesis that observers estimated the illumination from the objects' surface when surrounding context is not presented. The fact that performance was no better with glossy stimuli than with matte stimuli also supports this conclusion.

Second, the d-prime value for all models was generally worse or at most as high as the null model, implying that models that try to correct illumination using surrounding context independently at a single trial are not sufficient to predict human observers' behaviour. This would not necessarily rule out that humans used statistics from single trials, as there should be infinite cues based on image statistics that we did not simulate. Nevertheless, one important aspect about our experimental design was that environments 1 and 2 were always presented at the left side and the right side, respectively. Thus, human observers may be able to learn the chromatic properties of the environmental illumination over trials and apply those memories even for trials that did not have surrounding context. Based on this hypothesis, we implemented a model that integrates chromatic statistics over trials. In this context, the single trial model estimates the illumination using only the current trial.

The temporal model integrates statistics exponentially as shown in Figure 4.12. All models adopt the largest weighting on the current trial, but the weightings decrease towards past trials. The degree of the decrease is controlled by time constant τ. The models estimated and recorded the colour of illumination for each trial according to their chosen chromatic statistics (i.e. mean chromaticity, the chromaticity of brightest pixel, or the weighted-mean chromaticity), and averaged these estimates with weightings defined by the corresponding curve in Figure 4.12. Here, a time constant of 1 (red curve) is closest to the single-trial model. The time constant infinity (yellow line) means that the model average all presented trials with equal weightings. Other time constants act as models somewhere between these two extreme cases. Note that the model recorded the illuminant estimate indiscriminately for all conditions (specularity or the presence of surrounding context). Each participant completed 2,304 trials, and thus the model went through all the trials from first trial to the 2304th trial in the same order, and provided an answer for each trial. The model was run separately for each observers, as the order of trials was randomised differently for each observer.

Figure 4.13 indicates the results for these temporal integration models for all chromatic statistics. The rightmost panel indicates the mean values across the four conditions. The black dotted line shows the chance level. The dashed cyan line indicates the null model, and the magenta line indicates averaged performance across 7 human observers. First of all, regarding overall performance (upper panels), as τ increased, the performance of all models, except for brightest model, increased. The mean chromaticity model and the weighted mean chromaticity model (w = 3.0) reached almost perfect performance, and which is far better than human observer's average performance shown by the magenta line.

The lower panels in Figure 4.13 show the d-prime metric for all models. The maximum d-prime value within the condition is shown by the red arrow. For the matte condition with no surrounding context, it was shown that the d-prime of the weighted mean model (w = 3.0) increased as the time constant increased. It reached the highest value at time constant 10^4, which was the maximum value within this condition. The glossy condition with no surrounding context showed a very similar trend. When the surrounding context was available, the degree of improvement due to an increase in time constant was larger.

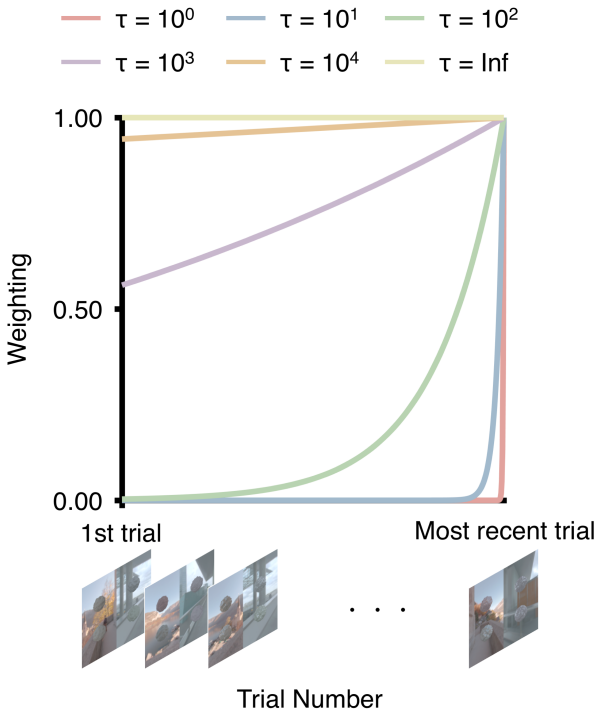


Figure 4.12: A diagram to show the weighting of each trial for the temporal-integration model. We tested six different time constants. A time constant of infinity corresponds to equal weighting for all trials.

For the single trial model (labelled as instant), the d-prime values for the weighted mean (w = 1.0 and w = 3.0) were at most as good as the null model. However, as the time constant increased, the d-prime values for these models considerably increased and reached its highest at 10^4 for the matte condition and 10^3 for the glossy condition. The rightmost panel shows the mean d-prime across all four conditions. We conclude that the weighted mean chromaticity model (w = 1.0 or 3.0) predicted observer's behaviour the best in experiment 1. Figure 4.13 shows results for specific noise values, but we note that findings held for other noise values as shown in the Appendix.

Overall, for the simplistic illuminant estimation model tested here, it was shown that taking the history of chromatic properties into account seems to be key to both solving the task and matching the model to human observer's behaviour. This implies that when the lighting environment is complex, our visual system can grasp the "gut feeling" of the chromatic properties of the surrounding environment by relatively simple spatial and temporal integration of surrounding context.

To further test this finding, we designed Experiment 2 to include a condition in which the useful information about the surrounding lighting environment was not available from a single trial. Thus, in order to perform the task for such a condition, human observers were forced to integrate information presented temporally over trials. To implement such an experimental condition, we utilised the directional asymmetry of the environment. In a lighting environment that is asymmetric such that illumination varies with direction, the lights falling on an object which are then reflected to the eye are very different from the part of the lighting environment that is presented in the surrounding context behind the object. Based on this idea, we selected two environmental illuminations that have high directional asymmetry.

In Experiment 1, we employed 20 possible camera angles which covered a wide range of the scene. Thus, by simply integrating the information over trials indiscriminately, our visual systems were able to find a representative influence of illumination. In Experiment 2, we specifically chose two camera angles aligned with the asymmetry of the illumination.

4.4 Experiment 2

4.4.1 Conditions

In Experiment 2, we again employed eight conditions: the factorial combination of two specular levels (matte and glossy), the presence of surrounding context (SC and No SC) and two congruence conditions (congruent and incongruent). Experiment 2 used smooth stimuli so that the geometry of specular reflections is clearer than with bumpy objects, expecting that this might help human observers use it more easily to estimate the influence of illumination hitting on the object's surfaces. Importantly, there were only two viewing angles for each environmental illumination in this experiment. For congruent and incongruent conditions, the viewing angle was selected such that the cone signal of the mean illumination falling on the object matched or mismatched the mean cone signals from the surrounding context, respectively. The method of selecting viewing angles is detailed in the next subsection. Half of the trials were congruent trials and the other half trials were incongruent trials. These two types of trials were interleaved in

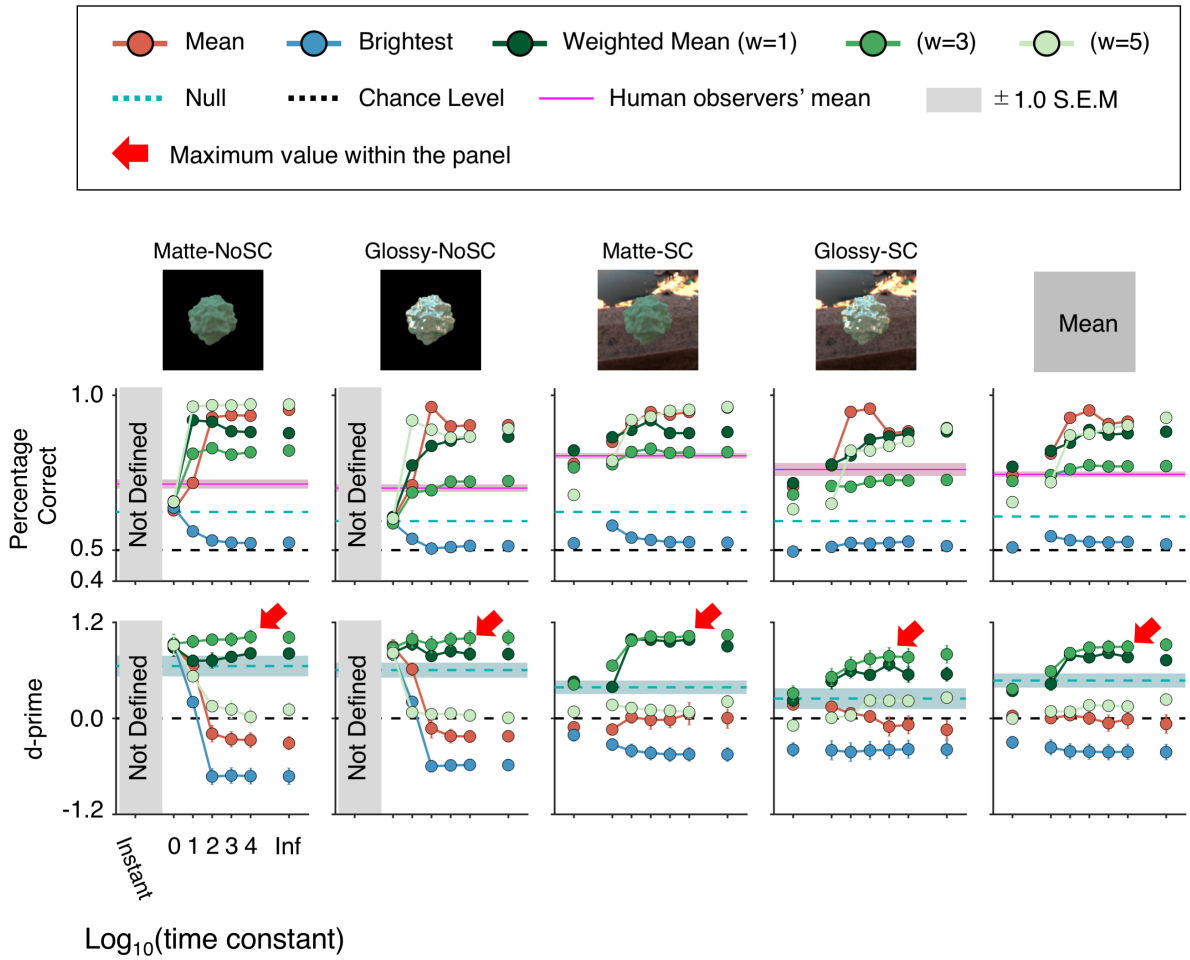


Figure 4.13: Percentage correct values (upper panel) and d-prime values (lower panel) across all conditions as a function of time constant. The results from mean chromaticity, brightest pixel, and weighted mean models are shown by circles, and the null model is shown by the dotted cyan line. The chance level is shown by the black dotted lines. The rightmost panel indicates the mean values across the four conditions. Human observers' mean performance ($\pm S.E.$) is indicated by the magenta line and shading. The red arrows identify the maximum values within each panel.

a random order within each block, and all observers except for TM were not informed which trials were congruent or incongruent. We used only 5 reflectances as shown in Figure 3. The full number of combinations of 5 reflectances (${}_5C_2$) were tested for each condition. Again, the assignment of reflectances to target and distractor objects or vice versa were treated as different trials. Also, a target object being placed under the left environment or the right environment were treated as different. Thus, one block consisted of 80 trials (${}_5C_2 \times 2 \times 2 \times 2$). All seven observers completed 4 conditions (gloss/matte, and SC/No SC), and thus 320 trials per observer were obtained and analysed.

4.4.2 Selection of environmental illumination and viewing angle

To implement congruent and incongruent conditions (where the spatial average of the cone signals from the surrounding context matches or mismatches the cone signals of mean illumination falling on the objects) we first selected two environmental illuminations whose mean cone signals change substantially depending on the viewing angle. Next, we rendered two mirror spheres placed under those two environmental illuminations. A mirror sphere reflects the illumination hitting the surface directly to the camera, and thus it allows us to precisely know the illuminations falling on the objects' surfaces. Viewing angles were changed from 0 degree to 350 degrees with 10 degree steps as shown in panels (a) and (b) of Figure 4.14. Then, the mean cone signals over the surfaces of the two mirror spheres were compared to the mean cone signals of the surrounding context region, and we chose the 2 viewing angles that maximise or minimise those two differences as shown in panel (c) and (d). Here, the magenta and cyan colours indicate the congruent and incongruent conditions, respectively. The mean cone signals across the two mirror objects are shown by the diamond symbols, and the mean cone signals across the surrounding context are shown by the circle symbols. As shown here, for the congruent condition (magenta), the circle and diamond symbols closely overlap indicating that mean cone signals are well matched. In contrast, the incongruent condition exhibits a large chromatic difference between circle and diamond symbols.

Figure 4.15 provides examples of actual stimulus presentations in Experiment 2. The right hand figure in each of panels (a) and (b) qualitatively shows how much the mean cone signals across objects and across surrounding context are separated. Clearly, for (b), the incongruent condition, the mean colour of the surrounding context is strongly biased toward bluish (left environment) and reddish (right environment) due to the presence of clear blue sky and the red items behind the stimuli, and consequently the two surround colours are very different from the mean colour across objects. We were interested in testing whether human observers can perform the task well even for incongruent conditions where the mean cone signals of the surrounding context do not help them to perform the task.

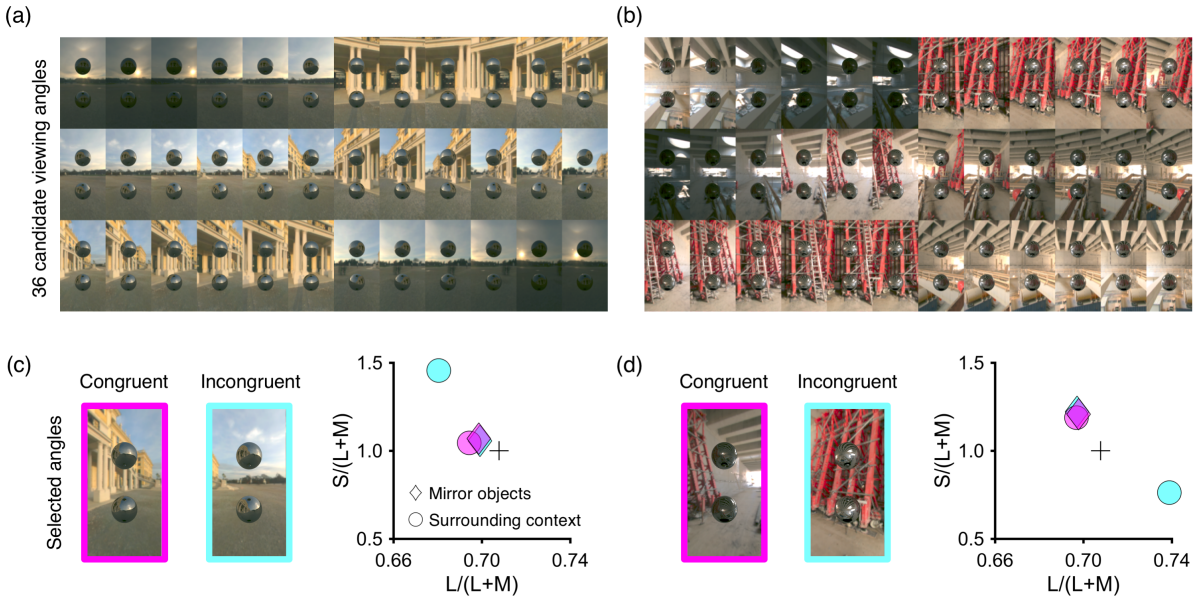


Figure 4.14: Generation of the "congruent" and "incongruent" conditions of Experiment 2. Panels (a) and (b) show environments 3 and 4 and all 36 candidate viewing angles rendered with mirror spheres under each environment. Panels (c) and (d) show the two selected viewing angles that maximise and minimise the difference between mean cone signals across the two mirror objects and the mean cone signals across the surrounding context in environments 3 and 4, respectively. The black cross indicates the chromaticity of equal energy white.

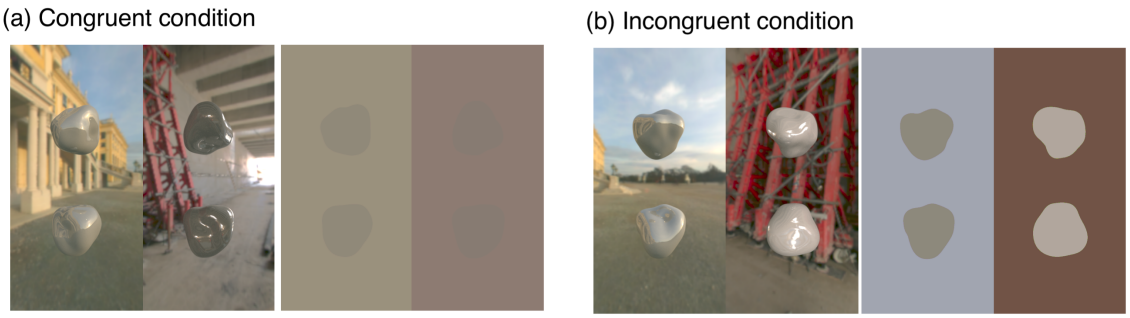


Figure 4.15: Examples of stimulus presentation for Experiment 2. Panel (a) shows a congruent condition where the mean cone signals of the surrounding context match the light falling on objects. Panel (b) shows an incongruent condition where the mean cone signals of the surrounding context differ from the light falling on the objects. The right-hand image in each panel shows the object regions and surrounding context regions filled by the sRGB colour of their mean cone signals.

4.4.3 Results and discussion

The experimental design was such that we expect particular predictions from the model, which we describe here before showing results from human observer. Figure 4.16 shows the results from the single trial model based on mean cone signals (mean chromaticity

weighted by luminance powered by 1.0) across surrounding context for (a) the congruent condition and (b) the incongruent condition. It was shown that for the congruent condition, the model works reasonably well for matte conditions, but performance worsened for the glossy condition. Correcting the influence of illumination led to nearly perfect performances for most trials for both matte and glossy conditions. For the incongruent condition, when there is no surrounding context, for both matte and glossy conditions, the null model does not work so well. For the surrounding context condition, the model performance dropped to chance level for most trials.

Note that these model predictions are as intended from the experimental design, and thus they provide a sanity check for the model implementation. Our actual interest here was to measure human observers' performances for those conditions. Panels (c) and (d) show the average performance across seven human observers. For (c) the congruent condition, similar to experiment 1, there was little difference between matte and glossy conditions, but performance seems to increase in the presence of surrounding context. Importantly, it was clearly shown in panel (d) that observers' performances did not drop much with the absence of the surrounding context, which indicates that human observers seem to be doing more than relying on the mean cone signals of surrounding context to compensate for different environmental illuminations.

As considered in Experiment 1, we implemented other candidate models that use the mean chromaticity, brightest pixel, and other luminance-weighted mean chromaticity (w = 3.0 and 5.0) of the *surrounding context region*, but these models based on single-trial statistics did not explain observers' performance. Also, we confirmed that models based on the mean chromaticity, brightest pixel, luminance-weighted mean chromaticity (w = 1.0, 3.0 and 5.0) of *objects' surfaces* did not provide a good account of observers' performance either.

Therefore, we imagined that the pattern of performance could be again explained by the temporal integration of chromatic information over trials. In other words, if by taking the temporal average of surrounding context, observers can utilise information from congruent trials, which then helps them to access relevant information about the lights hitting objects. Thus, we again implemented the same integration model as Experiment 1.

Figure 4.17 shows the d-prime values for all models as a function of the time constant of integration. This time, in addition to mean chromaticity, brightest pixel, and weighted mean chromaticity models, we implemented a congruent context model which always relies purely on the weighted mean chromaticity (w = 1.0) of the surrounding context presented in congruent conditions, regardless of whether the current trial is congruent or incongruent (shown by orange dotted lines). In a similar manner, we implemented an incongruent context model which relies solely on the weighted mean chromaticity (w = 1.0) of the surrounding context in incongruent trials (shown by blue dotted lines).

For congruent trials, the congruent context model shows the highest d-prime value for any condition. This implies that human observers were not simply integrating chromatic properties across all trials, but instead they relied on the chromatic statistics presented in the congruent trials and ignored the incongruent context. For incongruent trials, for matte objects with and no surrounding context, weighted mean chromaticity (w = 3.0) was the best predictor. Also, for glossy objects with surrounding context, the brightest pixel model is the best predictor. Other than these two conditions, the congruent context model is consistently the best predictor of human observers' performances. This would imply that even when observers are making a judgement for an incongruent trial, they were using cues obtained from congruent trials. Interestingly, the null model was not a good predictor of observers' performance for the congruent conditions, suggesting that observers were indeed performing some correction of the influence of the illumination.

Taken together, these results support the finding in Experiment 1 that observers temporally integrate chromatic signals from the surrounding context. However, rather surprisingly the integration seems to be more selective than just taking a temporal-spatial average of surrounding context across all trials. Instead, they seemed to rely more on congruent viewing angle trials, and to ignore the information presented in incongruent trials. Observers except for TM were not informed which trial belongs to incongruent and congruent conditions, but the trends were similar across observers.

How could observers selectively use the cues presented in congruent conditions? This effect cannot be explained by a simple integration, such as adaptation mechanism which is usually considered to be automatic and passive. Looking at the surrounding context of

environment 4 in panel (b), it is clear that reddish objects dominate the mean cone signals. It is possible that observers decided not to use an average across the whole surrounding context to estimate illumination because the redness clearly stems from the colour of surface rather than illumination. In any case, the results in Experiment 2 suggest that the visual system developed an effective reference to perform the task over trials.

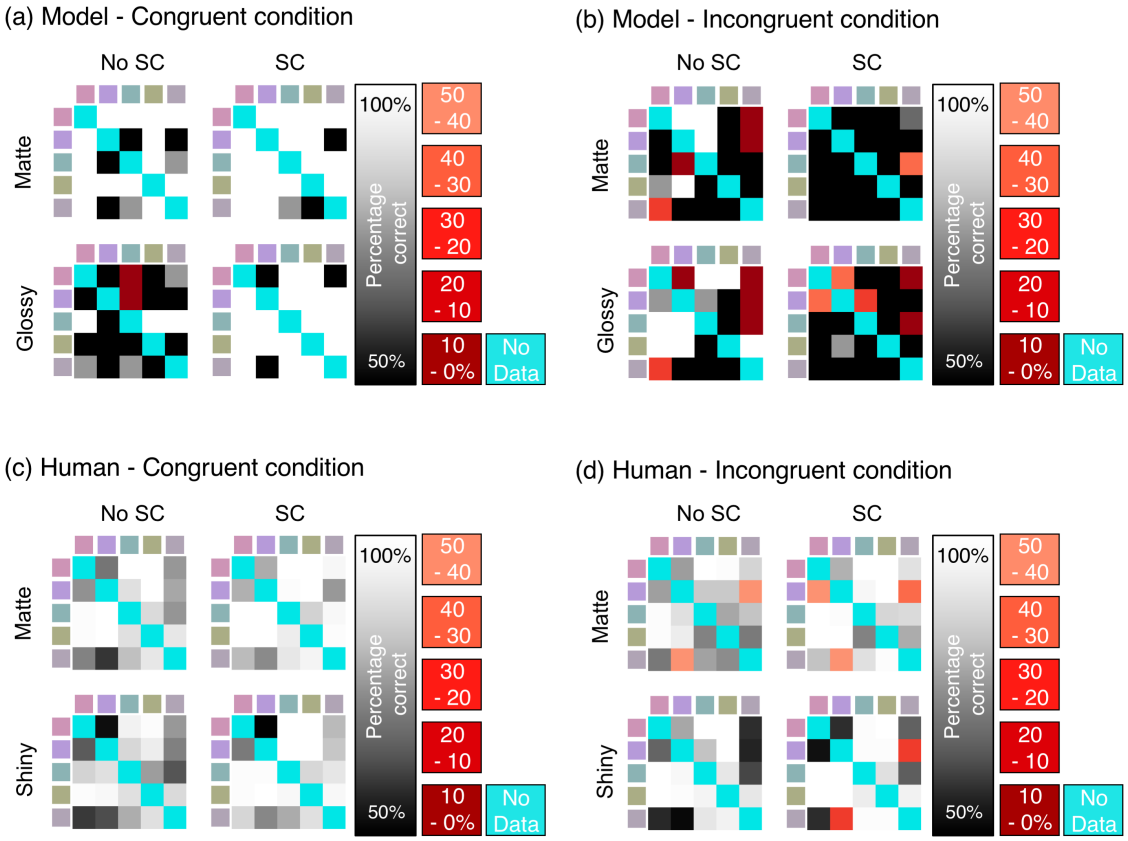


Figure 4.16: Panels (a) and (b) show model performance for congruent and incongruent conditions based on mean cone signals from the current trial. Note this is a single-trial model where past trials do not influence the present estimation of illuminant colour. Panel (c) and (d) show human observers' performance.

4.5 General discussion

The present study aimed to uncover strategies used by human observers to identify surface colours of matte and glossy objects when they are presented across complex lighting environments. We used computer graphics techniques to render objects under environmental illumination and tested human performance in a task that required

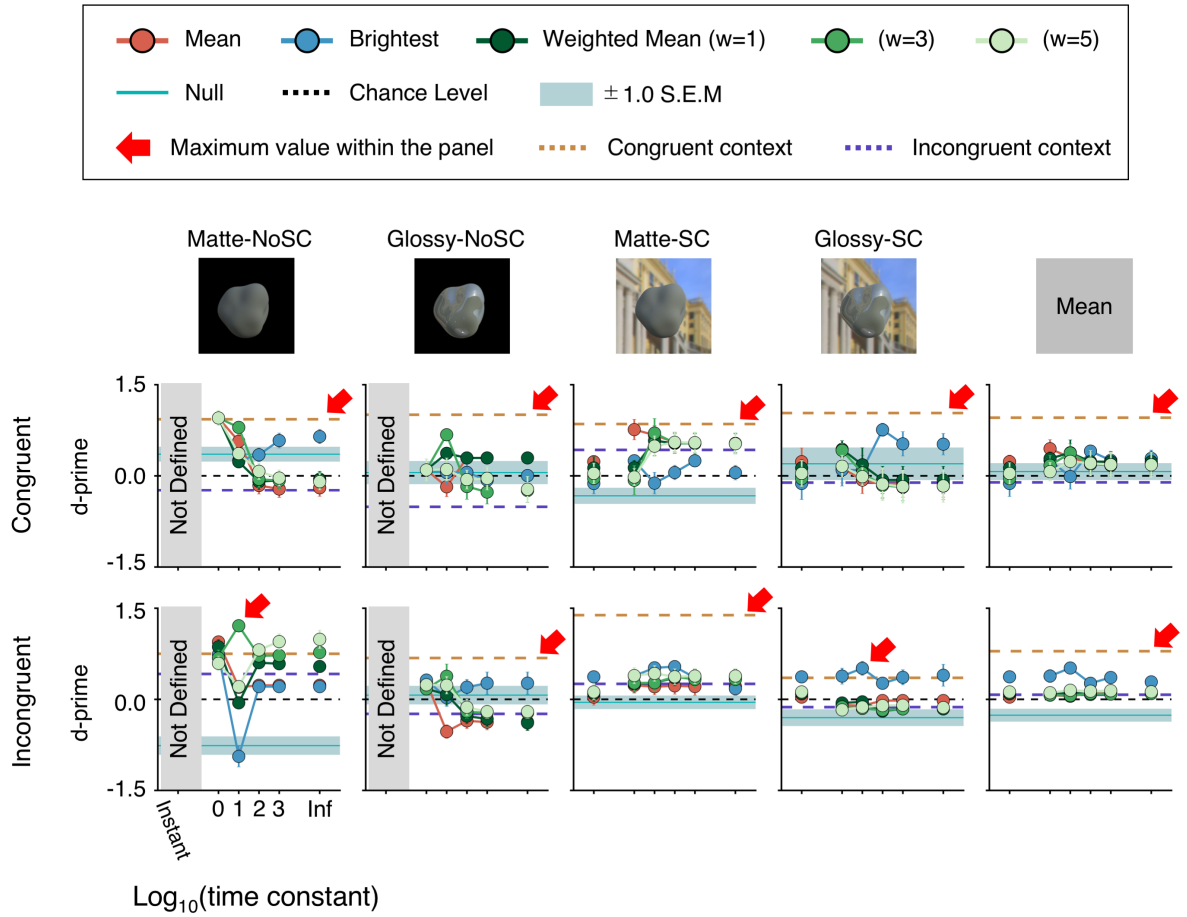


Figure 4.17: d-prime across all time constant. The results from mean chromaticity, brightest pixel, and weighted mean models are shown by circles, and the null model is shown by the solid cyan line. The chance level is shown by black dotted line. Maximum values within the panel is indicated by red arrows. The model which always try to estimate the illuminant colour from congruent condition is shown by orange dotted line, which rather surprisingly showed the highest d-prime values for most conditions.

identification of the target object across different lighting environments. It is not known how observers might estimate and compensate for illumination in an environment in which the spectral content of the illumination is direction dependent, meaning that objects are illuminated by complex light fields. We sought to understand patterns of human responses finding a computational model that responds according to known rules and tested whether that model mimicked human observers' performance. For the class of simple models we tested, in which a single illumination correction is performed, results suggested that a temporal integration model rather than a single-trial model was the best predictor of human observers' performance. This implies that a human observers may cope with the

complexity of illumination using temporal and spatial summation to provide a "best guess" of the chromatic bias in the illumination. Experiment 2 in addition suggested that if such temporal integration is employed by our visual system it seems to be more selective than simple integration across all trials. The computational demands required for human observers to estimate the exact influence of illuminations from every direction may be too expensive. Such a strategy that relies on summary statistics would make sense to deal with the complexity that visual system faces in the real world. However, although our simplistic models reached high levels of performance against the ground-truth, it should be noted that even the best predictor of human observers' performance (i.e. the temporal integration model) did not exceed the d-prime value 1.2, and therefore the models based on single white point correction are limited in their ability to describe human performance. Thus, we emphasise that our study cannot rule out the possibility that the computations driving behaviour are more complex than those in the models.

We decided to employ a selection-based colour constancy task as it is similar to what we do in everyday life. This study is not the first to use such a naturalistic task. For example, [139] showed that selection-based colour constancy could yield a similar degree of colour constancy to a matching-based task, but the degree of colour constancy was much larger for three-dimensional objects than two-dimensional flat simple stimuli. Our study is however the first one to introduce complexity of surface properties (specular reflection) and illuminant properties (environmental illumination) to the experimental stimuli. We found that specular reflection had little effect on performance. Overall, human observers can yield 70.7% percentage correct (average across matte and glossy conditions, and across seven observers) when there is no surrounding context, but their performance rose to 78.2% when contextual information was available.

In the following section, we separately discuss the effect of each experimental manipulation in the present study.

4.5.1 Specular reflection

Specular reflections can in theory provide a reliable cue to perform the current task because they provide reliable information about the lights falling on the object. Indeed,

the colour of highlights on the surface provides more reliable information than using the surrounding context which does not necessarily match with the lights falling on the object. While some studies suggested that specular reflection could be a cue to solve colour constancy [79], [129], [149], the present study found no effect. This is particularly notable in the no surround context conditions. There are at least three points to consider why we did not find a difference in performance for matte and glossy stimuli. First of all, unlike a single-spectrum illumination environment, the specular reflections in our stimuli exhibited a complex spatial pattern and their chromaticity varied dramatically from one point to another. Past studies that found an effect of specular reflection all used point light sources, which create specular highlights of specific geometry on the objects' surfaces. In addition, under point light sources, specular reflection introduces less chromatic variation across objects compared to under environmental illumination. This could be a potential reason why human observers could not benefit from the presence of specular reflection in our present study. When we finished the analysis of Experiment 1, we also suspected that one contributing factor for why we did not find an effect was the complexity of the spatial geometry of specular reflection with bumpy objects. Experiment 2 employed smoother objects, but still this did not produce a positive effect of specular reflection. Thus, the absent effect of specular reflection is unlikely to be due to the clarity of the reflected pattern on objects' surfaces. The third point is related to the fact that experimental tasks of colour constancy range widely, such as achromatic settings, asymmetric matching, illuminant discrimination, identification of surface colour and discrimination between illuminant change and reflectance change. It is possible that the effect of specular reflection varies across the type of task. Some studies that reported the effect of specular reflection employed achromatic settings [79] and discrimination between surface colour change and illumination colour change [129], [149]. Different methodologies might drive different aspects of colour constancy [110]. Because of the presence of large differences in experimental designs, inconsistent effects of specular reflection reported across studies may be not as contradictory as they appear. Alternatively, our visual system might not actually use information from specular reflection to implement colour constancy. It is good to remind ourselves that a theoretically useful cue may not be known or used by

the visual system. In any case, despite several claims regarding the positive effect of specular reflection in past studies, our study showed no effect of specular highlights under complex lighting environments, at least in our current experimental set up.

4.5.2 Temporal integration of cues from surrounding context

Both experiment 1 and 2 showed the temporal integration of chromatic signals from surrounding context is important to predict the pattern of human observers' performance. In practice, our visual system can implement temporal integration strategy simply by moving our head and eyes. However, the results from Experiment 2 and the associated modeling suggests that human observers *selectively* integrate chromatic signals but the way that might be achieved is still an open question. It's possible that the visual system is more spatially selective than implemented in our models, for example by ignoring the some part of the images and averaging the rest of the scene, but it is hard to verify using a computational model simply because there are too many ways to segment images. For example, in the right hand environment of panel (b) in Figure 4.15, the global mean colour of the surrounding context is clearly shifted towards a reddish colour, and thus there might be a signal that a simple spatial summation would give a biased estimation.

Compared to the large amount of literature on spatial integration mechanisms of human colour vision [21], [22], [25]–[27], the temporal mechanisms of human colour constancy have received much less attention [119], [164]. Studies on colour constancy by colour memory are slightly more numerous [84], [202]–[204]. However, even those studies employed simple stimuli, and did not use a selection-based task. The present study provides a first insight into the way human observers integrate the temporal context of surrounding information to implement colour constancy under real-world environments.

The suggestion from the present study that performance in a given trial depends on the history of exposures to previous trials is similar to ones in past related research [87], [164], which showed effects of temporal context in an appearance-based colour constancy task. Importantly, both studies suggested that the visual system might employ an attentionally gated adaptation mechanism because the degree of colour constancy decreased disproportionately to the reduction of presentation time of stimuli

requiring response. The results from our experiment 2 are in line with this observation because modelling temporal and spatial integration across all trials did not predict human observers' performance well. The experimental stimuli used in the previous study were two-dimensional matte surfaces, embedded in a variegated background illuminated by a uniform illumination. Therefore, our study extends their finding to experimental stimuli that are more complex in surface and illuminant properties. Also, another study that measured appearance-based colour constancy supports such a selective adaptation hypothesis [191].

4.5.3 Future directions

In the real world, it is rare for the visual system to experience an environment where there is only a single spectrum illumination. Unless it is a controlled environment such as a laboratory, an object receives both directly emitted lights from multiple light sources and inter-reflections from various directions. Nevertheless, the effect of complex lighting environments on colour constancy is still largely unexplored. Since the complete estimation of environmental illumination seems to be computationally implausible, the way our visual system internally represents its influence is an interesting research question. The present study employed a selection-based constancy task, but stepping back to a simpler methodology such as achromatic setting which allows a more direct measure of the perceived white point under complex lighting environments would be an interesting direction to take. When the illuminations change depending on direction, what sort of cues does our visual system use to determine the influence of illumination? Those ideas could potentially yield a surprising observation that may overwrite widely held views from past colour vision studies.

In this study, we tried to understand human colour constancy mechanisms using computational observers that try to solve a task using specific chromatic signals presented in a trial or trials. More specifically, our approach was to manually select the potentially useful chromatic statistics (e.g. mean chromaticity, brightest pixel, luminance-weighted mean chromaticity), and then to examine how well a model relying on such cues could explain human observers' performance. However, such choice of cues can be arbitrary

and, more importantly, if the visual system uses more sophisticated statistics we cannot predict, we will not be able to reach the strategy upon which human observers rely. We note that there should be many other chromatic statistics that this study could not test, and thus there are possibilities that the results can be explained by image features from a single trial without implementing temporal integration models. We consider this is a bottle-neck of this type of computational approach, and it needs to be refined in the future studies. For example, some behavioural studies have started to use machine learning algorithms to understand the human visual system's behaviours [205]–[209]. Low interpretability of machine learning algorithms is still a big challenge. Nevertheless, the potential of this approach to automatically extract the useful features hidden in images could be a powerful tool to explore the vast parametric space of real images, and potentially enable us to decode the complex behaviour of the human visual system.

Further experiments are needed to get a more comprehensive understanding of human colour vision mechanisms under complex lighting environments, but we believe the present study plays a role in providing insights into the mechanisms of human colour vision in response to complex inputs.

4.6 Appendix

Here we shows the noise dependence of our results. Figure 4.18 shows the case where no noise was introduced to the estimation of surface colour and illuminant colours. Figure 4.19 shows double the noise level compared to the results presented in the main text. In both cases, results are held almost constant. The further increase of noise level, of course eventually leads all models' performance to chance level and to lower d-prime. However, we note that our main claims are held across different noise levels.

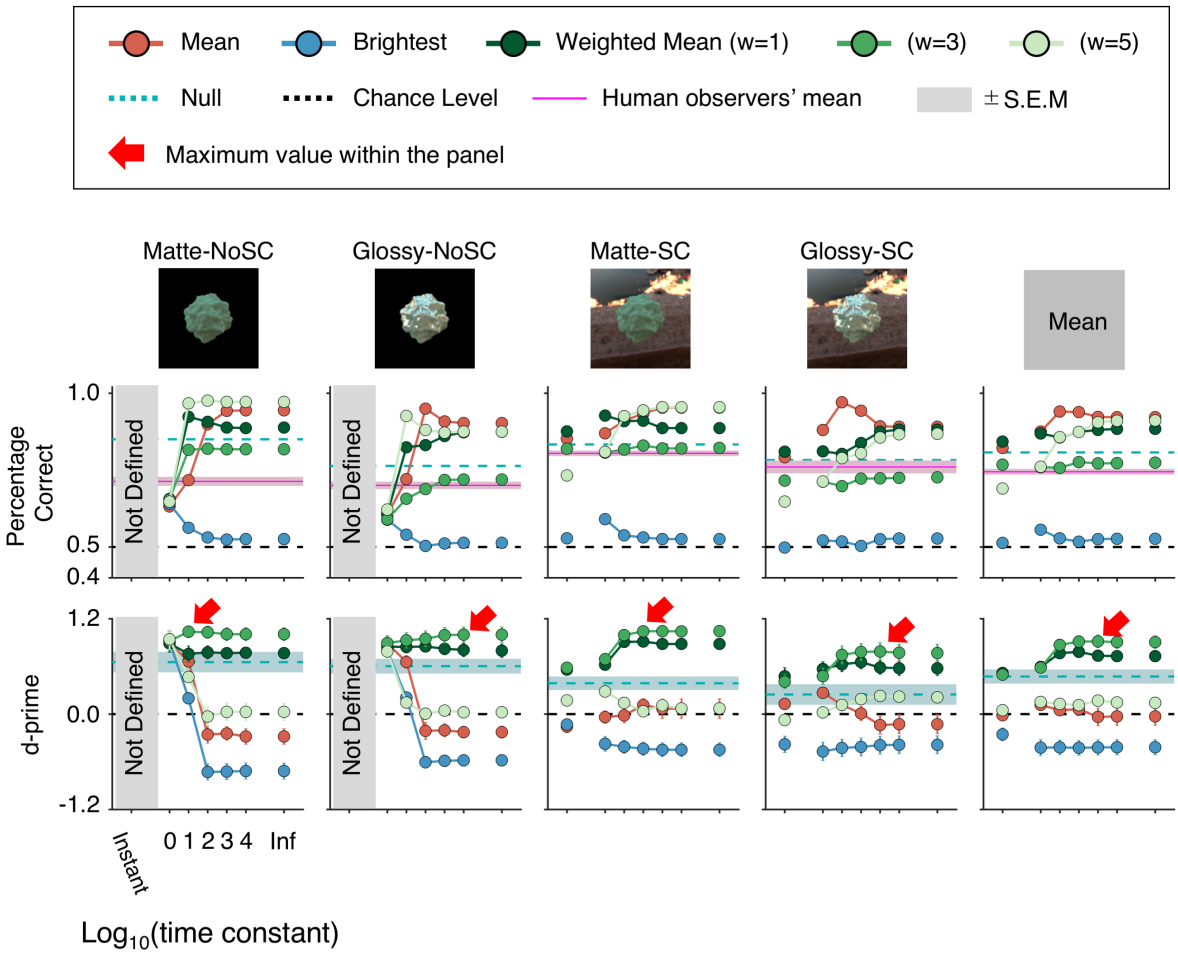


Figure 4.18: Percentage correct values (upper panel) and d-prime (lower panel) across all time constant when the no noise is introduced to the estimation of surface colour and illuminant colour. The results from mean chromaticity, brightest pixel, and weighted mean models are shown by circles, and the null model is shown by the dotted cyan line. The chance level is shown by black dotted line. Human observers' performance is indicated by magenta line. The red arrows show maximum values within the panel.

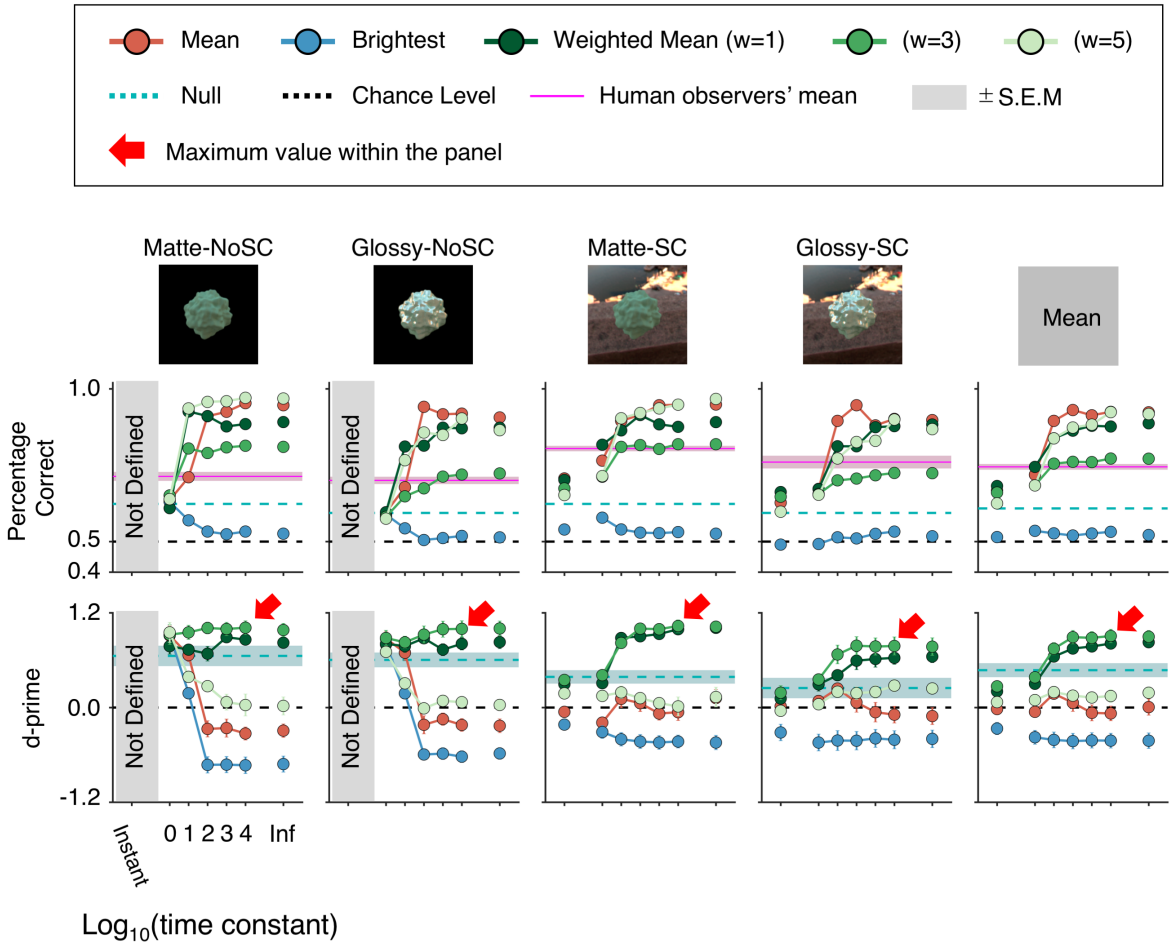


Figure 4.19: Percentage correct values (upper panel) and d-prime (lower panel) across all time constant when double the strength of noise is introduced to the estimation of surface colour and illuminant colour. The results from mean chromaticity, brightest pixel, and weighted mean models are shown by circles, and the null model is shown by the dotted cyan line. The chance level is shown by black dotted line. Human observers' performance is indicated by magenta line. The red arrows show maximum values within the panel.

5

Hyperspectral environmental illumination maps: characterising directional spectral variation in natural environments

5.1 Introduction

The vast majority of objects do not themselves emit light and are therefore visible only by virtue of the incident light they reflect. Characterizing the properties of natural lighting is therefore important in understanding the physical inputs that visual sensors receive in everyday life. The variation in manufactured indoor lighting is relatively accessible (e.g.[210]), but variation in natural daylights requires systematic measurement. In past studies, vigorous efforts have been made to characterize the spectral variation of illumination in natural outdoor environments. The most widely known is the CIE daylight model [211], which accounts for the major variations in the spectral content of natural daylight using only three basis functions. Other studies show similar statistical trends where the chromaticites of natural illuminants cluster closely around the daylight locus [212]. It is also well understood how the spectrum of daylight can change over time and atmospheric conditions [213]. In addition, recent work has expanded these observations to twilight [214]. The focus of these studies has been to characterize the dominant illumination in a scene. Such measurements have significantly advanced our

understanding of the human visual system, especially of mechanisms underlying the perception of objects' surface properties, such as colour and shape. For example, in order to extract information about surface colour, the visual system might exploit the constraint that the illumination in a natural environment is likely daylight [64].

However, the fact that the spectral content of local illumination varies as a function of position in a complex scene suggests a significant influence of nearby objects in reflecting light and thereby modifying its spectral content away from the daylight locus [180]. Indeed, several studies have modelled the effect of inter-reflection on the resultant signals from surfaces and illuminants [215]–[217]. Work has also been undertaken to characterize the *directional* variation in lighting, with particular emphasis on the directional distribution of intensity [170], [218]. The so-called "light from above" heuristic has received a lot of attention as a candidate Bayesian prior to interpret the shape of objects from the variation in shading across their surface [219], [220]. More recent work suggests that visual behaviours are sensitive to the complex directional structure of natural lighting beyond "light from above" [170], [221].

The present study measured the directional variation in spectral content of environmental illumination. Natural illumination mostly originates from two sources: (i) direct emission from a light source and (ii) secondary (or higher-order) reflection from other surfaces in the scene. The spectral shape of light in the secondary reflection is modified by the surface spectral reflectances of the objects. If we imagine an object placed at a specific point in a scene, the spectral content of illumination hitting the object may vary significantly as a function of incident angle. In computer graphics this directional dependency of illumination is approximated using an environmental illumination map [222], which, at each pixel in the map, stores the intensity of a light coming from a particular direction towards a single point in the scene (e.g. an object or an eye). Environmental illumination maps have proved to be important in achieving a realistic appearance when rendering computer-graphics generated images of objects. However, existing databases of environmental illumination maps store only three-channel (RGB) measurements of light, and their use in perceptual science has been limited to characterising intensity variation [167]. Consequently, the spectral variation in directional lighting is largely unexplored.

In this chapter, we report new hyperspectral measurements of environmental illumination, which specify the spectral content of incident light from every direction at a point in a scene. The analysis of the collected datasets revealed that the colours of illuminations are not limited to falling along the daylight locus, and instead can deviate significantly from this locus. It was also clearly shown that there is a striking directional variability in illumination in natural scenes. This variation was observed mainly as a function of elevation angle rather than azimuth angle. For outside scenes, lights from above tend to be intense and dominated by short wavelengths (typically evoking a bluish percept), whereas lights from below are likely to have one log unit lower intensity and lower colour temperature. It was also found that indoor scenes have relatively uniform illumination across angle.

5.2 Method

5.2.1 Hyperspectral imaging system

The hyperspectral imaging system consists of a low-noise Peltier-cooled digital camera that has a spatial resolution of 1344 × 1024 pixels (C4742-95-12ER, Hamamatsu Photonics K.K., Hamamatsu, Japan) equipped with a tunable liquid-crystal filter (VariSpec, model VS-VIS2-10-HC-35-SQ, Cambridge Research & Instrumentation, Inc., MA). Focal length was adjusted before each measurement and the aperture was set to F16 to achieve a large depth of field. The intensity resolution was 12-bits per pixel which is linear with light intensity, confirmed with in-house laboratory measurements. Spectral measurements ranged from 400 nm to 720 nm in 10 nm steps. Each wavelength acquisition required a different exposure time to integrate a sufficient amount of light. Thus, before each image acquisition, we determined the exposure time using an automatic procedure so that the maximum pixel value in the acquired image fits within 85±5 % of the upper limit of dynamic range, per individual wavelength. Wide-field images were obtained by imaging a 3-inch diameter mirror sphere (AISI 52100 chrome steel ball, Simply Bearings Ltd, Lancashire, UK) with the hyperspectral camera, and four such images were stitched together to produce the full 4π steradian panorama. Immediately after an image was

acquired, we measured the reflected light from a flat gray reflectance surface corresponding to Munsell N7, oriented perpendicular to the imaging axis and placed immediately adjacent to the mirror sphere, using a spectrophotomer with traceable calibration (PR-650, Photo Research Inc., Chatsworth, CA) for the purpose of spectral radiance calibration. A more detailed specification of the system can be found elsewhere [223], [224].

5.2.2 Measurement

All measurements were performed in Braga, Portugal. We collected data in eight outdoor scenes and four indoor scenes at Museu Nogueira da Silva and at the University of Minho, Gualtar Campus. Since the present imaging system is the wavelength scanning type, the time taken to sample all wavelengths was relatively long compared to other imaging methods. Although changes in lighting environment were inevitable during the measurement, we minimised the impact by choosing a time when the location of the sun was relatively stable (rather than when the sun is rising or setting). Before starting a measurement, we set up the imaging system as shown in Fig. 5.1(a). The mirror sphere was placed on top of an optical support attached to a tripod. The height between the middle of the sphere and the ground was fixed at 158 cm. The zoom of the lens was maximized to allow the greatest physical distance between the sphere and the imaging apparatus, thereby minimizing the size of the image of the apparatus reflected in the mirror sphere. The hyperspectral camera was positioned 89 cm from the sphere so that the image of the mirror sphere fits just within upper and lower limit of the overall image, as shown in Fig. 5.1(b). The spatial resolution of each sphere region was a 1024 $\times$ 1024. In the optical conditions of acquisition, the system has an acceptance angle of about 6 deg of visual angle and a line-spread function close to a Gaussian with standard deviation approximately 1.3 pixels (around 0.35 arcmin) at 550 nm. Thus, two standard deviations correspond to 0.70 arcmin producing a spatial resolution at least as good as that of the human eye which is around 1.0 arcmin. Before each acquisition, the camera alignment was adjusted to horizontal using a spirit level. To avoid direct sunlight, all outside measurements were performed in shadow. The mirror sphere reflects lights from a wide range of incident angles towards the camera, but there are two issues if we try

to construct an environmental illumination map from only a single image of the sphere. Firstly, the scene behind the sphere is a "blind spot" for the sphere. Secondly, the center of the sphere has the reflection of the imaging apparatus itself. To overcome these issues, four images of the sphere were taken, separated by 90 degrees, as shown in Fig. 5.1(c). The exposure time of each wavelength acquisition was determined by the aforementioned automatic procedure before the measurement and it was kept constant for all angles. Finally, a dark-field image was acquired with the same exposure time for calibration purposes using the hyperspectral camera with the lens cap on. Images were processed, transformed to an equirectangular image, like a world map, and stitched together to create a full panorama image as described in the next subsection.

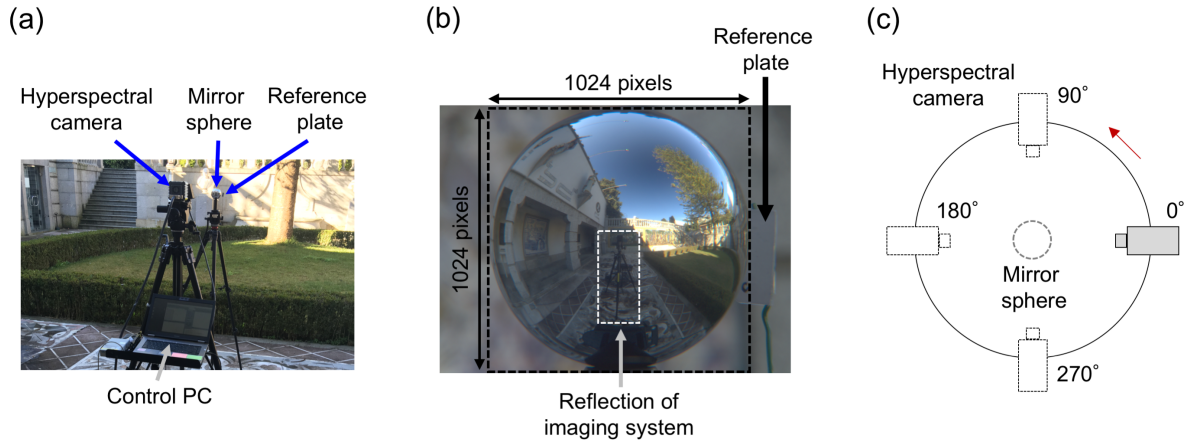


Figure 5.1: How to measure environmental illumination. (a) Set-up of the hyperspectral imaging system. The control PC was placed behind the camera so that we can erase it later along with the tripod. (b) Mirror sphere with a reference plate for calibration purposes. The spatial resolution of the sphere region was 1024 × 1024. (c) Schematic illustration of the measurement geometry. To cover the "blind-spot" of the sphere and to erase the reflection of the imaging apparatus, measurements were taken from four different angles.

5.2.3 Image Processing

Figure 5.2 shows the image-processing pipeline to produce an environmental illumination map from a set of mirror-sphere images. First, (i) each raw image $I_{Raw}(x,y,\lambda)$ of the sphere was corrected for average dark noise in the imaging system using spectral images obtained with the camera occluded, and for spatial in-homogeneity due to off-axis vignetting using spectral images obtained with a uniform transmittance diffuser in front of the lenses and

illuminated with spatially uniform light. Let the dark field image be $I_{DF}(x,\ y,\lambda)$ and the flat field image be $I_{FF}(x,y,\lambda)$. The corrected image $I_1(x,y,\lambda)$ is given by Eq.(5.1).

$$I_1(x,y,\lambda) = \frac{I_{Raw}(x,y,\lambda) - I_{DF}(x,y,\lambda)}{I_{FF}(x,y,\lambda) - I_{DF}(x,y,\lambda)} \tag{5.1}$$

At this stage, the image has arbitrary units. Thus, secondly, (ii) we further calibrated the image so that each pixel has units of spectral radiance (W·sr/m^2/nm). Our approach here is to find a spatially-global, wavelength-dependent scaling function that equates the spectrum at the gray reference plate recorded by the hyperspectral image to the spectrum that was independently measured in spectral radiance by a spectrophotomter PR650. This procedure is independent of any assumptions about the geometry of the light from the scene. Let (u,v) be the region on the gray reference plate, a spectrum reflected back from this surface measured by PR650 be $S(u,v,\lambda)$ and spectra at the corresponding region in the hyperspectral image be $I_1(u,v,\lambda)$. The desired scaling function $w_{SRC}(\lambda)$ is given by Eq.(5.2).

$$w_{SRC}(\lambda) = \frac{S(u,v,\lambda)}{I_1(u,v,\lambda)} \tag{5.2}$$

Then, we scaled whole hyperspectral image using this factor as shown in Eq.(5.3).

$$I_2(x,y,\lambda) = I_1(x,y,\lambda) \cdot w_{SRC}(\lambda) \tag{5.3}$$

Thirdly, as shown in Eq.(5.4), (iii) any saturated pixels were removed from the image using a binary mask $M(x,y)$, where saturated and non-saturated pixels have zero and one (so non-saturated pixels stay as they are) at any wavelength, respectively. After the calculation, the resultant blank pixels were filled-in using linear interpolation of horizontal nearest pixels because our data suggested spectral shapes are more similar in the horizontal direction (seen later in Results and discussion).

$$I_3(x,y,\lambda) = fillin\left\{I_2(x,y,\lambda) \cdot M(x,y)\right\} \tag{5.4}$$

Fourth, (iv) the wavelength-dependent reflection function of the mirror sphere $SR(\lambda)$ measured by spectrophotometer (UV-3600 Plus UV-VIS-NIR, Shimadzu, Kyoto, Japan) was corrected by Eq.(5.5).

$$I_4(x, y, \lambda) = I_3(x, y, \lambda)/SR(\lambda) \tag{5.5}$$

Fifth, (v) since regions close to the edge of the mirror sphere receive (and integrate) lights from a wider range of angles in the scene, this intensity compression was corrected by a cosine function of the angle θ of the incoming lights towards the camera from a specific point on the sphere $w_I(x,y)$ as in Eq.(5.6).

$$I_5(x, y, \lambda) = I_4(x, y, \lambda) \cdot w_I(x, y) \tag{5.6}$$

Sixth, (vi) for each pixel independently, we smoothed the spectrum using a method that preserves the peak based on a locally weighted quadratic regression with span of 25% that was optimised to minimise information loss using a leave-one-out cross-validation method [225] as in Eq.(5.7). We chose a method that can be applied to individual spectra as opposed to a method based on spatial averaging to preserve spatial resolution. The main source of noise is presumably residual dark noise. Equation (5.1) only subtracts the average level of dark noise, and a random noise around the average remains. There are other sources of noise such as photon noise (noise in the light itself) or cosmic ray noise (due to the camera being exposed for so long).

$$I_6(x, y, \lambda) = smooth\,\{I_5(x, y, \lambda)\} \tag{5.7}$$

Seventh, (vii) each sphere image was converted to equirectangular coordinates [226] as in Eq.(5.8). This is a transformation in which the horizontal coordinate represents the azimuth and the vertical coordinate indicates the elevation. The elevation zero corresponds to the height of the camera.

$$I_6(x, y, \lambda) \rightarrow I_7(\phi, \theta, \lambda) \tag{5.8}$$

Eighth, (viii) image registration was performed using the image from the 560 nm channel as a reference to correct for chromatic aberration as in Eq.(5.9). The affine transformation was determined to optimise the alignment, quantified by a cross correlation [227].

$$I_8(\phi, \theta, \lambda) = registration\left\{I_7(\phi, \theta, \lambda)\right\} \tag{5.9}$$

Ninth, (ix) we removed the reflection of the imaging system using an image from another angle $I_{8_2}(\phi, \theta, \lambda)$ and a gaussian window $Gaussian(\phi, \theta)$ as in Eq.(5.10). Here, the centre coordinate and the standard deviation (width) of the gaussian function were carefully selected for each image by hand so that the imaging system was eliminated from the final image.

$$I_9(\phi, \theta, \lambda) = I_8(\phi, \theta, \lambda)\left\{1 - Gaussian(\phi, \theta)\right\} + I_{8_2}(\phi, \theta, \lambda)Gaussian(\phi, \theta) \tag{5.10}$$

Finally, tenth, (x) four sphere images were stitched together based on the SIFT automatic feature detection algorithm [228] to produce the 4π steradian full-panorama environmental illumination map. Image features were found based on the V(λ)-weighted intensity image, and the same stitching transformation was applied to all wavelengths independently.

Parts of the methods described here are analogous to those described in a recent tutorial paper [229].

5.3 Results and Discussion

Figure 5.3 shows the sRGB representation of the collected hyperspectral environmental illumination maps for eight outdoor scenes and four indoor scenes (Fig. 5.3(a) and 5.3(b) respectively). Figure 5.3(c) and 5.3(d) show the corresponding chromatic distributions of these maps expressed in the MacLeod-Boynton chromaticity diagram [17]. The horizontal axis is L/(L+M), which expresses the ratio of signals in the long and middle-wavelength sensitive cones, while the vertical axis indicates S/(L+M), which expresses the signal in the short-wavelength cones in relation to a combination of long and middle-wavelength sensitive cone signals. We used the Stockman and Sharpe cone fundamentals [18] to

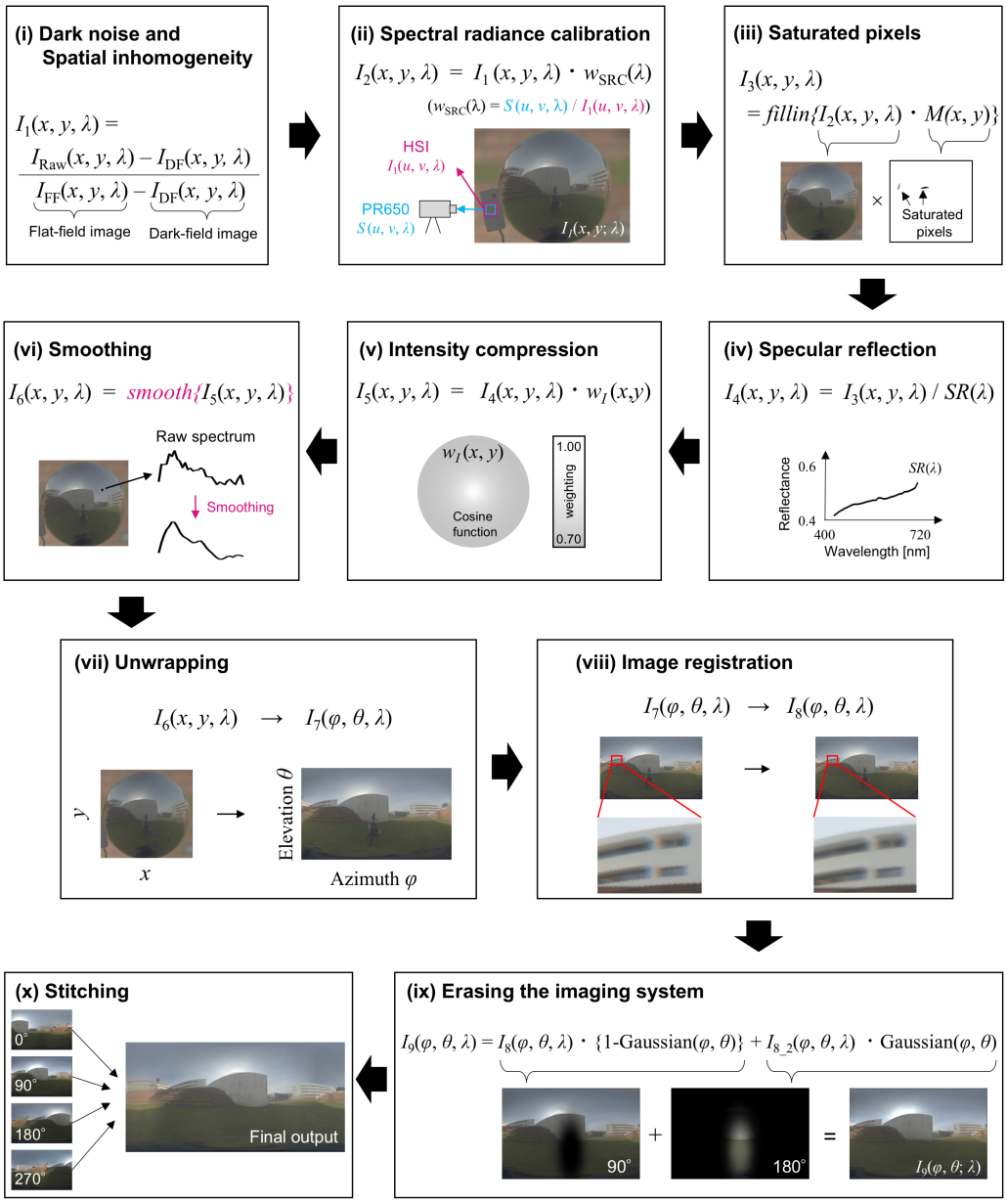


Figure 5.2: The image-processing pipeline. (i) A dark field image was measured with the imaging system with the lens cap on, and was subtracted from a raw image. Also, spatial inhomogeneites of the imaging system were corrected. (ii) Spectral radiance calibration was performed by equating the measurement from the hyperspectral imaging system to the reference measurement from the PR650 spectrophotometer. (iii) Any saturated pixels were removed, and filled-in using linear interpolation of horizontal nearest pixels. (iv) Specular reflection of the mirror sphere was measured by spectrophotometer and corrected. (v) The edge of the sphere receives and integrates lights from a wider region of the surrounding scene. This intensity compression was corrected using a cosine function of radius. (vi) Smoothing based on locally-weighted quadratic regression was performed for each pixel independently. (vii) The image coordinates were transformed to equirectangular projection. (viii) Image registration was performed to minimize the effects of chromatic aberration. (ix) The tripod was erased from each sphere image using an image from another angle. (x) Finally, four images taken from different angles were stitched together based on an automatic feature-detection algorithm to create a full-panorama hyperspectral environmental illumination map.

calculate LMS cone excitations. It is clear that the chromaticities of the illumination spectra distribute around the daylight locus, but that they also deviate substantially from the locus. This was consistent for both outdoor and indoor scenes. Thus, natural lighting is not only limited to the daylight locus, but instead includes a wider range of chromaticities. The white cross indicates the chromaticity of the mean spectrum across all pixels. For outdoor scenes, it has generally lower L/(L+M) and higher S/(L+M) in relation to the chromaticity of equal energy white shown by the intersection of gray dotted lines. For indoor scenes, the mean chromaticity is slightly shifted toward lower values of S/(L+M).

More importantly, Fig. 5.4 shows that there is a variation in spectral shape and intensity level depending on incident angle. Each plot on the right-hand panel indicates the mean spectrum across all pixels within the corresponding rectangle depicted in the left-hand panel. The background colour of each plot indicates the log luminance level. For the campus scene (Fig. 5.4(a)), lights from above ($\theta > 0$) tend to have high intensities in the order of thousands cd/m^2 and high energy around the short-wavelength region. In contrast, lights reflected back from the ground and other objects, which are dominated by secondary (or higher-order) reflections, tend to have one log unit lower luminance and lower correlated colour temperature. In this scene, lights from below seem to have high energy at long-wavelengths presumably because vegetation tends to have high reflectance in the near infrared [230]. The variation across azimuth angles (ϕ) is much smaller. For the canteen scene (Fig. 5.4(b)) there is substantially less variation in intensity as a function of direction. Its spectral shape shows signatures of the daylight spectrum and of the fluorescent lighting, additionally modulated by inter-reflections. Collectively the measured datasets show that directional spectral variation seems to be highly systematic, and different between outdoor and indoor scenes. Also, these data clearly show that the "light from above" heuristic does not hold universally in natural environments because of the presence of secondary reflections. In summary, the obtained datasets revealed large directional spectral variation in real-world visual environments, though the amount of directional variation depends on the scene, as shown in the Appendix.

The black dotted curve in each plot represents the spectrum reconstructed using tristimulus values of the spectrum and the daylight-basis assumption (see Eqs.(2a), (4a),

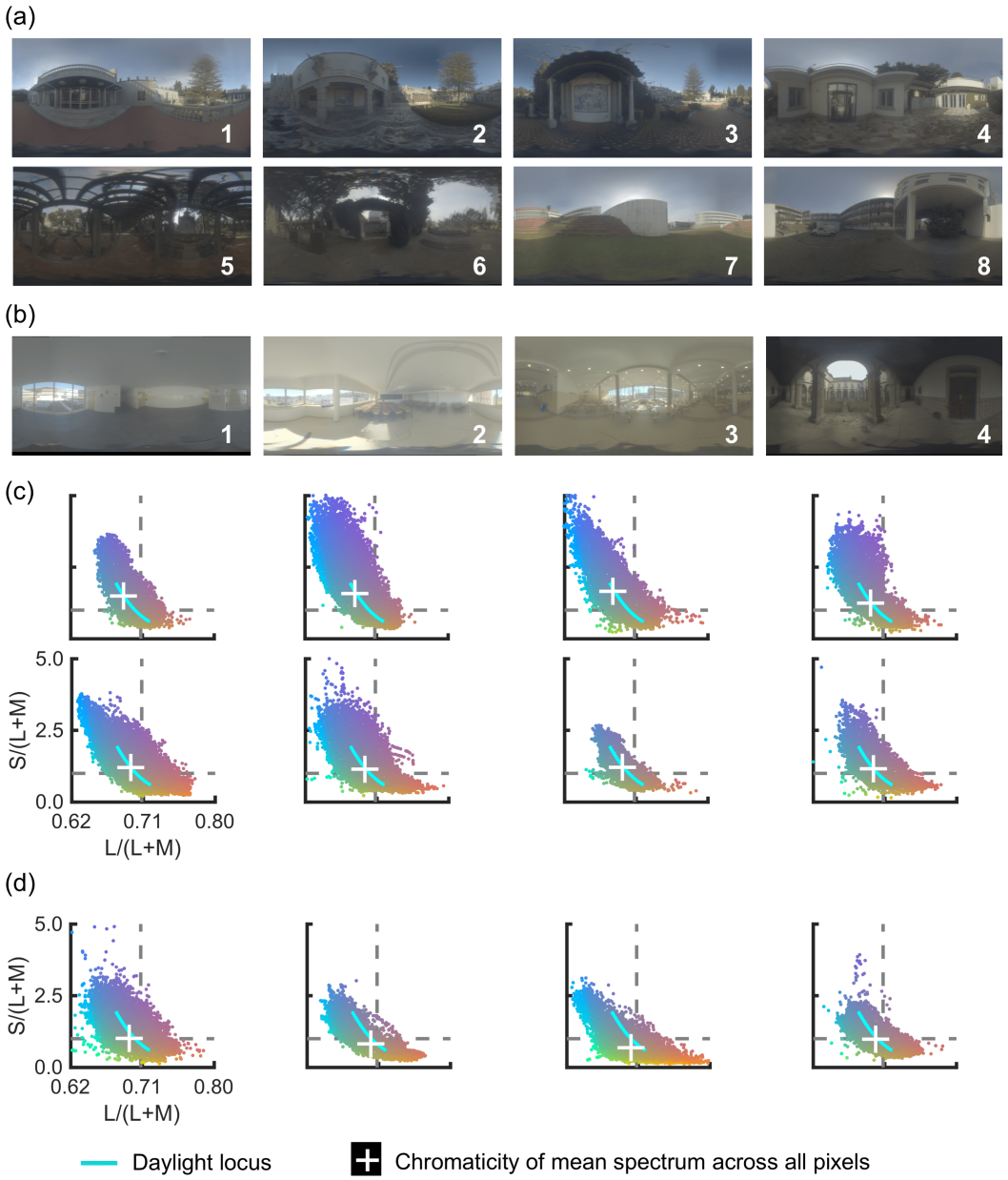


Figure 5.3: sRGB images of hyperspectral environmental illumination at (a) 8 outdoor scenes and (b) 4 indoor scenes. (c) and (d) show the corresponding chromatic distributions in the MacLeod-Boynton chromaticity diagram for each scene. The cyan line indicates the daylight locus. The gray dotted lines show the chromaticity of equal energy white.

(5a), [211]). Strictly speaking, these Eqs. can be applied only to chromaticities on the daylight locus, but here we attempt to test the extent to which a linear weighted sum of daylight basis functions accounts for observed spectral variation. The reconstructed spectrum based on the daylight assumption was normalized so that it had an equal area under the curve as that of the measured spectrum. It is clear for the campus scene Fig.

5.4(a) that the reconstructed spectra are not well matched to the real spectra for lower regions of the image, corresponding to low elevation angles. As expected, this shows that not all natural illuminant spectra can be recovered from daylight basis functions in a scene where secondary reflection exists. For the canteen scene Fig. 5.4(b), spectral shapes are more uniform across incident angles. Overall, spectral shapes rather resemble the CIE daylight function, presumably because the artificial overhead lighting from fluorescent tubes was mixed with daylight entering the room from the large window. It is notable that the regions oriented towards the large window are one to two log units higher in luminance.

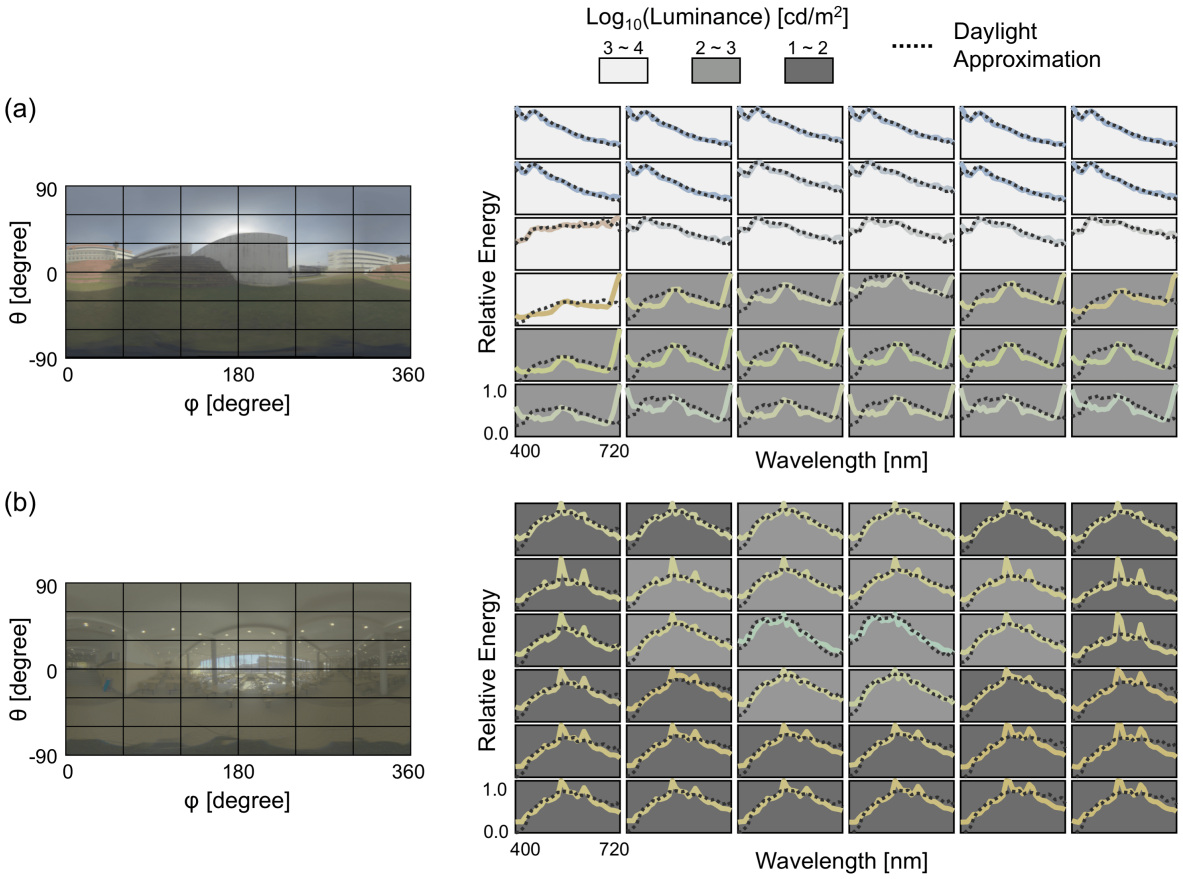


Figure 5.4: Directional spectral variation as a function of azimuth ϕ and elevation θ. (a) Example of one outdoor scene (campus). Each plot on the right-hand panel represents the mean spectrum within the corresponding rectangle drawn on the left-hand image. The colour of each curve is the sRGB colour of the spectrum. The background colour represents the log luminance level. The black dotted line indicates the spectrum derived from CIE daylight basis functions. (b) The case for one indoor scene (canteen). See Appendix for 10 other scenes.

It should be noted this trend depended on the scene, as shown in the Appendix. To

see whether the similarity between observed spectra and reconstructed spectra by CIE daylight basis functions differs between the upper hemisphere (where directly emitted lights dominate) and the lower hemisphere (where the secondary reflection dominates), we calculated the root mean square error (RMSE) between the measured spectra and the reconstructed spectra for all panels in Fig. 5.4. Note that in Fig. 5.4, all spectra were scaled so that the maximum value of the spectrum becomes one within each panel for the sake of visibility. However, when we calculate RMSE, all spectra used for calculation were scaled so that they have the same area under the curve. This normalization allows us to compare spectral shape whilst ignoring intensity differences. Due to this normalization procedure, the units of these RMSE values were arbitrary. The RMSE values were separately averaged across the upper 18 panels ($\theta >= 0$) and the lower 18 panels ($\theta < 0$) for each scene. Figure 5.5 shows the mean RMSE values across upper panels and lower panels for (a) 8 outdoor scenes and (b) 4 indoor scenes.

A two-way repeated-measures analysis of variance (ANOVA) was performed for (a) outdoor scenes with scene (scene 1 to scene 8), and hemisphere (upper and lower) as within-subject factors for the RMSE values. For (a) outdoor scenes, the main effect of hemisphere, scene and the interaction between two factors were all significant ($F(1,17) = 512.5$, $\eta^2 = 0.968$, $p < .00001$; $F(7,119) = 40.5$, $\eta^2 = 0.704$, $p < .00001$; $F(7,119) = 52.6$, $\eta^2 = 0.756$, $p < .00001$). The analysis of interaction revealed that for scenes 1, 5, 6, 7 and 8, the lower parts showed significantly higher RMSE values ($F(1,17) = 26.5$, $p = .000081$; $F(1,17) = 43.4$, $p < .00001$; $F(1,17) = 9.34$, $p = .00715$; $F(1,17) = 214.3$, $p < .00001$; $F(1,17) = 52.8$, $p < .00001$, respectively). In contrast, scenes 2, 3, 4 did not show significant differences between the upper and lower hemispheres ($F(1,17) = 3.68$, $p = .072$; $F(1,17) = 2.16$, $p = .160$; $F(1,17) = 3.38$, $p = .084$, respectively).

For (b) indoor scenes, the main effect of hemisphere was not significant ($F(1,17) = 1.35$, $\eta^2 = 0.282$, $p = .261$). However, the main effect of scene was significant ($F(3,51) = 73.4$, $\eta^2 = 2.08$, $p < .00001$). The interaction between the two factors was not significant ($F(3,51) = 0.59$, $\eta^2 = 0.186$, $p = .624$). Thus, for indoor scenes, the dissimilarity of the measured spectra to that recovered from the daylight basis functions is similar between upper and lower hemispheres.

Overall, for outside scenes, secondary reflection alters the spectrum shape substantially, and thus the daylight assumption does not adequately account for the observed spectral variation. In contrast, for indoor scenes, dissimilarity (RMSE) was small overall compared to outdoor scenes, and the difference in applicability of the daylight model between upper and lower hemispheres was essentially absent.

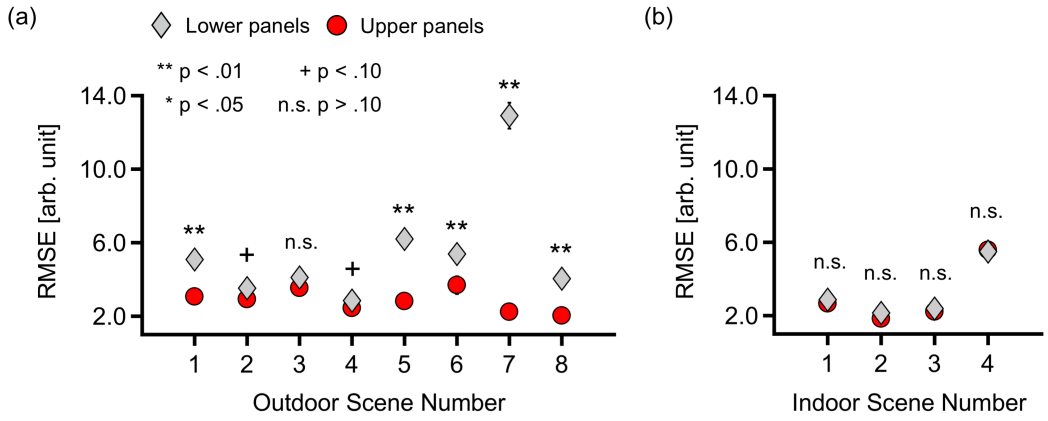


Figure 5.5: Average RMSE between the measured spectra and reconstructed spectra based on the CIE daylight assumption, across the upper 18 panels (red circle symbols) and the lower 18 panels (gray diamond symbols). The panels (a) and (b) show 8 outdoor scenes and 4 indoor scenes, respectively. The p-values to show whether the upper part and lower part have significantly different RMSE values are indicated by signs above each symbol. The error bars indicate ± S.E. but their lengths were mostly smaller than the symbols.

Figure 5.6 shows how (a) cone signals, (b) opponent signals (L/(L+M), S/(L+M) and luminance) and (c) correlated colour temperature, shift as a function of the angles of elevation and azimuth. For example, in the plot of L cone values in panel (a), the data point labelled 1 indicates the average L cone signal across all pixels within rectangle 1 (shown at the right hand side). Circle symbols show the average across 8 outdoor scenes and diamond symbols show average across 4 indoor scenes. Error bars indicate standard error.

When the elevation angle changes from high to low (upper row), the L cone signal decreases about 2 log units in outdoor scenes. For indoor scenes, the change was smaller across elevation angle. We found similar trends for M cone and S cone signals.

In contrast, when azimuth angle changes (lower row), L cone signals did not show a systematic shift for outdoor scenes. Overall level of L cone excitation was lower for indoor scenes. Again, the trends were similar to other cone classes.

The analysis of opponent signals (shown in panel(b) indicates that the luminance channel exhibits a similar trend to the cone signals. For the L/(L+M) signal, when elevation changes, indoor scenes did not show much change from around the chromaticity of equal energy white indicated by the pink dashed horizontal line. However, outdoor scenes changed from greenish region (lower L/(L+M)) to the chromaticity of equal energy white. We do not see much shift of L/(L+M) signals as a function of azimuth change although the overall L/(L+M) level was lower for outdoor scenes than for indoor scenes. Regarding the S/(L+M) signal in outdoor scenes, S/(L+M) drops significantly as elevation angle lowered, from bluish to neutral around equal energy white. Indoor scenes were relatively constant across elevation change. Similar to L/(L+M), S/(L+M) signals did not change much across azimuth angle for both outdoor and indoor scenes, but again overall values were well separated.

Finally, the shift of correlated colour temperature (CCT) across changes in elevation angle also confirms that lights coming from above in outdoor scenes have substantially higher colour temperature. In contrast, indoor scenes show constantly lower CCT. CCT was also relatively constant across azimuth angle, but outdoor scenes had systematically higher CCT than indoor scenes.

To compare the uniformity of spectral shape across direction, Fig. 5.7 shows heat maps in which each pixel indicates the RMSE between the measured spectrum and the mean spectrum across all pixels in that scene. Values at the top left of each image indicate the averaged RMSE across all pixels. If the scene had the same spectrum at every pixel, the heatmap would be uniformly black. Note that the mean spectrum across all pixels was calculated based on raw data, and thus it was dominated by intense spectra. However, again, when we calculated the RMSE, all spectra were normalized to have equal area so that we can compare the dissimilarity of spectral shape whilst ignoring intensity differences. Overall, the mean RMSE seems to be smaller for indoor scenes than for outdoor scenes. The mean of averaged RMSE values (shown at the top left in each panel) across 8 outdoor scenes was 7.95 and that for indoor scenes was 5.49, which was shown to be significantly different ($p = 0.0121$, $t(10) = 3.06$, t-test). Therefore, the indoor scenes have rather more uniform directional structure of spectral shapes across all

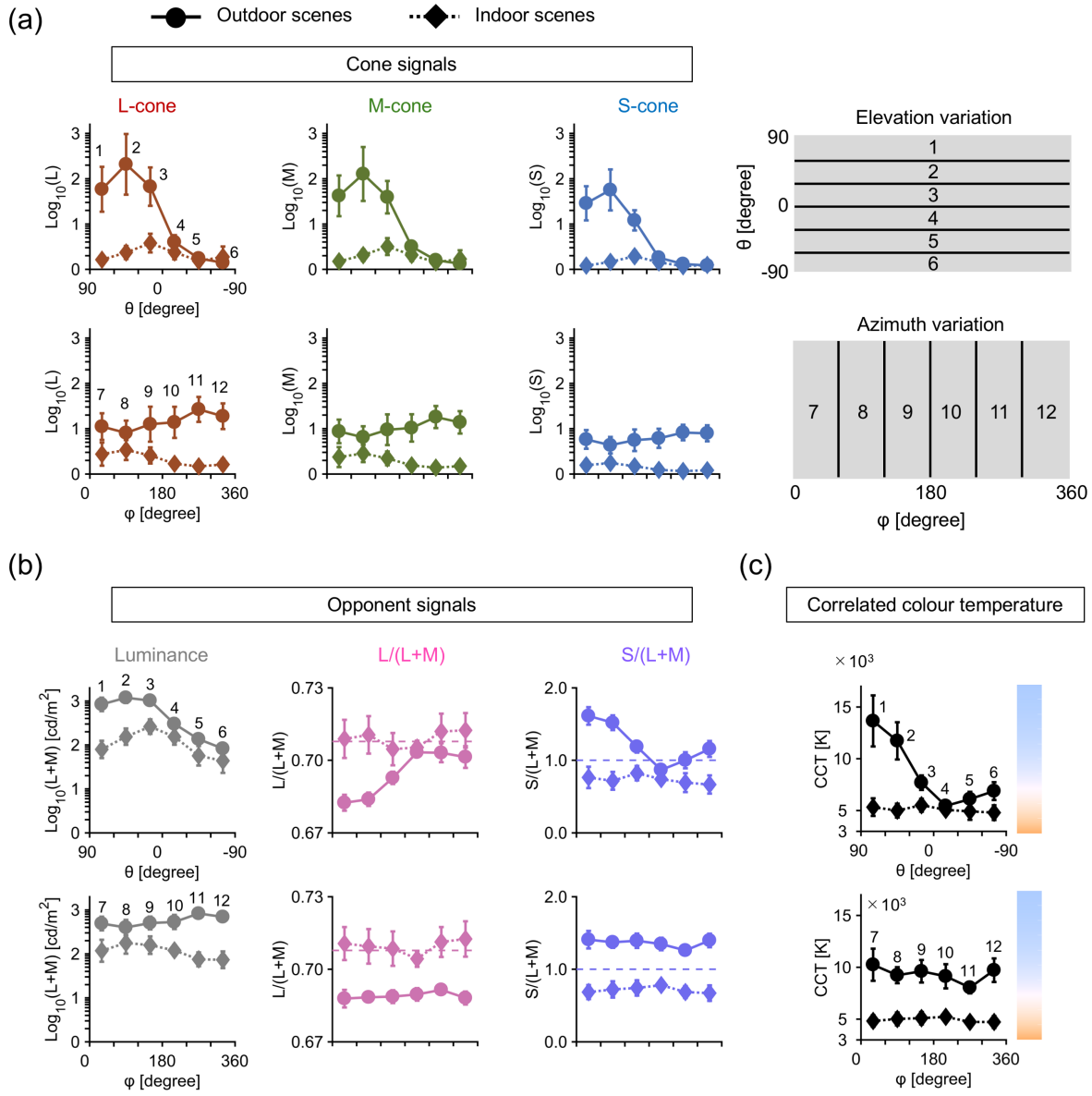


Figure 5.6: Shift of (a) cone signals, (b) opponent signals, and (c) correlated colour temperature as a function of elevation θ and azimuth ϕ. Each point indicates the average across all pixels within the corresponding rectangle region indicated on the environmental illumination map, which is shown at the right had side. Circle symbols show the average across 8 outdoor scenes, and diamond symbols show the average across 4 indoor scenes. The error bars indicate ± S.E.. In the plots of L/(L+M) and S/(L+M), the chromaticity of equal energy white is shown by pink and blue horizontal dash lines, respectively.

directions than outdoor scenes. We also note that calculating RMSE of two normalized spectral distributions by their norm is not the only way to quantify the shape difference. However, we confirmed that a different metric - the Pearson correlation coefficient based on 33 values from 400 nm to 720 nm with 10 steps between measured spectrum and

reconstructed spectrum based on CIE daylight functions - leads to similar conclusions.

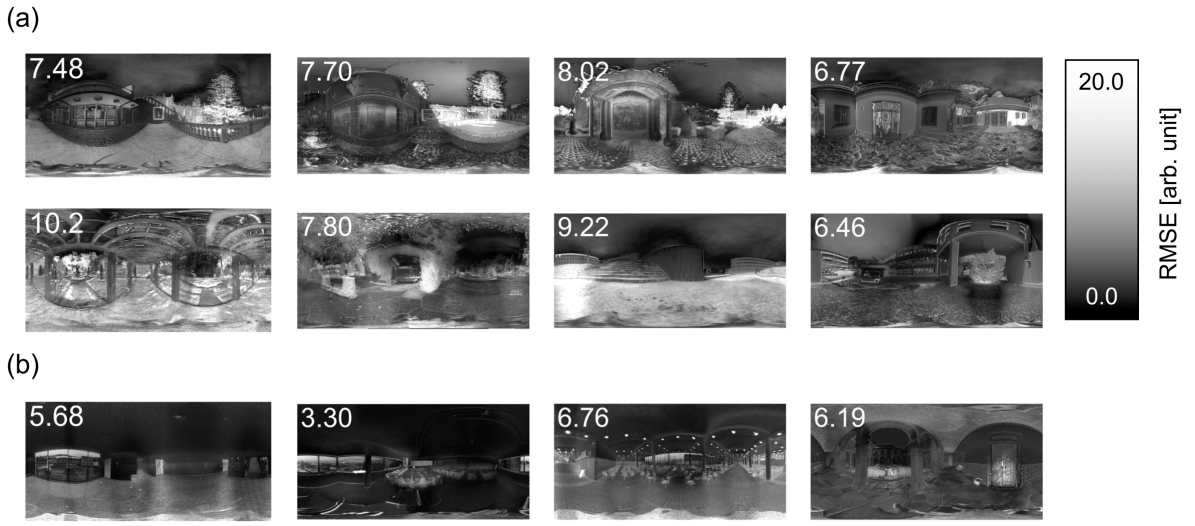


Figure 5.7: Heatmaps to show the spectral variation for (a) outdoor scenes and (b) indoor scenes. Each pixel indicates the RMSE between measured spectrum and the mean spectrum across all pixels. The units are arbitrary. Note spectra in all pixels and the mean spectra were normalized to have equal area to allow us to compare the spectral shape. This calculation was performed independently for each scene so that the overall luminance differences within each scene did not affect the calculation. The number at the top left of each image indicates the mean RMSE value across all pixels.

Additionally, we characterize the spatial structure of the measured environmental illumination maps, to understand the variation in spatial scale from one scene to another, and also from one spectral channel to another. In other words, our intention here is to quantify how finely the intensity of a specific wavelength channel changes depending on angular direction. For example, short-wavelength variation might be spatially low-pass while long-wavelength variation has higher power in high frequency components. For 2D image planes, the power spectrum of the 2-D Fourier transformation would answer this question, but since environmental illumination maps are specified in spherical coordinates we used a spherical-harmonic decomposition. Figure 5.8(a) shows a schematic illustration. The right-hand part of the panel shows the series of spectral images that comprise the original hyperspectral image (left-hand side). The basic concept of spherical-harmonic decomposition is analogous to that of the Fourier transformation. Instead of decomposing a signal to sine and cosine functions as in Fourier transformation, spherical-harmonic decomposition uses different basis functions defined in spherical coordinates known as

legendre functions. Spherical harmonics adopt two parameters, namely degree and order, that control the frequency and alignment of the legendre basis functions. A detailed explanation is documented elsewhere (e.g. [231]). In Fig. 5.8(b) and 5.8(c) we show the power spectra (i.e. the absolute values of the spherical harmonic coefficients) as a function of degree l. Coefficients were summed across orders, and each curve was normalized independently so that the highest point within the curve becomes zero. Each point was connected by spline interpolation. For all scenes, the power is highest at zero or small degrees, and decreases rapidly as the degree increases. Scenes marked by a plus symbol at the bottom left in the panel show a particularly strong wavelength dependency in that the power of short wavelengths decreases more rapidly than that of longer wavelengths. Therefore, for some scenes longer wavelengths seem to have finer spatial structure. However, note that this wavelength dependency was not observed for all scenes.

One important implication from the present findings relates to human perceptual colour constancy, which is the ability of the human colour vision system to support the stable perception of an object's surface colour regardless of illumination. Most previous studies of human colour constancy have used a simplified experimental setting where an object is illuminated by a single-spectrum illumination [29]. Under such lighting environments, the computation to achieve colour constancy could be conceptualised as first estimating the illumination colour and then discounting its influence from the whole visual field. This approach may be difficult in real-world situations because an object receives different lights from every direction. Moreover, by slightly changing head or object position, the object samples a different set of incident lights. The measured data also raise an interesting question about which white point our visual system adapts to when there is more than one candidate white point in the scene. For example, adaptation to lights from above and to lights from below would result in substantially different retinal adaptation states. Nevertheless, various research suggests that our perception of surface properties of objects such as shape, colour or gloss is not critically impaired under such complex lighting environments [141], [164]–[166], [168], [169], [171], [172], [192], [193]. One might argue that the visual system would not have access to the full illumination map at a glance. However, it is possible that we accumulate information over time by eye or

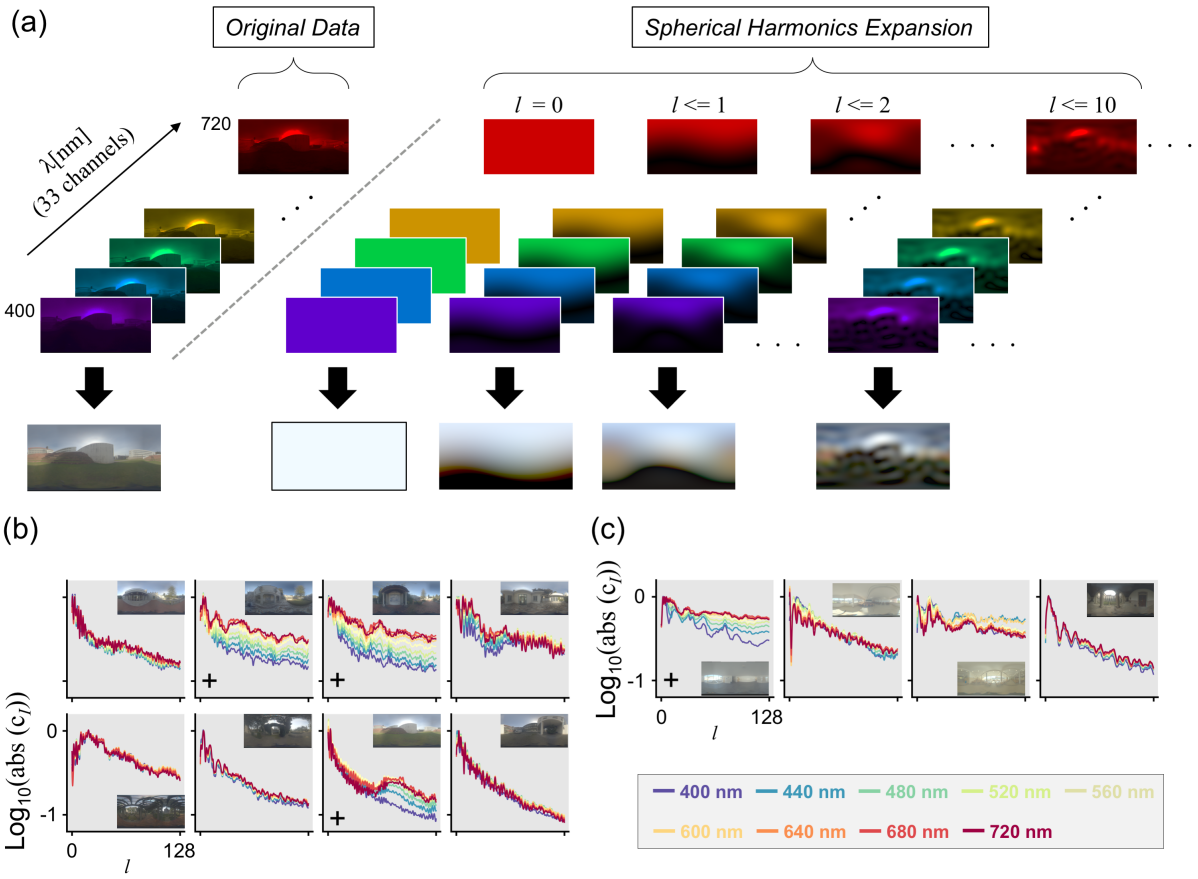


Figure 5.8: Application of spherical-harmonic expansion to environmental illuminations. (a) Spherical-harmonic expansion was performed independently for each wavelength. The variable l represents the degree and the right-hand panel shows the reconstructed images up to a specified degree level. (b) Power spectra of spherical-harmonic expansion for wavelengths from 400 nm to 720 nm every 40 nm as a function of degree l for outdoor scenes. (c) Power spectra of indoor scenes.

head movements [87]. A recent behavioural study suggests that the visual system might indeed summarise the influence of illumination under a complex lighting environment via a relatively simple strategy such as spatio-temporal averaging of the illumination map [232]. Although further evidence is desired, this line of research will potentially provide interesting and novel insights towards understanding human vision in more realistic environments. Another notable feature is the failure of the daylight assumption for lights from below for several scenes. This is consistent with one study on colour constancy that suggests the visual system may not actually rely on the daylight assumption [116].

One limitation of the present study is related to the spatial resolution of the image

due to the use of a mirror sphere and limited resolution of the hyperspectral camera itself (the image of the sphere was sampled at 1024×1024 pixels). At the edge of the sphere, the reflected image of a scene is highly compressed so spatial resolution is low for these regions. However, by stitching four images we were able to eliminate the left and right edges of the image where compression is most severe. Specifically, we used only the central 82.0% of each sphere image and, since the compression follows a cosine function (as described in the panel (v) in Fig. 5.2), the azimuthal compression factor was always less than 27.3%. Since the elevation angle of the camera was not changed, we were not able to eliminate the compressed regions at the top and bottom of the sphere. Thus, light from directly above and directly below was more sparsely sampled but, in general, there was less spatial variation in the scenes in these directions. For future measurements, we plan to introduce a fish-eye lens which can more uniformly sample illuminations across incident angles compared to a mirror sphere. Spatial resolution might additionally be lowered due to some of the post-processing of the image, such as image registration (correction of lateral chromatic aberration) or image stitching, since these manipulations might introduce misalignment.

The second limitation is that the present measurements avoided direct sunlight, and thus the dynamic range was also limited. Hyperspectral measurement of a single sphere image can take over 10 minutes to capture sufficient light for all wavelengths. We prioritized the data-acquisition time over the dynamic range, which would require a change of the exposure time, because otherwise the lighting environment is likely to change during the measurement. Methods to measure multispectral environmental illuminations have been improving over recent years [233], [234], and it will be important to measure additional hyperspectral datasets for further generalization of the findings. Nevertheless, the present study provides an important first step toward understanding the complex nature of environmental illumination, especially in the spectral domain.

5.4 Conclusion

Based on novel datasets collected in this study, we found that the chromaticites of natural illuminants are not restricted to the daylight locus, but instead show a much wider chromatic variation than previously thought. More importantly, the data show that natural lighting environments possess significant directional spectral variation. The degree of variation seems to be much larger in outdoor scenes than in indoor scenes.

Supplemental Information

Figs. 5.9 and 5.10 show directional spectral variation for other outdoor and indoor scenes, respectively. Here, we see that the degree of spectral variation depends somewhat on the scene. However, we observe a similar general trend as described in the main text. That is, for outdoor scenes as shown in Fig. 5.9, light from above tends to be more bluish and high luminance whilst lights from below have less energy in the short-wavelength region and lower intensity. Indoor scenes have relatively uniform spectral shape regardless of direction, but regions that include large windows or that are open to the sky have particularly high luminance.

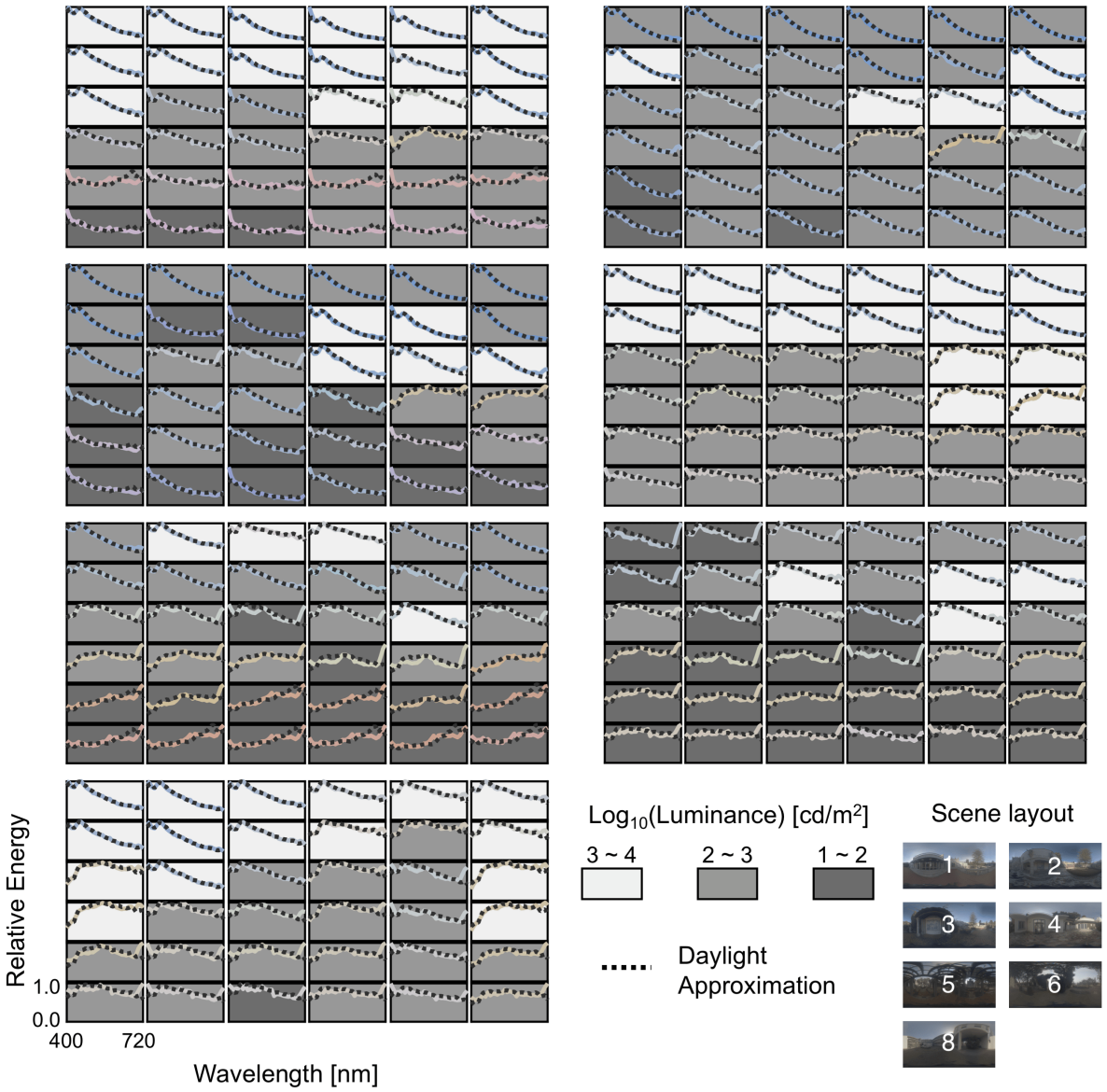


Figure 5.9: Directional spectral variation for the remaining 7 outdoor scenes. The way the data is plotted follows Fig. 5.4. The colour of each curve is the sRGB colour of the spectrum. The background colour represents the log luminance level. The black dotted line indicates the spectrum derived from CIE daylight basis functions.

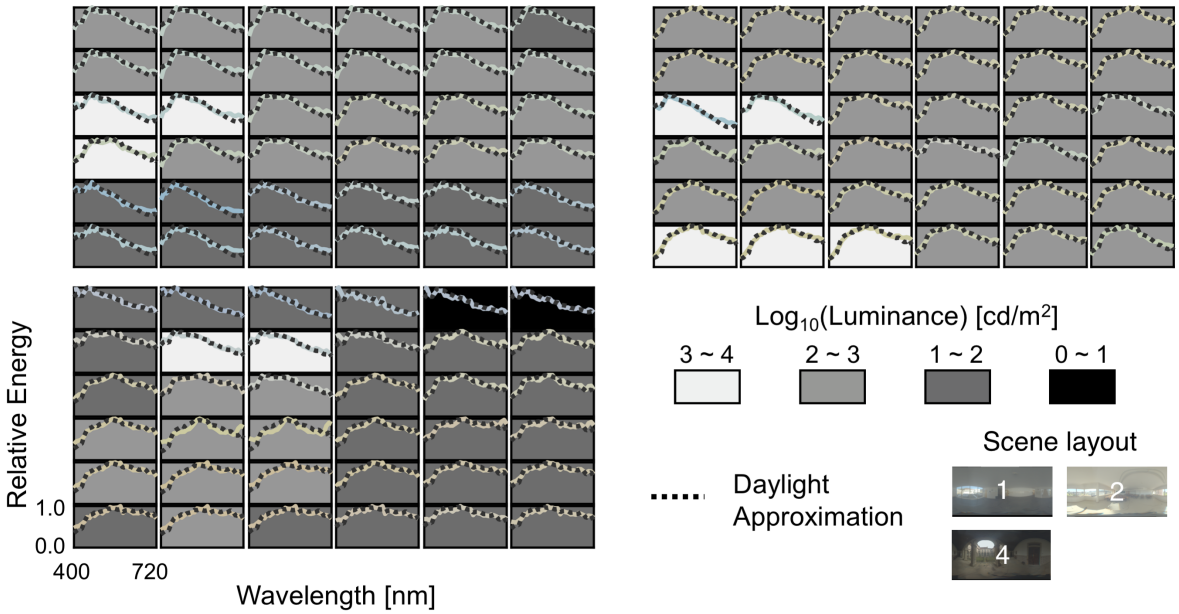


Figure 5.10: Directional spectral variation for the remaining 3 indoor scenes. The colour of each curve is the sRGB colour of the spectrum. The background colour represents the log luminance level. The black dotted line indicates the spectrum derived from CIE daylight basis functions.

6

Breaking metamerism using directional spectral variation in illumination

6.1 Introduction

Illuminant metamerism is known to impair our ability to identify objects by their colour, but its prevalence has traditionally been evaluated under single-spectrum environments, where the scene is lit by illumination with only one spectral power distribution. This study considers the presence of directional spectral variation that is present in natural environments, and estimates the prevalence of illuminant metamerism in such environments. The analysis showed that simply by tilting surfaces, both trichromats and dichromats can disambiguate roughly the half of the indistinguishable reflectance pairs found under a single-spectrum environment.

Colour plays a diagnostic role in object recognition. It allows us to segregate objects from their backgrounds [1] and find them more quickly [235]. Under daylight conditions, human perception of surface colour is mediated by the responses from the three classes of cone in our retinae. When light reflected from an object's surface reaches our retinae, the light is converted to cone signals. The surface colour of the object is essentially determined by the relative balance among those three numbers. Importantly, at this very first stage of visual processing, information about wavelengths is lost (the so-called

principle of univariance [2]). The input to human colour vision is fundamentally limited in this way, and thus any higher level process is also subject to this limitation. Because of this, two lights that have distinct spectral energy distributions can elicit the same cone signals. This phenomenon is known as metamerism. For example, a monochromatic light at 580 nm elicits a yellow percept, but it is possible to generate a colour that appears identical by balancing the amount of two monochromatic lights at 520 nm and 620 nm. In a similar manner, reflected lights from two objects that have different spectral reflectances, can elicit different cone signals under one illumination but the same under another illumination. As a result, two objects that are distinguishable under one illumination become indistinguishable to our visual system under another (see Supplementary Material for an example). This is called illuminant metamerism [19] and has been recognised as a problem for colour vision systems. Because it occurs at the cone level and there is nothing that later visual processing can do to break this conditional colour matching. In this sense, metamerism has been considered to be as a hard limit on human colour vision that is supported by only a limited number of cone classes. Considerable effort has been made over decades to quantify the impact of illuminant metamerism on human colour vision [236], [237].

In past studies, however, the prevalence of metamerism has been evaluated under a simplified environment where two surfaces are illuminated by a single-spectrum illuminant, and the consequence of having multiple illuminations has not been explored. Importantly, recently collected hyperspectral environmental illumination maps [194] revealed that natural environments hold a significant directional spectral variation. As a consequence, a flat surface placed in the natural environment will sample different spectra depending on its tilted angle. If we imagine a pair of two flat surfaces (i.e. paper), these might appear the same at a certain tilted angle, but it might be possible to disambiguate them simply by tilting the surfaces toward another direction.

For dichromats, who possess only two classes of cones, illuminant metamerism is more frequent, and therefore it is a more severe problem [238]. However, Broackes [239] argued that it is possible for dichromats to recover the lost dimension of cone signals. He showed theoretically that dichromats can solve some metamerism by changing

the illumination or by changing viewing angle to look at surfaces to allow surfaces to sample a different illumination from a different part of the sky. It is therefore an interesting question to ask how much dichromats can benefit from the directional spectral variation that natural environments offer.

This study aimed to revisit illuminant metamerism in the light of environmental illumination. We quantified the impact of metamerism on trichromats and dichromats under complex lighting environments where the spectral content of lights vary from one direction to another.

6.2 Method

To quantify the frequency of illuminant metamerism, we firstly simulated a matte planar surface using computer-graphics rendering techniques. Panel (a) in Figure 6.1 shows the geometry of a camera and a planar surface that is tilted towards the camera by 45 degrees. The surface was completely matte (no specular reflection) and tiltable so that it can sample different spectra from different directions in the scene. We use an environmental illumination map when rendering the scene. We used 19,123 spectral reflectances for the evaluation drawn from the standard object colour spectra database for colour reproduction evaluation (SOCS, ISO/TR 16066:2003). The scenes were then rendered multispectrally from 400 nm to 700 nm with 10 steps using the Mitsuba renderer [176].

The goal of our analysis here was to find the proportion of indistinguishable reflectance pairs out of 1.83×10^8 possible pairs, given a particular tilted angle, and then to ask how many of those metamers we can break by tilting the surface. As a criterion of being distinguishable, we chose to use Δ E = 1.0 in CIECAMO2-UCS colour space [20]. We used Brettel's model [240] to convert XYZ tristimulus values for trichromats to XYZ for dichromats. Intuitively speaking, the transformation is performed by squashing one dimension of XYZ colour space that dichromats cannot access, and the converted values were transformed to coordinates of CIECAMO2-UCS. Δ E values for dichromats were calculated in this transformed CIECAMO2-UCS space.

Panel (b) shows variation in the appearance of the surface that has a flat reflectance when it is tilted towards various directions under the set environmental illumination. When the surface is tilted upwards (θ>=45), the surface mainly samples directly emitted lights from the sky, whilst when the surface is tilted downwards (θ<45), the surface receives mainly inter-reflections from the ground. The panel shows that surface colour changes dramatically depending on the tilted angle. Panels (c) and (d) show spectra reflected back from surfaces at five different tilted angles along θ (blue and green squared surfaces in Panel (b)) and ϕ (red and green squared surfaces in Panel (b)) and their chromaticities in the MacLeod-Boynton chromaticity diagram [17]. Note that the rendered surfaces had a slight spatial spectral variation in their reflected lights, and we spatially averaged those spectra. We see that their mean spectrum changes substantially depending on the tilted angle. Also, the chromaticity of the reflected lights also largely shifts along the daylight locus (shown by black solid line) depending on the tilted angle, and the variation seems to be larger along the θ direction (sky versus ground) than the ϕ direction (compass direction).

Figure 6.2 illustrates the conceptual idea of how to solve indistinguishable pairs using directional spectral variation. Panel (a) shows the case for trichromats. We chose two surface reflectances as shown at the right hand-side of the panel. As shown by the yellow circle and diamond symbols, at a certain tilted angle, the chromaticities of the reflected lights are too close for the visual system to discriminate. However, the surfaces' chromaticities shift differently by tilting them to another angle, and as a result they become distinguishable. Panel (b) illustrates the case for a deutan observer who lacks M cones. At a certain angle, the two chromaticities of the two reflected lights are on the deutan confusion line, on which colours differ only in M cone excitation. However, by tilting the surfaces, the chromaticities of the reflected lights move away from the confusion line, which allows deutans to see that they are different.

6.3 Results and Discussion

In Figure 6.3, we show the percentage of indistinguishable reflectance pairs that are disambiguated by tilting the surface, relative to five reference frames (the surfaces marked

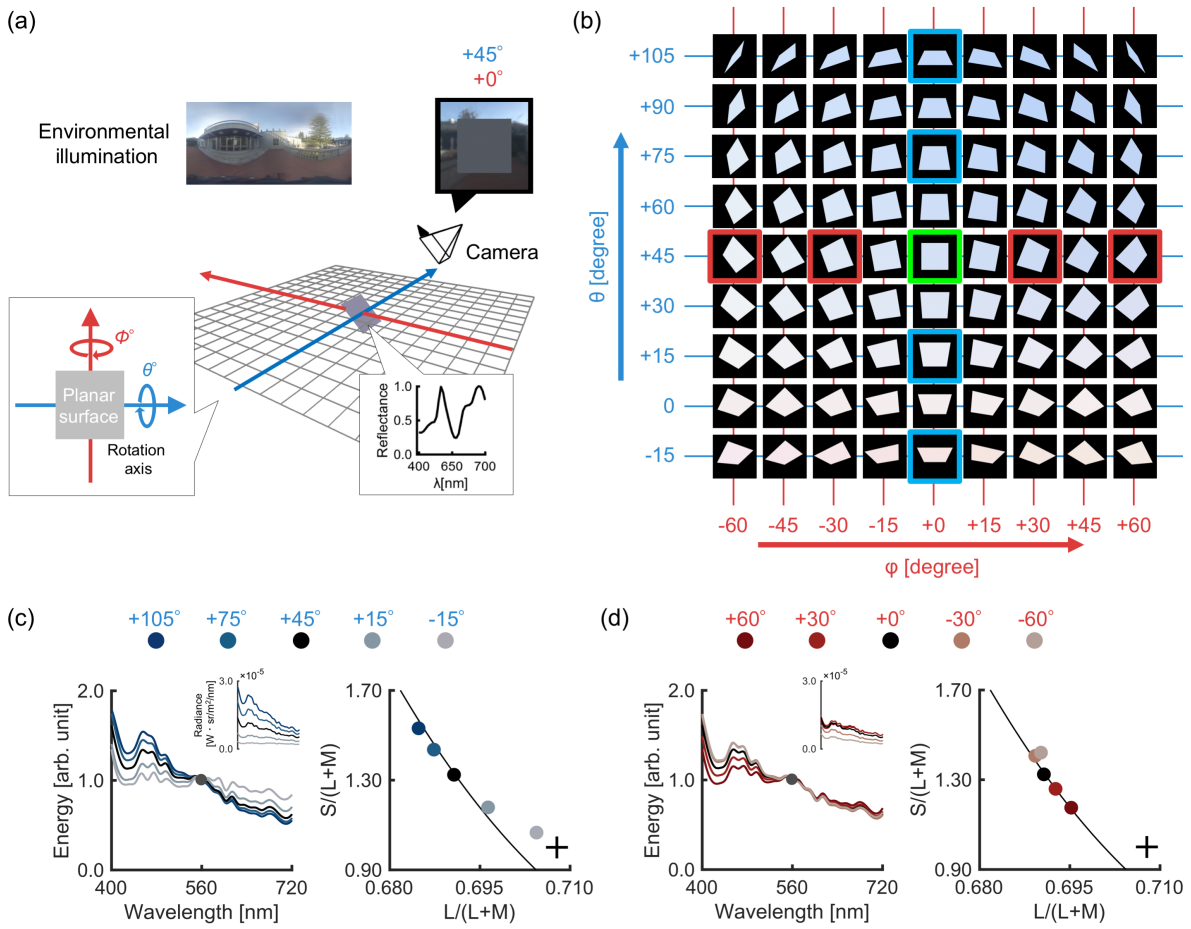


Figure 6.1: Spectral variation in the reflected lights from a white planar surface under an environmental illumination. (a) Standard set-up of the scene geometry. The planar surface was tilted 45 degrees towards the camera. From this angle the object was tilted ± 60 degrees along θ and ϕ directions as shown at the left-side. Panel (b) shows rendered planar surfaces at different tilted angles. Panel (c) shows the variation of the spatial-mean of reflected spectra from the surface at five different elevation angles θ (blue and green squared papers). Each spectrum was normalised at 560 nm, but the small inset plot shows spectra in absolute units. The right small figure in the panel (c) indicates the chromaticity of each spectrum. Panel (d) shows the variation of spatial-mean of reflected spectra from the surface at five different azimuth angles ϕ (red and green squared papers)

by green and cyan squares in panel (b) of Figure 6.1). If there were 100 indistinguishable reflectance pairs at the 45 degrees reference frame (shown by green), and after tilting the surface to all other directions, 70 reflectance pairs are still found to be indistinguishable, the percentage of solved metameric pairs would be 30%. We repeated the same analysis for the other 4 reference frames, and the results were averaged. Panel (a) shows the results for individual scenes, and the averages across scenes are shown in panel (b). The

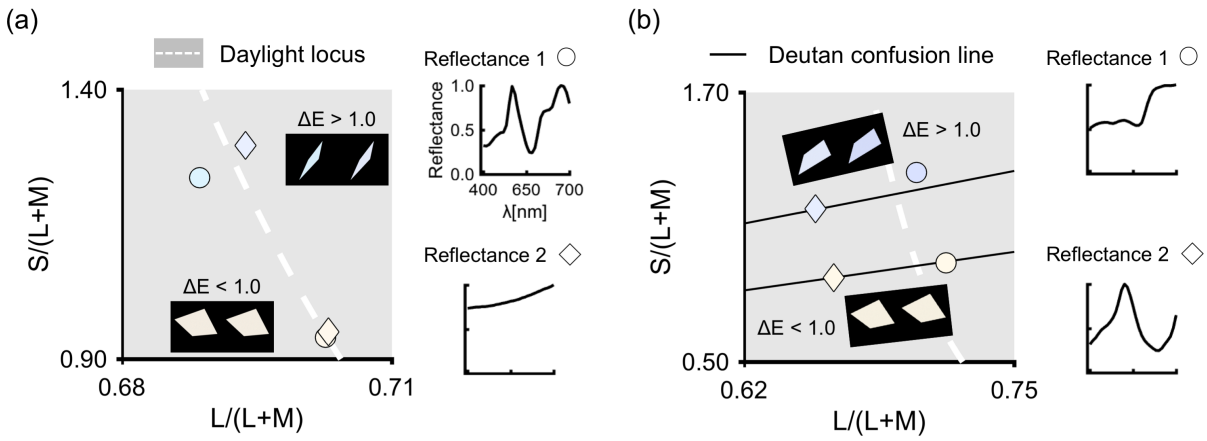


Figure 6.2: An illustration of the conceptual idea of breaking metamerism using directional spectral variation in illumination. (a) Under a particular environmental illumination, two surfaces at particular tilted angles have too similar colours to disambiguate. However, simply by tilting them to an other direction, the chromaticities are sufficiently separated for us to make a discrimination. Panel (b) shows the case for a deutan observer who misses M cones. At a particular tilted angle, two surfaces fall on the deutan confusion line, but tilting the surfaces can move the chromaticities so they move away from the confusion line.

gray, green, red, and blue bars indicates the results for trichromat, deutan, protan, and tritan observers, respectively. The averaged data in panel (b) revealed that trichromats can break 58.5% of indistinguishable pairs in outdoor scenes and 45.8% in indoor scenes simply by tilting the surface. Dichromats can break a slightly higher proportion (deutan: 70.9%, protan: 68.6%, tritan: 74.6% in outdoor scenes, and deutan: 57.7%, protan: 53.4%, tritan: 57.0% in indoor scenes). The main effects of scene type (outdoor or indoor) and colour vision type (trichromat, deutan, protan and tritan) were significant ($F(1,8) = 7.45$, $\eta^2 = 0.965$, $p < .05$; $F(3,24) = 49.9$, $\eta^2 = 2.50$, $p < .00001$) while the interaction between them was not ($F(3,24) = 1.71$,$\eta^2 = 0.463$, $p > .10$). Multiple comparisons using Bonferroni's correction found that relative improvements are larger for any type of dichromat than for trichromats and for tritan compared to protan.

The study of human colour vision is a historically rich research domain in visual neuroscience, and illumination plays a key role in object colour perception. However, early studies typically used a simplified environment where only a single illumination uniformly hits objects' surfaces (known as "Mondrian stimuli" [29]). In such situations, illuminant metamerism could present a significant challenge for the human colour vision system. However, in the present study, we found that the spectral directional variation present in

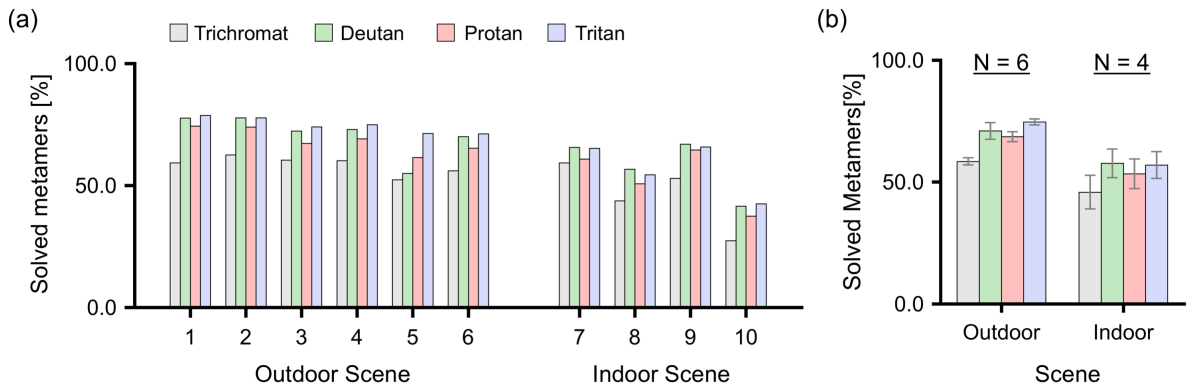


Figure 6.3: Results of the analysis. (a) The proportion of solved metameric pairs for 6 outdoor and 4 indoor scenes. The gray, green, red and blue bars indicate the results for trichromats, deutan, protan and tritan, respectively. Values indicate the percentage of broken metameric pairs averaged across five reference frames (blue and green squares in Figure 6.1). (b) Average across outdoor scenes and indoor scenes.

natural environments may allow human colour vision, regardless of whether the observer is a trichromat or dichromat, to distinguish roughly half of the reflectance pairs that cannot be discriminated under a conventional single-spectrum illuminant. In practice, the short term memory of colour across different angles would allow us to implement such a strategy. Some recent studies have suggested that illuminant metamerism might be relatively rare [241], [242]. Here, we are suggesting that metamerism is even rarer in practice thanks to the presence of directional spectral variation.

This study highlights the importance of using experimental stimuli that capture the complexity of the illumination environment in which our visual system operates in daily life. We believe this study challenges existing view, which have arisen from the use of simplified environments. Combining recently measured light-measurement data with computer-graphics rendering techniques has provided a new insight into metamerism under real-world illumination. Directional variations offered by natural environments compensate to some extent for the biological limitations of trichromatic and dichromatic visual system.

6.4 Supplementary materials

6.4.1 An example of illuminant metamerism

Figure 6.4 shows an example of illuminant metamerism. Panel (a) shows two objects that have two distinct surface spectral reflectances are placed under a single point light source which has a greater energy around the short-wavelengths (illuminant 1). The lights reflected back from these two surfaces elicit two different sets of cone signals as shown below the panel, and this is why we are usually able to distinguish these two surfaces. However, when the spectrum changes to illuminant 2 shown in panel (b), the reflected lights change accordingly, and those lights happen to produce exactly the same cone signals from the two objects. In this case these appear identical even though the two objects still have distinct surface reflectances. This is called illuminant metamerism.

6.4.2 Set of surface spectral reflectances

In this study, we used *the standard object colour spectra database for colour reproduction evaluation* (SOCS, ISO/TR 16066:2003) to quantitatively evaluate the number of indistinguishable reflectance pairs under particular environmental illuminations. The original database had 53,486 reflectances measured from a variety of items in the natural environment, including painting materials (e.g. ink or dyes), flowers, leaves, natural objects (Krinov datasets) and human skin. However, the 53,486 reflectances generates 1.43×10^9 candidate pairs, which was too large a search to implement with our current computing resources. Also, the database contains many reflectance pairs that are physically very similar. In an extreme case, a pair of two identical surface reflectances is indistinguishable under any illumination, and thus using these spectra leads to an overestimation of the number of metameric pairs. Thus, we decided to reduce the number of reflectances using the following criteria. The chromaticities of all reflectances, when rendered under equal energy white, are shown in Figure 6.5(a). The intersection of the gray dotted lines shows the chromaticity of equal energy white. The panel shows histograms of all chromaticities along L/(L+M) and S/(L+M) axes. The peaks of histograms are slightly shifted towards higher L/(L+M) and lower S/(L+M), indicating that the there are more orange surfaces

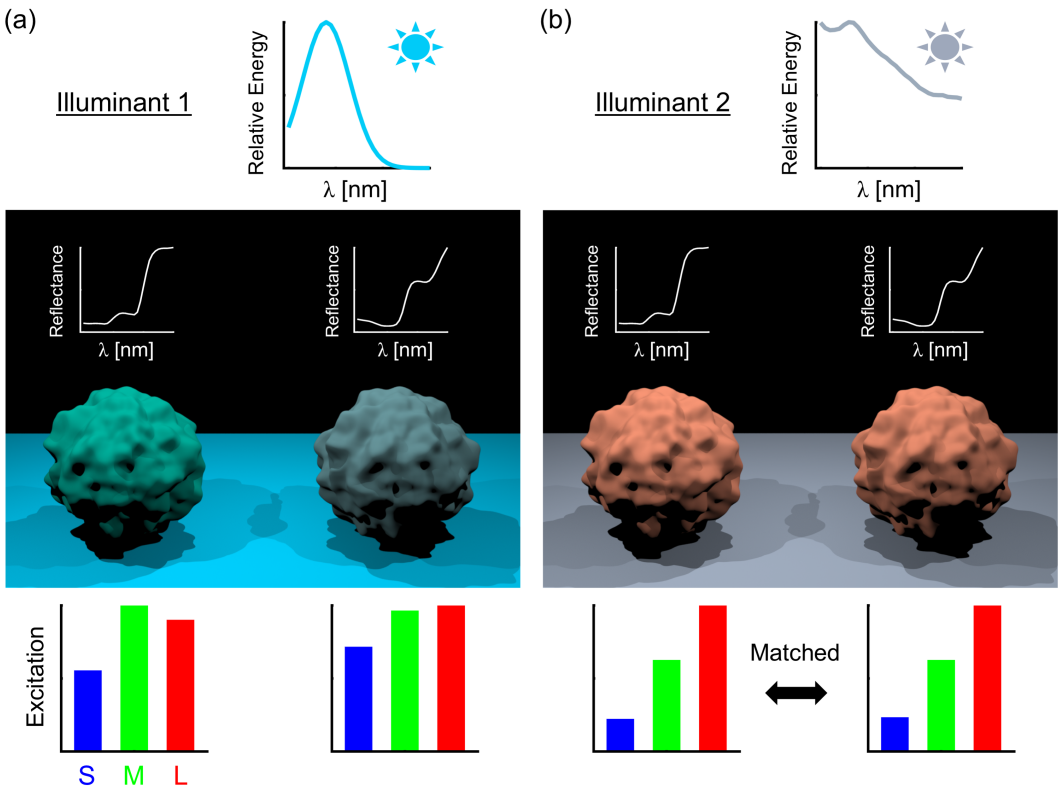


Figure 6.4: An example of illuminant metamerism. The panel (a) is a situation where two objects, which have different surface spectral reflectances, illuminated by a single spectrum shown above. In this case, the reflected lights from these two objects generate different cone signals, which allow us to perceive different surface colours. However, when illumination changes to the illuminant 2 as shown in the panel(b), updated reflected lights elicit the same cone signals, in which case we are not able to discriminate two objects. This problem is widely known as illuminant metamerism.

in the database. Thus, for the evaluation presented in the main text, we identified reflectance pairs that have a Pearson's correlation coefficient of more than 0.999 and removed one of reflectance from each pair such that peaks of the histograms correspond more closely to the white point (the chromaticity of equal energy white). Furthermore, to protect against bias due to the choice of the set of reflectances, we analysed a database that had been further reduced by setting the criterion correlation coefficient to 0.99. The result will be shown later in the Supplementary information.

6.4.3 Environmental illumination

Figure 6.6 shows the 10 environmental illumination maps (6 outdoor scenes and 4 indoor scenes) used for the analysis. Two scenes that were reported in our previous study

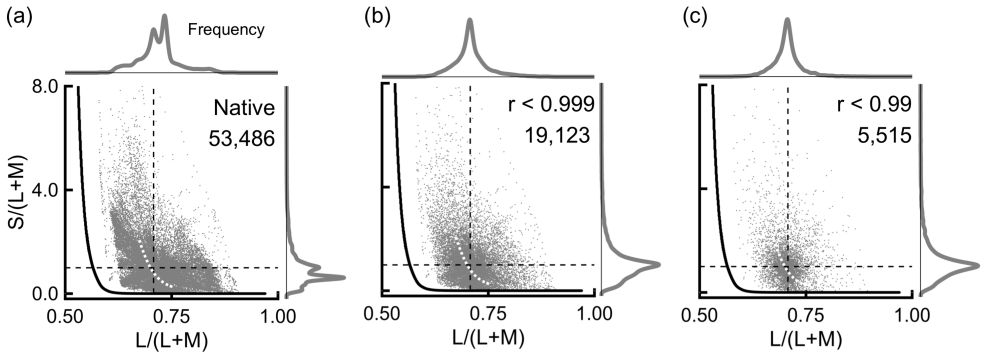


Figure 6.5: Chromatic distributions and histograms of the reflectance datasets for evaluation when rendered under equal energy white. The intersection of the gray dotted lines indicates the chromaticity of equal energy white. The dotted white line indicates the daylight locus. Panel (a) shows the case for the original database. Panels (b) and (c) indicate reflectance sets after too similar reflectances were removed from the original database ($r > 0.999$ and, 0.99, respectively).

[194] were excluded from the evaluation as there was a slight misalignment in the scene capture that might introduce an artefact in the analysis. More details on the properties of these environments, such as the chromatic distribution of all pixels and directional spectral variation, are found in the previous study.

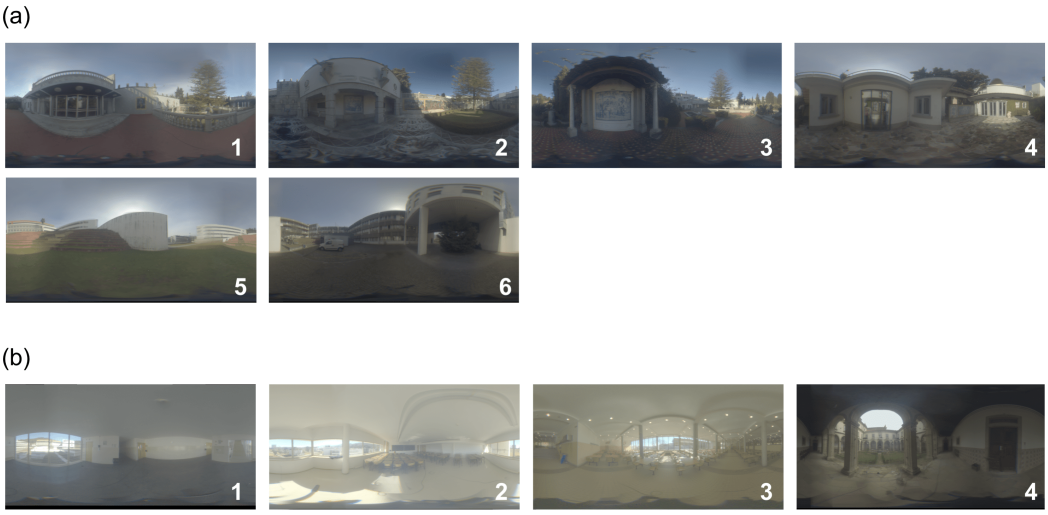


Figure 6.6: sRGB representations of the 10 environmental illumination maps used in this study. Each pixel of the image stores information about the spectrum (from 400 nm to 720 nm with 10 nm steps) coming from a particular direction in the scene towards a single point (e.g. an eye or a camera). More detailed information is given in the previous study [194].

6.4.4 Criteria to define indistinguishable surfaces

Past studies have used a variety of metrics to estimate the frequency of metamerism. For example, one early study used distance in XYZ tristimulus values [236], while a more recent study [242] chose to evaluate differences in a uniform colour space that accounts for the sensitivity of our visual system to colour difference in different regions of the space. There is no gold standard, but the present study used Δ E = 1.0 in CIECAMO2-UCS, which is a uniform colour space developed by the CIE committee and currently adopted as a standard colour space within which to evaluate chromatic differences. There are several parameters that need to be defined to construct this colour space. For the white point W_p, we used averaged XYZ values across 81 rendered surfaces, that are shown in panel (b) in Figure 6.1), that all had a reflectance equal to 1.0 across all wavelengths. The calculation of the average was done in XYZ colour space. For the adapting field luminance L_A, we used the mean XYZ values (2 degree, CIE1931) across all pixels in the environmental illumination. For Y_b, we used the ratio between the luminance of the adapting field (L_A) and the luminance of the white point (W_p). These parameters were designed to incorporate the scene's white point W_p, the luminance level of the background L_A and observers' adaptation state Y_b that are known to modulate colour appearance. These parameters were separately set for each environment. For viewing condition, there are three options in CIECAMO2-UCS: dark, dim, and average. We set it to "average" because this is defined by CIE as the condition appropriate for viewing the colour of a surface.

6.4.5 Supplementary Analysis

Panel (a) in Figure 6.7 shows the absolute numbers of indistinguishable pairs (the whole bar) found when the surface was at 45 degrees tilted angle (the surface outlined by green in the panel (b) in Figure 6.1) and the number of disambiguated pairs (white region in the bar) produced by tilting papers to all other angles. The colours gray, green, red and blue indicate the type of colour vision. To plot each bar, we calculated the number of metameric pairs separately for each of five different reference points (blue and green squares in Figure 6.1), and the numbers of solvable pairs by tilting the surface, and the results from the five reference points were averaged. Panel (b) indicates the average across six outdoor

scenes (left group) and across four indoor scenes (right group). Panel (b) shows that out of 1.83×10^8 candidate pairs trichromats can reduce the number of indistinguishable pairs from 1.19×10^5 to 6.70×10^4 in outdoor scenes and from 1.14×10^5 to 7.85×10^4 in indoor scenes. Dichromats' absolute number of metameric pair was roughly one order of magnitude more. However, relatively speaking, as shown in the main text (Fig 6.3), the proportions of solved metameric pairs were larger for dichromats than for trichromats.

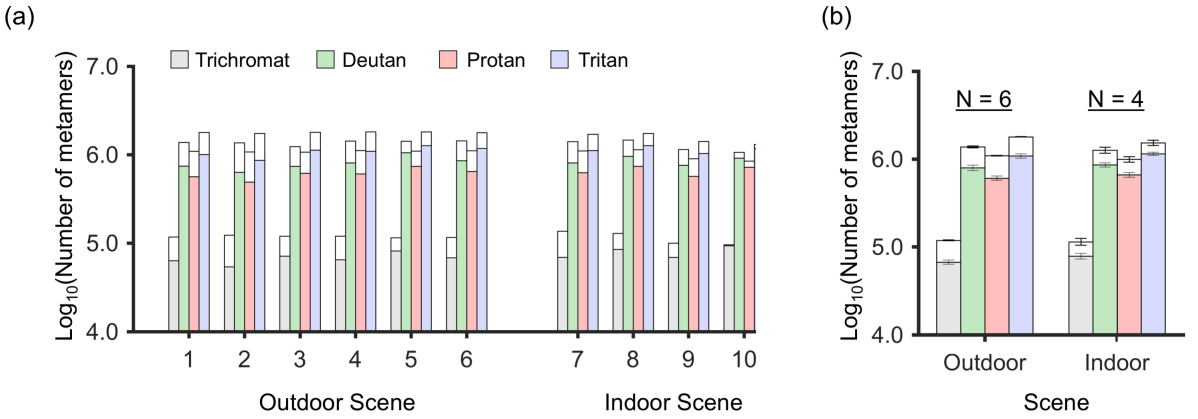


Figure 6.7: Absolute numbers of indistinguishable pairs on a log scale. (a) Number of disambiguated pairs for 6 outdoor and 4 indoor scenes. The gray, green, and blue bars indicate the results for trichromats, deutan, protan and tritan observers, respectively. The whole bar indicates the number of indistinguishable pairs found when the surface was tilted at a 45 degree angle (the surface outlined by green in the panel (b) in Figure 6.1). The white region of the bar indicates the number of indistinguishable pairs that can be solved using directional spectral variation. (b) Average across 6 outdoor scenes and 4 indoor scenes.

We also checked the extent to which the findings hold across different criteria. We analysed the data exactly the same way except for using different Δ E values to determine the indistinguishable surfaces. Figure 6.8 (a) and (b) show the relative number of broken metameric pairs as a function of Δ E taking values of 0.5, 1.0, and 2.0 for outdoor scenes and indoor scenes, respectively. The results suggest that the relative percentages change slightly in response to the Δ E values, but importantly, the difference between trichromats and dichromats seems to change little.

We ran a two-way ANOVA for (a) outdoor scenes, with factors of colour vision type (four levels) and ΔE values (three levels). The main effects of colour vision type and ΔE values were both significant ($F(3,15) = 31.3$, $p < .01$; $F(2,10) = 312.5$, $p < .01$). Also, the interaction between the two factors was significant ($F(6,30) = 7.95$, $p < .01$).

Further analysis of the interaction revealed that the effect of colour vision type was significant at every value of ΔE ($F(3,15) = 31.6$, $p < .01$ for $\Delta E = 0.5$; $F(3,15) = 32.9$, $p < .01$ for $\Delta E = 1.0$; $F(3,15) = 17.2$, $p < .01$ for $\Delta E = 2.0$). Also, the effect of ΔE was significant for every colour vision type ($F(2,10) = 2086.6$, $p < .01$ for trichromat; $F(2,10) = 234.0$, $p < .01$ for deutan; $F(2,10) = 74.4$, $p < .01$ for protan; $F(2,10) = 488.8$, $p < .01$ for tritan). Furthermore, multiple comparisons using Bonferroni's correction (corrected p value 0.5) showed that at any value of ΔE, deutan, protan, tritan all showed higher values than trichromats, and tritan showed higher values than protan. In addition, similarly, for any colour vision type, the solved proportion was higher for ΔE 0.5 than ΔE 1.0 and 2.0 and for ΔE 1.0 than ΔE 2.0.

In the same way, we conducted a two-way ANOVA for (b) indoor scenes. The main effect of colour vision type and ΔE values were both significant ($F(3,9) = 16.5$, $p < .01$; $F(2,6) = 185.1$, $p < .01$). The interaction between the two factors was again significant ($F(6,18) = 22.5$, $p < .01$). Moreover, the analysis of the interaction revealed that the effect of colour vision type was significant at ΔE 0.5 and 1.0 ($F(3,9) = 18.4$, $p < .01$; $F(3,9) = 25.6$, $p < .01$, respectively), but it was not significant at ΔE 2.0 ($F(3,9) = 3.65$, $p > .05$). Also, the effect of ΔE was significant at every colour vision type ($F(2,6) = 37.3$, $p < .01$ for trichromat; $F(2,6) = 197.5$, $p < .01$ for deutan; $F(2,6) = 146.8$, $p < .01$ for protan; $F(2,6) = 195.9$, $p < .01$ for tritan). Furthermore, multiple comparisons using Bonferroni's correction (corrected p value 0.5) showed that at ΔE 0.5 and 1.0, deutan, protan, tritan all showed higher values than trichromats. However, at ΔE 2.0, there was no significant difference. In addition, for deutan, protan and tritan, the solved proportion was higher for ΔE 0.5 than ΔE 1.0 and 2.0 and for ΔE 1.0 than ΔE 2.0. In contrast, for trichromats, ΔE 0.5 was higher than ΔE 1.0 and 2.0, but the difference between ΔE 1.0 and ΔE 2.0 was not significant.

In summary, the statistical tests detailed above showed that the claim that dichromats benefit relatively more than trichromats holds except in the case of ΔE 2.0 in indoor scenes.

As discussed in the section of 6.4.2 *Set of surface spectral reflectances*, some of the reflectance pairs were physically too similar, and we removed some of them from the analysis. We tested whether our findings might change under different criterion for

similarity. Thus, we removed reflectance pairs using correlation coefficient 0.99 as a cut-off and then re-ran the whole analysis. Figure 6.8 (c) and (d) show that for both outdoor and indoor scenes the relative proportions of solved metameric pairs are essentially constant when we remove more or less similar reflectance pairs.

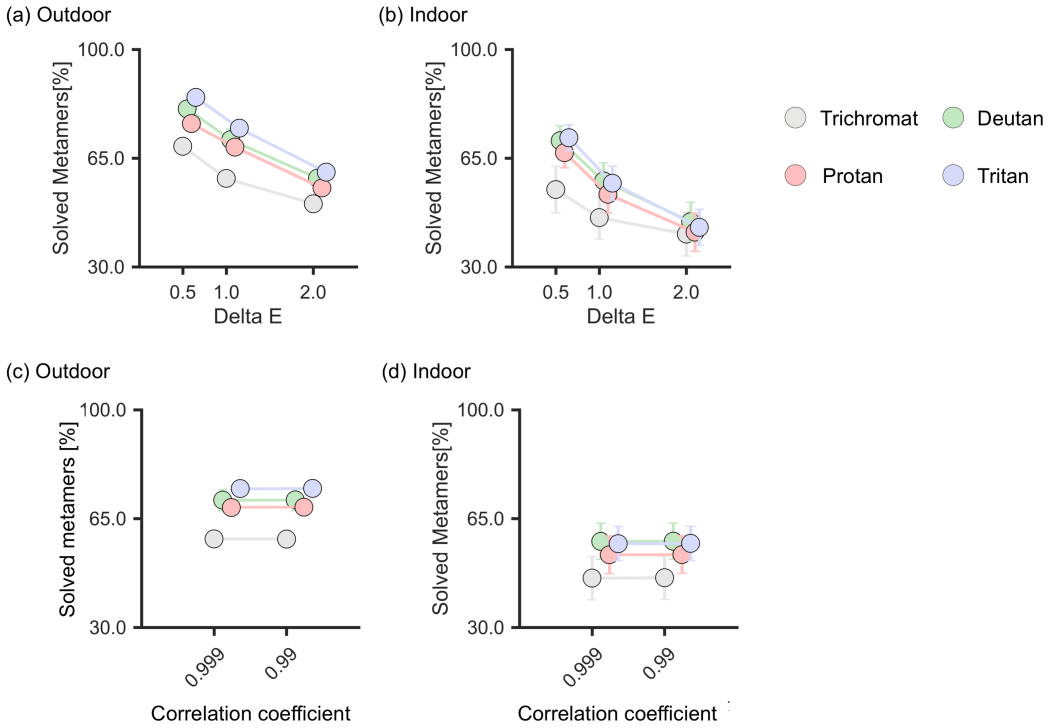


Figure 6.8: In panels (a) and (b), the relative proportion of solved metameric pairs are plotted as a function of three ΔE levels for outdoor and indoor scenes, respectively. The percentage decreases as a function of Δ E, but importantly, the difference between trichromats and dichromats seems to be relatively constant. In panels (c) and (d), the relative proportion of solved metameric pairs plotted as a function of the cut-off correlation coefficient for similarity of reflectance spectra in the test dataset for outdoor and indoor scenes, respectively. The relative proportion of solved indistinguishable pairs are kept constant as a function of cut-off correlation coefficient.

The final question we asked was which of the surface orientations are effective to break metamerism for each of the five different reference orientations. Panel (a), (b), (c) and (d) in Figure 6.9 show the case for trichromat, deutan, protan and tritan observers. Each pixel in the heatmap indicates the percentage of metameric pairs that can be broken if we tilt a surface towards the specified orientation relative to the centre frame (paper in Figure 6.1 outlined in green). For both outdoor (left six panels) and indoor scenes (right four panels), all colour vision types indicate that tilting towards the sky

is particularly effective to disambiguate the metameric pairs. However, interestingly, panel (a) shows that tilting towards lower angles, which samples inter-reflections from the ground, is not very effective for trichromats. In contrast, deutans seem to be able to solve some metameric pairs using lights from below.

Figure 6.10 quantifies this observation. Panel (a) shows the average percentage of solved metameric pairs across the upper 36 pixels in the heatmap (θ>0), and panel (b) shows the average across the lower 36 panels (θ<0). The case in Figure 6.9 is plotted in the centre column (labelled as reference 3) and it is also shown how this trend changes when the reference orientation changes along the elevation axis (from number 1 to 5). When the elevation angle is high (for example, for reference 1 or 2), dichromats showed slightly higher values than trichromats for the average across the upper 36 pixels, and the difference is much larger for the average across the lower 36 pixels. Thus, in this conditional sense, directional variation generated by inter-reflection benefits dichromats more than trichromats. However, note this observation is a dependent on the reference orientation. As the elevation angle for the reference point gets lower, this trend weakens.

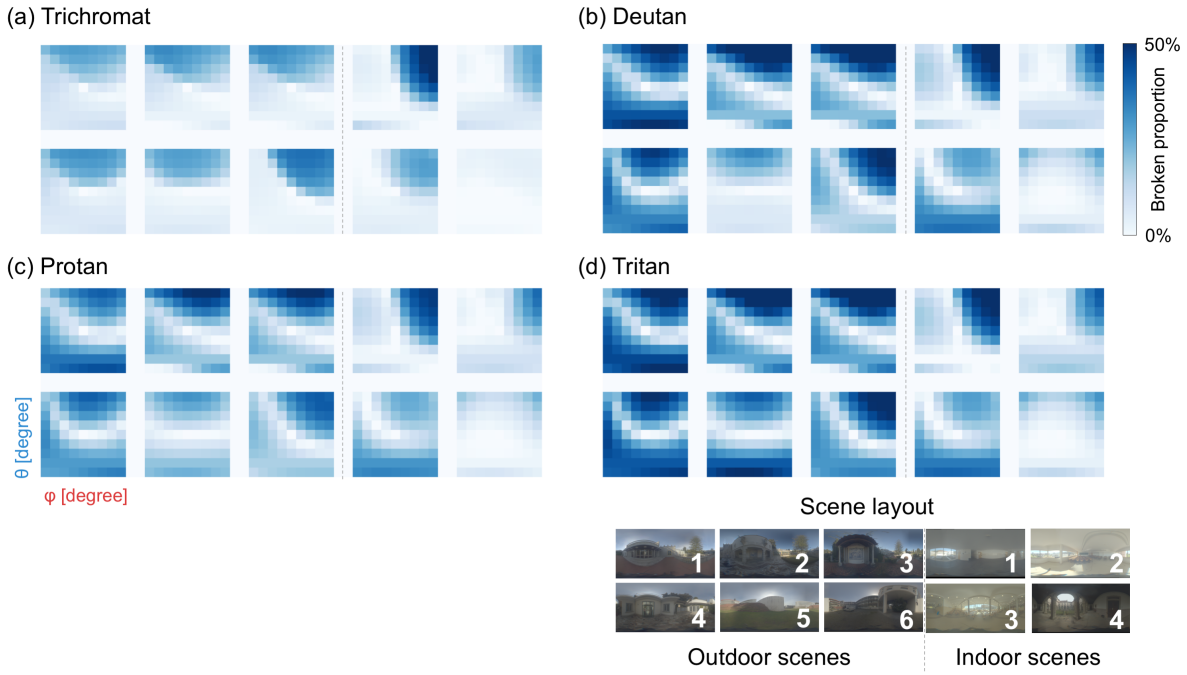


Figure 6.9: Heatmaps to show which orientations of the samples solve indistinguishable surfaces more efficiently. Each pixel in the heatmap shows the percentage of solved metameric pairs at a particular tilted angle relative to the reference oriented at 45 degrees (the surface outlined by green in the panel (b) in Figure 6.1). Different subpanels correspond to different scenes, as indicated by the images in the "scene layout" subpanels. The different panels indicate different colour vision types.

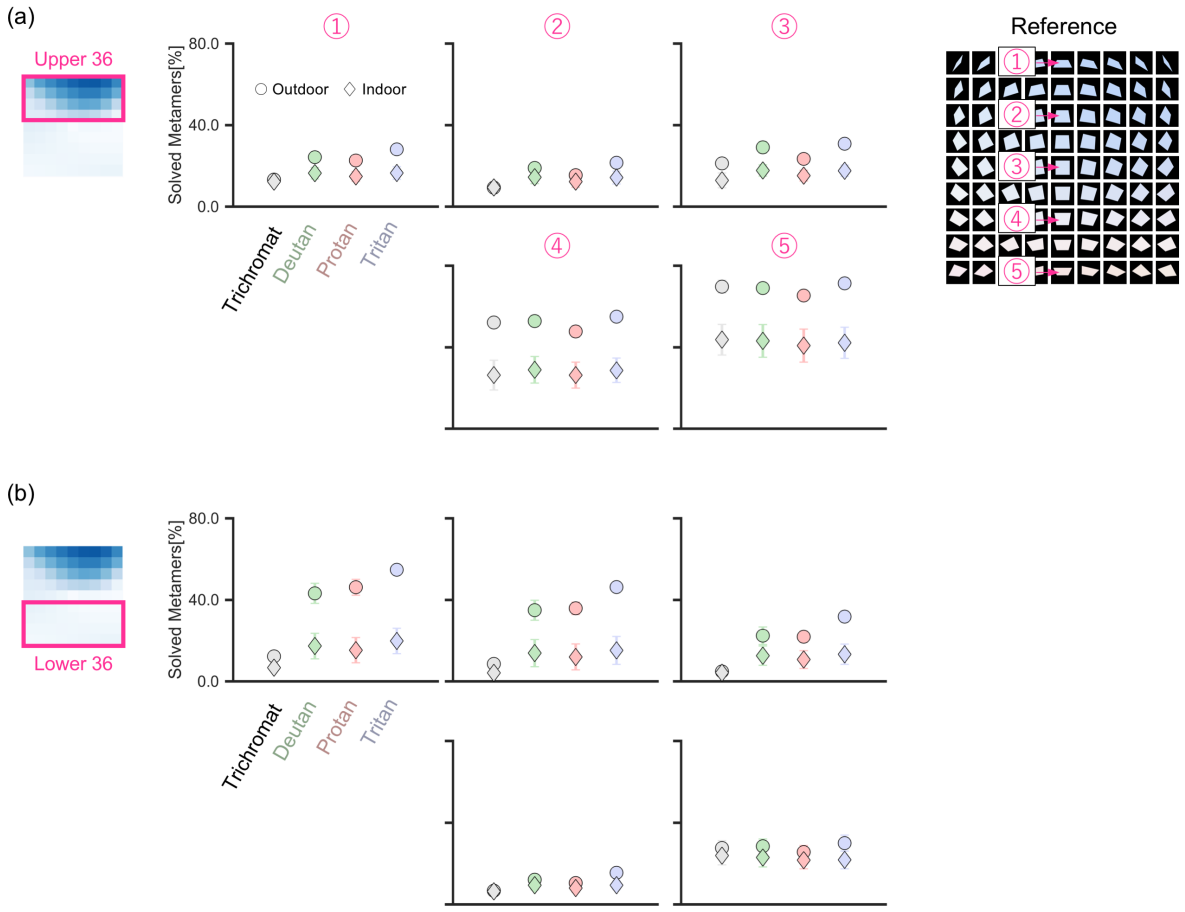


Figure 6.10: The average proportion of solved metameric pairs using (a) upper illumination and (b) lower illumination. Each point shows the average across 6 outdoor scenes (circle symbols) or 4 indoor scenes (diamond symbols) for each colour vision type. The results for five reference points are shown in different columns.

7 Summary and Conclusion

This project aimed to identify mechanisms that support stable perception of surface colour, with a particular emphasis on objects embedded in complex environmental illumination. We used computer graphics techniques to generate experimental stimuli, by which we moved away from the use of oversimplified stimuli that were extensively used in past studies. We approached the question by combining psychophysical experiments, computational modeling and hyperspectral measurement.

This chapter briefly summarises the main findings of each chapter. We also provide an overall summary of the DPhil project and discuss potential future directions.

7.1 Chapter 2

When the brightest is NOT the best: illuminant estimation based on highlight geometry

Illuminant estimation is known as a mathematically ill-posed problem as the spectral composition of reflected light from an object depends both on surface spectral reflectance and illuminant composition. Thus, our visual system rarely has direct access to illuminant colour within a scene. Rather, the visual system must rely on indirect cues to infer the influence of scene illumination. One such algorithm is to use the signal returned from a white surface that has a flat surface spectral reflectance. However, the difficulty

arises in identifying such signals, or approximations to them from the real world. The "brightest is white" assumption provides a heuristic in which the brightest element in a scene is assumed to provide the most informative cue about the illuminant colour. This heuristic has also been suggested to identify specular highlights which by definition provide information about the illuminant. In this chapter, we first showed that this principle does not always hold when a glossy objects has a high contrast texture over its surface. We then ran an experiment, to test whether or not observers' performances are consistent with application of this heuristic. Observers were asked to discriminate illuminant colour changes versus object colour changes. The results showed that observers' performances improved under a condition in which the brightest is white assumption collapses. This observation directly suggests that our visual system does not simply rely on the brightest element in a scene to estimate illuminant colour. Instead, we propose an alternative hypothesis: observers use the geometry of the surface and the illumination to select highlight regions, even when they are not the brightest elements in the scene. The findings from the present chapter potentially overwrite a notion that holds for simplified stimuli when the surface is matte and has uniform surface spectral reflectance over the surface, which was widely believed for decades.

7.2 Chapter 3

Reflectance discrimination under complex environmental illumination

Chromatic discrimination is of fundamental interest in the study of human colour vision. Many studies were conducted in the past mainly to characterise how colour vision is tuned in colour space by measuring discrimination ellipses. For example, the visual system might be more sensitive to variation in the red-green direction than in the blue-yellow direction. However, past studies mostly used a simple patches of light, and thus discrimination ability for more complex experimental stimuli such as surfaces that exhibit specular reflection is still unclear. Moreover, the effects of illumination, especially complex environmental illumination, on discrimination ellipses have never been

explicitly studied. Thus, the purpose of this experiment was to evaluate our ability to discriminate surface colours of matte and glossy objects placed under complex lighting environments. Experiment 1 was designed to explore the candidate factors that might affect our ability to discriminate the surface spectral reflectance of matte and glossy objects under complex environmental illumination. Experiment 2 was aimed specifically to investigate whether our visual system exploits the statistical chromatic regularity in natural environments to compensate for surface variation introduced by properties of environmental illumination. The results of Experiment 1 suggested that (i) specular highlights do not seem to have systematic effects on discrimination thresholds, (ii) the shape of objects also has little effect, but interestingly (iii) discrimination ellipses are tuned along the direction in which the chromaticities of all illuminations in environmental illuminations extend. The results of Experiment 2 in chapter 3 showed that this tuning effect seems to be dominated by the chromatic distribution of the environment used for rendering rather than depending exclusively on a statistical invariance of natural environments. However, the chromatically-inverted environment additionally gave rise to larger discrimination ellipses than the typical environment. Overall, it was shown that our ability to discriminate surface colours is more affected by the the properties of the illumination within which objects are placed, than properties of the surface of the object such as shape and specular reflection.

7.3 Chapter 4

Reflectance identification across complex environmental illuminations

In chapter 3, we measured the ability to discriminate surface colours of objects that are all placed under the same complex lighting environments. Chapter 4 tested our ability to identify surface colours of objects placed across different environmental illuminations. To address such a question, we designed a selection-based colour constancy experiment. We presented four objects simultaneously, such that the two objects on the left were placed under different environmental illumination than the two on the right. Three

objects shared the same surface spectral reflectance, while the fourth, target object had a different spectral reflectance. Observers were asked to pick the target object. The key point of this experimental design was that observers were not able to pick the target objects unless they compensated for the complex influence of environmental illumination. Results suggested that specular reflection has little effect on observers' performance. The effect of including surrounding context in the stimulus presentation was significant, suggesting that participants estimated the influence of illumination using surrounding information. We also constructed a computational observer model to simulate observers' pattern of behaviour. Given models based on simple statistics, it was shown that rather than trying to estimate the influence of illumination from a single trial, models that used spatial and temporal averages of surrounding colours over trials predicted human observers' performance better. This result indicates that visual system may exploit the history of chromatic signals in order to learn the representative chromatic statistics of complex environmental illumination.

7.4 Chapter 5

Hyperspectral environmental illumination maps: characterising directional spectral variation in natural environments

Characterising illumination is crucially important to understand human vision because it helps us to identify possible range of physical inputs to our visual system in daily life. However, previous studies have mostly measured lights that come from above, and have reported that chromaticities are limited to a line known as daylight locus. However, in the real world, we know that lights emitted from the dominant light sources bounce within a scene, and thus an object placed in such environment should receive modified spectra that would be different from every direction. This chapter aimed to investigate this directional dependency of illumination, using a hyperspectral imaging system and a mirror sphere that captures lights from wide incident angles. All measurements were conducted in collaboration with Prof. Sergio Nascimento at University of Minho, Portugal. We collected 12 environmental illumination maps, which store at each pixel the information about

the incident light coming from a particular direction towards single point in the scene. The analysis of the collected environmental illuminations (8 outdoor and 4 indoor scenes) revealed that natural environments possess a significant amount of directional variation. For outdoor scenes, such variation of spectra was mainly as a function of elevation. Lights from above tended to be more intense and have greater energy around short-wavelengths. In contrast, lights from below were much weaker, and they had lower correlated colour temperatures. In addition, lights from below were not well described by daylight models. Such observations provide new insights in understanding human colour vision. In daily life, we presumably receive systematic spatial pattern of light inputs to our retina, meaning that photosensitive cells in our retina are non-uniformly adapted. Over the life time, amount and spectral composition of light that each cone receptor can receive seems to be restricted by this lighting gradation in natural environments. This directional dependency makes the problem of illuminant metamerism smaller, which is discussed in chapter 6.

7.5 Chapter 6

Breaking metamerism using directional spectral variation in illumination

In chapter 5, we measured the directional dependency of illumination. These data suggested new insights into the study of illuminant metamerism. Illuminant metamerism is the term used to describe the situation in which two distinct surfaces appear the same under one illumination but appear different under different illumination. This metamerism happens when the reflected lights from two distinct surfaces happen to invoke same triplet of cone signals. Thus, illuminant metamerism is a biological limit of human colour vision. However, directional variation in lighting means that when we tilt a surface, it will sample different illumination, which potentially disambiguate metameric pairs. We performed theoretical analyses based on measured illumination maps that showed that we can indeed solve around 50 % of illuminant metamerism that occurs in conventional single-spectrum environments. This revisit of metamerism with new

hyperspectral measurements and simulation based on multispectral rendering provided a exciting new perspective into the limits of human colour vision in real world environment.

7.6 Overall summary

The list below summarises the key findings of this project.

1. Using experimental stimuli that incorporate some of the complexities of real world surfaces can reveal that heuristics that work with simplified stimuli (a) may not work for complex stimuli, and (b) do not describe human perceptual performance. We suggest the alternative strategy to estimate the influence of illumination using a geometry of specular highlights.

2. The human perceptual ability to discriminate surface colours was not dependent on surface properties (gloss vs matte, which has a large effect on image properties), but was strongly affected by the colour distribution of the lighting environment.

3. Human colour constancy works well in a complex environment where illuminations hitting an object vary as a function of incident angle. If the visual system achieves this with simple illuminant estimation strategies, modelling suggest that it does so by using the history of chromatic signals we receive over time.

4. The directional spectral distribution in natural environments is large and systematic, and the trends in outdoor scenes and indoor scenes are different. The collected database makes it possible to characterise the spatio-spectral nature of visual inputs in our everyday life.

5. The complexities of illumination not only cause difficulties for human colour vision, they can play an important role in compensating the biological limitation in human colour vision system by helping us to solve illuminant metamerism.

6. It is critical for studies of human colour vision to use realistic experimental stimuli that truly reflect the complex aspects of the real world. It is now possible to implement such stimuli thanks to technological advances (for example, in computer graphics techniques), and further studies along these lines are expected to expand knowledge in the field in a way that past classic studies might have missed.

7.7 Future extensions of the current project and concluding remarks

None of the experiments we conducted in this project included motion cues. There are several studies that have suggested that motion can be a useful cue for perception of surface properties such as shape [243], [244], gloss [245] and potentially aid colour constancy [80], [149]. In everyday life, the geometrical relationship between an object, the illumination and our eye position constantly changes. When any of those factors changes, our retinal image can accordingly change dramatically. Thus, it is possible that our visual system evolved and learned a way to utilise motion cues to solve visual tasks. For example, specular reflection and diffuse reflection moves at different speed as an object or our head moves, and thus this might help us to differentiate which part of the object surface is dominated by specular reflection or diffuse reflection [149]. In our reflectance identification experiment, observers' performance changed little when specular reflections were presented, but the addition of a motion cue may possibly induce the effect.

Our analysis of metamerism under environmental illumination provided a new insight into human colour vision. We found around on average 50% of metameric pairs can be disambiguated using directional spectral variation. However, are there any other environments that reduce the impact of metamerism further? If yes, what sort of factors determine the size of the effect? Ultimately, can we design a lighting environment in a way to diminish the disadvantage of dichromats? These questions can be addressed by theoretical analysis of computer rendered objects. Such findings are not only important for the community of colour vision research, but can be also directly applied to colour universal design of lighting environments to aid colour deficient observers.

The effect of directional dependence exhibited by environmental illumination has received little attention in past studies, and thus many questions in colour constancy remain unanswered in these environments. It might be interesting for example to revisit the classic colour constancy experiment with more complex stimuli. For example, achromatic adjustment of objects would answer questions about the point to which human visual systems adapt. Under the classical single colour illumination, the visual

system's adaptation was simple and often found to be somewhere close to the colour of the illumination. However, under the environmental illumination that offers many candidate illuminant points, where does visual system adapt to? Is it simply the mean chromaticity or less apparent chromatic signals hidden in environmental illuminations? Our computational modeling in the reflectance identification (chapter 4) suggested that the visual system seems to adapt to the temporal and spatial average of cone signals. However, direct ways of measuring the neutral point of the visual system may provide further evidence.

This thesis provide three experimental investigations, one set of hyperspectral measurements of environmental illumination maps, and theoretical work based on measured hyperspectral data, using computer rendered objects that reflect some of the complex properties of surface and illumination that the natural environment holds. I believe that the projects provided maybe a small, but hopefully important, step toward understanding the mechanisms that support human colour vision, especially colour constancy for glossy objects under complex lighting environments.

References

[1] B. C. Regan, C. Julliot, B. Simmen, F. Viénot, P. Charles-Dominique, and J. D. Mollon, "Fruits, foliage and the evolution of primate colour vision", *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 356, no. 1407, pp. 229–283, 2001.

[2] W. A. Rushton, "Pigments and signals in colour vision", *The Journal of physiology*, vol. 220, no. 3, pp. 1–31, 1972.

[3] R. M. Boynton, *Human Color Vision*. Holt, Rinehart and Winston, 1979.

[4] L. M. Hurvich and D. Jameson, "An opponent-process theory of color vision", *Psychological Review*, vol. 64, no. 6, pp. 384–404, 1957.

[5] R. H. Masland, "The fundamental plan of the retina", *Nature Neuroscience*, vol. 4, no. 9, pp. 877–886, 2001.

[6] D. M. Dacey and B. B. Barry, "The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type", *Nature*, vol. 367, no. 24, pp. 731–735, 1994.

[7] D. M. Dacey, "Circuitry for color coding in the primate retina", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 2, pp. 582–588, 1996.

[8] J. Krauskopf, D. R. Williams, and D. W. Heeley, "Cardinal Directions of color space", *Vision Research*, vol. 22, pp. 1123–1131, 1982.

[9] R. L. De Valois and K. De Valois, "A Multi-Stage Color Model", *Vision Research*, vol. 33, no. 8, pp. 1053–1065, 1993.

[10] M. S. Livingstone and D. H. Hubel, "Anatomy and physiology of a color system in the primate visual cortex", *Journal of Neuroscience*, vol. 4, no. 1, pp. 309–356, 1984.

[11] Y. Xiao, Y. Wang, and D. J. Felleman, "A spatially organized representation of colour in macaque cotical area V2", *Nature*, vol. 421, no. 6922, pp. 535–539, 2003.

[12] P. H. Schiller, "On the specificity of neurons and visual areas", *Behavioural Brain Research*, vol. 76, no. 1-2, pp. 21–35, 1996.

[13] K. R. Gegenfurtner, "Cortical mechanisms of colour vision", *Nature Reviews Neuroscience*, vol. 4, no. 7, pp. 563–572, 2003.

[14] S. G. Solomon and P. Lennie, "The machinery of colour vision", *Nature Reviews Neuroscience*, vol. 8, no. 4, pp. 276–286, 2007.

[15] B. R. Conway, "Color vision, cones, and color-coding in the cortex", *Neuroscientist*, vol. 15, no. 3, pp. 274–290, 2009.

[16] A. M. Derrington, J. Krauskopf, and P. Lennie, "Chromatic mechanisms in lateral geniculate nucleus of macaque", *The Journal of Physiology*, vol. 357, no. 1, pp. 241–265, 1984.

[17] D. I. A. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance.", *Journal of the Optical Society of America*, vol. 69, no. 8, pp. 1183–1186, 1979.

[18] 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, no. 13, pp. 1711–1737, 2000.

[19] G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae (2nd edition)*. Wiley-Blackwell, 2000.

[20] R. M. Lou, G. Cui, and C. Li, "Uniform colour spaces based on CIECAM02 colour apperance model", *Color Research and Application*, vol. 31, no. 4, pp. 320–330, 2006.

[21] D. Jameson and L. M. Hurvich, "Perceived color and its dependence on focal, surrounding, and preceding stimulus variables", *Journal of the Optical Society of America*, vol. 49, pp. 890–898, 1959.

[22] S. K. Shevell, "Dual Role of Chromatic Backgrounds", *Vision Research*, vol. 18, no. 1971, pp. 1649–1661, 1978.

[23] K. T. Blackwell and G. Buchsbaum, "The effect of spatial and chromatic parameters on chromatic induction", *Color Research & Application*, vol. 13, no. 3, pp. 166–173, 1988.

[24] O. Creutzfeldt, B. Lange-Malecki, and E. Dreyer, "Chromatic induction and brightness contrast: a relativistic color model", *Journal of the Optical Society of America. A, Optics and image science*, vol. 7, no. 9, pp. 1644–1653, 1990.

[25] M. F. Wesner and S. K. Shevell, "Color perception within a chromatic context: the effect of short-wavelength light on color appearance", *Vision Research*, vol. 34, no. 3, pp. 359–365, 1994.

[26] R. O. Brown and D. I. A. MacLeod, "Color appearance depends on the variance of surround colors", *Current Biology*, vol. 7, no. 11, pp. 844–849, 1997.

[27] R. B. Lotto and D. Purves, "An empirical explanation of color contrast", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 23, pp. 12 834–12 839, 2000.

[28] G. Monge, "Memoire sur quelques phénomenes de la vision", *Annales de Chimie*, vol. 3, pp. 131–147, 1789.

[29] L. Arend and A. Reeves, "Simultaneous color constancy", *Journal of the Optical Society of America. A, Optics and image science*, vol. 3, no. 10, pp. 1743–1751, 1986.

[30] H. E. Smithson, "Sensory, computational and cognitive components of human colour constancy", *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1458, pp. 1329–1346, 2005.

[31] D. H. Foster, "Color constancy", *Vision Research*, vol. 51, no. 7, pp. 674–700, 2011.

[32] S. K. Shevell, *The Science of Color*. Elsevier Science, 2003.

[33] R. Lafer-Sousa, H. K. L., and B. R. Conway, "Striking individual differences in color perception uncovered by 'the dress' photograph", *Current Biology*, vol. 25, R523–R548, 2015.

[34] A. D. Winkler, L. Spillmann, J. S. Werner, and M. A. Webster, "Asymmetries in blue-yellow color perception and in the color of 'the dress'", *Current Biology*, vol. 25, no. 13, R547–R548, 2015.

[35] M. Melgosa, L. Gõmez-Robledo, M. Isabel Suero, and M. D. Fairchild, "What can we learn from a dress with ambiguous colors?", *Color Research and Application*, vol. 40, no. 5, pp. 525–529, 2015.

[36] L. Schlaffke, A. Golisch, L. M. Haag, M. Lenz, S. Heba, S. Lissek, T. Schmidt-Wilcke, U. T. Eysel, and M. Tegenthoff, "The brain's dress code: How The Dress allows to decode the neuronal pathway of an optical illusion", *Cortex*, vol. 73, pp. 271–275, 2015.

[37] K. R. Gegenfurtner, M. Bloj, and M. Toscani, "The many colours of 'the dress'", *Current Biology*, vol. 25, no. 13, R543–R544, 2015.

[38] D. H. Brainard and A. C. Hurlbert, "Colour Vision : Understanding #TheDress", *Current Biology*, vol. 25, R549–R568, 2015.

[39] V. M. Hesslinger and C. C. Carbon, "TheDress: The role of illumination information and individual differences in the psychophysics of perceiving white-blue ambiguities", *i-Perception*, vol. 7, no. 2, pp. 1–10, 2016.

[40] K. Rabin, B. Houser, C. Talbert, and R. Patel, "Blue-black or white-gold? Early stage processing and the color of 'the dress'", *PLoS ONE*, vol. 11, no. 8, pp. 1–10, 2016.

[41] K. Vemuri, K. Bisla, S. Mulpuru, and S. Varadharajan, "Do normal pupil diameter differences in the population underlie the color selection of #thedress?", *Journal of the Optical Society of America A*, vol. 33, no. 3, A137–A142, 2016.

[42] S. Aston and A. Hurlbert, "What #theDress reveals about the role of illumination priors in color perception and color constancy", *Journal of Vision*, vol. 17, no. 1–18, p. 4, 2017.

[43] L. D. Daoudi, M. Kunchulia, and M. H. Herzog, "The role of one-shot learning in # TheDress", *Journal of Vision*, vol. 17, pp. 1–7, 2017.

[44] E. L. Dixon and A. G. Shapiro, "Spatial filtering, color constancy, and the color-changing dress", *Journal of Vision*, vol. 17, no. 3, pp. 1–20, 2017.

[45] L. Hugrass, J. Slavikova, M. Horvat, A. A. Musawi, and D. Crewther, "Temporal brightness illusion changes color perception of 'the dress'", *Journal of Vision*, vol. 17, no. 5, pp. 1–7, 2017.

[46] R. Lafer-Sousa and B. R. Conway, "#TheDress: Categorical perception of an ambiguous color image", *Journal of Vision*, vol. 17, no. 12, pp. 1–30, 2017.

[47] O. A. Mahroo, K. M. Williams, I. T. Hossain, E. Yonova-Doing, D. Kozareva, A. Yusuf, I. Sheriff, M. Oomerjee, T. Soorma, and C. J. Hammond, "Do twins share the same dress code? Quantifying relative genetic and environmental contributions to subjective perceptions of 'the dress' in a classical twin study", *Journal of Vision*, vol. 17, no. 1, pp. 1–7, 2017.

[48] M. Toscani, K. R. Gegenfurtner, and K. Doerschner, "Differences in illumination estimation in #thedress", *Journal of Vision*, vol. 17, no. 1, pp. 1–14, 2017.

[49] K. Uchikawa, T. Morimoto, and T. Matsumoto, "Understanding individual differences in color appearance of '#TheDress' based on the optimal color hypothesis", *Journal of Vision*, vol. 17, no. 8, pp. 1–14, 2017.

[50] P. Wallisch, "Illumination assumptions account for individual differences in the perceptual interpretation of a profoundly ambiguous stimulus in the color domain: 'The dress'", *Journal of Vision*, vol. 17, no. 4, pp. 1–14, 2017.

[51] C. Witzel, C. Racey, and J. K. O'Regan, "The most reasonable explanation of "the dress": Implicit assumptions about illumination", *Journal of Vision*, vol. 17, no. 2, pp. 1–19, 2017.

[52] E. H. Land, "The Retinex Theory of Color Vision", *Sci. Am.*, vol. 237, no. 6, pp. 108–128, 1977.

[53] G. Buchsbaum, "A spatial processor model for object colour perception", *Journal of the Franklin Institute*, vol. 310, no. 1, pp. 1–26,

[54] K. Uchikawa, K. Fukuda, Y. Kitazawa, and D. I. A. MacLeod, "Estimating illuminant color based on luminance balance of surfaces", *Journal of the Optical Society of America A*, vol. 29, no. 2, A133–A143, 2012.

[55] J. Golz and D. I. A. MacLeod, "Influence of scene statistics on colour constancy", *Nature*, vol. 415, no. 6872, pp. 637–640, 2002.

[56] K. Fukuda and K. Uchikawa, "Color constancy in a scene with bright colors that do not have a fully natural surface appearance", *Journal of the Optical Society of America A*, vol. 31, no. 4, A239–A246, 2014.

[57] T. Morimoto, K. Fukuda, and K. Uchikawa, "Effects of surrounding stimulus properties on color constancy based on luminance balance", *Journal of the Optical Society of America A*, vol. 33, no. 3, A214–A227, 2016.

[58] J. J. Granzier, E. Brenner, F. W. Cornelissen, and J. B. Smeets, "Luminance-color correlation is not used to estimate the color of the illumination", *Journal of vision*, vol. 5, no. 1, pp. 20–27, 2005.

[59] J. Golz, "The role of chromatic scene statistics in color constancy: Spatial integration", *Journal of Vision*, vol. 8, no. 13, pp. 1–16, 2008.

[60] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms - Part I: Methodology and experiments with synthesized data", *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 972–984, 2002.

[61] G. Arjan, G. Theo, and J. van de Weijer, "Computational Color Constancy: Survey and Experiments", *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 2475–2489, 2011.

[62] J. T. Barron, "Convolutional color constancy", *Proceedings of the IEEE International Conference on Computer Vision*, pp. 379–387, 2015.

[63] S. Bianco, C. Cusano, and R. Schettini, "Color constancy using CNNs", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, no. 2, pp. 81–89, 2015.

[64] D. H. Brainard and W. T. Freeman, "Bayesian color constancy", *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 14, no. 7, pp. 1393–1411, 1997.

[65] D. H. Brainard, P. Longère, P. B. Delahunt, W. T. Freeman, J. M. Kraft, and B. Xiao, "Bayesian model of human color constancy", *Journal of Vision*, vol. 6, no. 11, pp. 1267–1281, 2006.

[66] L. T. Maloney and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance", *Journal of the Optical Society of America A*, vol. 3, no. 1, pp. 29–33, 1986.

[67] L. Maloney, *Color Vision: From Genes to Perception*. Cambridge University Press, 2000, pp. 387–422.

[68] D. H. Brainard, W. A. Brunt, and J. M. Speigle, "Color constancy in the nearly natural image I. Asymmetric matches", *Journal of the Optical Society of America A*, vol. 14, no. 9, pp. 2091–2110, 1997.

[69] J. von Kries, *Theoretische Studien über die Umstimmung des Sehorgans*. Holt, Rinehart and Winston, 1902, pp. 143–158.

[70] M. Bloj, C. Ripamonti, K. Mitha, R. Hauck, S. Greenwald, and D. H. Brainard, "An equivalent illuminant model for the effect of surface slant on perceived lightness", *Journal of Vision*, vol. 4, no. 9, pp. 735–746, 2004.

[71] D. H. Brainard and L. T. Maloney, "Surface color perception and equivalent illumination models", *Journal of Vision*, vol. 11, no. 5, pp. 1–18, 2011.

[72] H. Boyaci, K. Doerschner, J. L. Snyder, and L. T. Maloney, "Surface color perception in three-dimensional scenes", *Visual Neuroscience*, vol. 23, no. 3-4, pp. 311–321, 2006.

[73] M. Hedrich, M. Bloj, and A. I. Ruppertsberg, "Color constancy improves for real 3D objects", *Journal of vision*, vol. 9, no. 4, pp. 1–16, 2009.

[74] J. M. Kraft, S. I. Maloney, and D. H. Brainard, "Surface-illuminant ambiguity and color constancy: Effects of scene complexity and depth cues", *Perception*, vol. 31, no. 2, pp. 247–263, 2002.

[75] Y. Mizokami, M. Ikeda, and H. Shinoda, "Color constancy in a photograph perceived as a three-dimensional space", *Optical Review*, vol. 11, no. 4, pp. 288–296, 2004.

[76] T. Morimoto, Y. Mizokami, H. Yaguchi, and S. L. Buck, "Color Constancy in Two-Dimensional and Three-Dimensional Scenes: Effects of Viewing Methods and Surface Texture", *i-Perception*, vol. 8, no. 6, pp. 1–20, 2017.

[77] S. M. C. Nascimento, V. M. N. de Almeida, P. T. Fiadeiro, and D. H. Foster, "Effect of scene complexity on colour constancy with real three-dimensional scenes and objects", *Perception*, vol. 34, no. 6, pp. 940–950, 2005.

[78] B. Xiao, B. Hurst, L. MacIntyre, and D. H. Brainard, "The color constancy of three-dimensional objects", *J. Vis.*, vol. 12, no. 2012, pp. 1–15, 2012.

[79] J. N. Yang and S. K. Shevell, "Stereo disparity improves color constancy", *Vision Research*, vol. 42, no. 16, pp. 1979–1989, 2002.

[80] A. Werner, "Color constancy improves, when an object moves: High-level motion influences color perception", *Journal of Vision*, vol. 7, no. 14, pp. 1–14, 2007.

[81] L. Wan and K. Shinomori, "Possible influences on color constancy by motion of color targets and by attention-controlled gaze", *Journal of the Optical Society of America A*, vol. 35, no. 4, B309–B323, 2018.

[82] J. J. M. Granzier and K. R. Gegenfurtner, "Effects of memory colour on colour constancy for unknown coloured objects", *i-Perception*, vol. 3, no. 3, pp. 190–215, 2012.

[83] T. Hansen, M. Olkkonen, S. Walter, and K. R. Gegenfurtner, "Memory modulates color appearance", *Nature Neuroscience*, vol. 9, no. 11, pp. 1367–1368, 2006.

[84] E. W. Jin and S. K. Shevell, "Color memory and color constancy", *Journal of the Optical Society of America A*, vol. 13, no. 10, pp. 1981–1991, 1996.

[85] M. Olkkonen, T. Hansen, and K. R. Gegenfurtner, "Color appearance of familiar objects: effects of object shape, texture, and illumination changes", *Journal of vision*, vol. 8, no. 5, pp. 1–16, 2008.

[86] C. Witzel, H. Valkova, T. Hansen, and K. R. Gegenfurtner, "Object knowledge modulates colour appearance", *i-Perception*, vol. 2, no. 1, pp. 13–49, 2011.

[87] H. E. Smithson and Q. Zaidi, "Colour constancy in context: Roles for local adaptation and levels of reference", *Journal of Vision*, vol. 4, no. 9, pp. 693–710, 2004.

[88] H. E. Ives, "The Relation between the Color of the Illuminant and the Color of the illuminated Object", *Transactions of Illuminating Enginerring Society*, vol. 7, pp. 62–72, 1912.

[89] B. B. Lee, D. M. Dacey, V. C. Smith, and J. Pokorny, "Horizontal cells reveal cone type-specific adaptation in primate retina", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 25, pp. 14 611–14 616, 1999.

[90] O. D. Creutzfeldt, S. Kastner, X. Pei, and A. Valberg, "The neurophysiological correlates of colour and brightness contrast in lateral geniculate neurons - II. Adaptation and surround effects", *Experimental Brain Research*, vol. 87, no. 1, pp. 22–45, 1991.

[91] J. L. Barbur and K. Spang, "Colour constancy and conscious perception of changes of illuminant", *Neuropsychologia*, vol. 46, no. 3, pp. 853–863, 2008.

[92] S. Zeki, "The representation of colours in the cerebral cortex", *Nature*, vol. 284, no. 5755, pp. 412–418, 1980.

[93] S. J. Schein and R. Desimone, "Spectral properties of V4 neurons in the macaque", *Journal of Neuroscience*, vol. 10, no. 10, pp. 3369–3389, 1990.

[94] S. Zeki, "Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours", *Neuroscience*, vol. 9, no. 4, pp. 741–765, 1983.

[95] M. Kusunoki, K. Moutoussis, and S. Zeki, "Effect of background colors on the tuning of color-selective cells in monkey area V4", *Journal of Neurophysiology*, vol. 95, no. 5, pp. 3047–3059, 2006.

[96] H. M. Wild, S. R. Butler, D. Carden, and J. J. Kulikowski, "Primate cortical area V 4 important for colour constancy but not wavelength discrimination", *Nature*, vol. 313, no. 10, pp. 133–135, 1985.

[97] V. Walsh, D. Carden, S. R. Butler, and J. J. Kulikowski, "The effects of V4 lesions on the visual abilities of macaques: hue discrimination and colour constancy", *Behavioural Brain Research*, vol. 53, no. 1-2, pp. 51–62, 1993.

[98] A. Hurlbert, "Colour vision: Primary visual cortex shows its influence", *Current Biology*, vol. 13, no. 7, pp. 270–272, 2003.

[99] F. W. Cornelissen, J. J. van Es, and T. Vladusich, "Fmri of relational color constancy in human visual cortex [vss meeting abstract]", *Journal of Vision*, vol. 6, no. 6, p. 389, 2006.

[100] E. Baumgartner, D. Weiss, and K. R. Gegenfurtner, "Decoding color constancy in fmri [vss meeting abstract]", *Journal of Vision*, vol. 16, no. 12, p. 392, 2016.

[101] L. E. Arend, A. Reeves, J. Schirillo, and R. Goldstein, "Simultaneous color constancy: papers with diverse Munsell values", *Journal of the Optical Society of America A*, vol. 8, no. 4, pp. 661–672, 1991.

[102] D. H. Brainard and B. A. Wandell, "Asymmetric color matching: how color appearance depends on the illuminant", *Journal of the Optical Society of America A*, vol. 9, no. 9, pp. 1433–1443, 1992.

[103] F. W. Cornelissen and E. Brenner, "Simultaneous colour constancy revisited: an analysis of viewing strategies", *Vision Research*, vol. 35, no. 17, pp. 2431–2448, 1995.

[104] I. Kuriki and K. Uchikawa, "Limitations of surface-color and apparent-color", vol. 13, no. 8, pp. 1622–1636, 1996.

[105] K. H. Bäuml, "Simultaneous color constancy: How surface color perception varies with the illuminant", *Vision Research*, vol. 39, no. 8, pp. 1531–1550, 1999.

[106] J. M. Troost and C. M. M. De Weert, "Naming versus matching in color constancy", *Perception & Psychophysics*, vol. 50, no. 6, pp. 591–602, 1991.

[107] D. H. Foster, S. M. C. Nascimento, K. Amano, L. Arend, K. J. Linnell, J. L. Nieves, S. Plet, and J. S. Foster, "Parallel detection of violations of color constancy", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 14, pp. 8151–8156, 2001.

[108] V. M. N. De Almeida, P. T. Fiadeiro, and S. M. C. Nascimento, "Color constancy by asymmetric color matching with real objects in three-dimensional scenes", *Visual Neuroscience*, vol. 21, no. 3, pp. 341–345, 2004.

[109] I. J. Murray, A. Daugirdiene, H. Vaitkevicius, J. J. Kulikowski, and R. Stanikunas, "Almost complete colour constancy achieved with full-field adaptation", *Vision Research*, vol. 46, no. 19, pp. 3067–3078, 2006.

[110] A. Reeves, K. Amano, and D. H. Foster, "Color constancy: Phenomenal or projective?", *Perception & psychophysics*, vol. 70, no. 2, pp. 219–228, 2008.

[111] J. S. Werner and J. Walraven, "Effect of chromatic adaptation on the achromatic locus: The role of contrast, luminance and background color", *Vision Research*, vol. 22, no. 8, pp. 929–943, 1982.

[112] M. D. Fairchild and P. Lennie, "Chromatic adaptation to natural and incandescent illuminants", *Vision Research*, vol. 32, no. 11, pp. 2077–2085, 1992.

[113] L. E. Arend, "How much does illuminant color affect unattributed colors?", *Journal of the Optical Society of America A*, vol. 10, no. 10, pp. 2134–2147, 1993.

[114] J. M. Kraft and D. H. Brainard, "Mechanisms of color constancy under nearly natural viewing", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 1, pp. 307–312, 1999.

[115] P. B. Delahunt and D. H. Brainard, "Color constancy under changes in reflected illumination", *Journal of Vision*, vol. 4, no. 9, pp. 764–778, 2004.

[116] ——, "Does human color constancy incorporate the statistical regularity of natural daylight?", *Journal of vision*, vol. 4, no. 2, pp. 57–81, 2004.

[117] J. Golz, "Colour Constancy: Influence of Viewing Behaviour on Grey Settings", *Perception*, vol. 39, no. 5, pp. 606–619, 2010.

[118] Y. Mizokami and H. Yaguchi, "Color constancy influenced by unnatural spatial structure", *Journal of the Optical Society of America A*, vol. 31, no. 4, A179–A186, 2014.

[119] T. Hansen, S. Walter, and K. R. Gegenfurtner, "Effects of spatial and temporal context on color categories and color constancy", *Journal of Vision*, vol. 7, no. 4, pp. 1–15, 2007.

[120] R. Ma, N. Liao, P. Yan, and K. Shinomori, "Categorical color constancy under RGB-LED light sources", *Color Research and Application*, vol. 43, no. 5, pp. 655–674, 2018.

[121] M. Olkkonen, C. Witzel, T. Hansen, and K. R. Gegenfurtner, "Categorical color constancy for real surfaces", *Journal of Vision*, vol. 10, no. 9, pp. 1–22, 2010.

[122] S. Schultz, K. Doerschner, and L. T. Maloney, "Color constancy and hue scaling", *Journal of Vision*, vol. 6, no. 10, pp. 1102–1116, 2006.

[123] B. Berlin and P. Kay, *Basic color terms*. Berkeley:University of California Press, 1969.

[124] R. M. Boynton and C. X. Olson, "Locating basic colors in the OSA space", *Color Research & Application*, vol. 12, no. 2, pp. 94–105, 1987.

[125] D. H. Foster, B. J. Craven, and E. R. H. Sale, "Immediate colour constancy", *Ophthalmic and Physiological Optics*, vol. 12, no. 2, pp. 157–160, 1992.

[126] B. J. Craven and D. H. Foster, "An operational approach to colour constancy", *Vision Research*, vol. 32, no. 7, pp. 1359–1366, 1992.

[127] V. M. N. de Almeida, P. T. Fiadeiro, and S. M. C. Nascimento, "Effect of scene dimensionality on colour constancy with real three-dimensional scenes and objects", *Perception*, vol. 39, no. 6, pp. 770–779, 2010.

[128] R. C. Baraas, D. H. Foster, K. Amano, and S. M. C. Nascimento, "Color constancy of red-green dichromats and anomalous trichromats", *Investigative Ophthalmology and Visual Science*, vol. 51, no. 4, pp. 2286–2293, 2010.

[129] R. J. Lee and H. E. Smithson, "Low levels of specularity support operational color constancy, particularly when surface and illumination geometry can be inferred", *Journal of the Optical Society of America A*, vol. 33, no. 3, A306–A316, 2016.

[130] A. Werner, "The influence of depth segmentation on colour constancy", *Perception*, vol. 35, no. 9, pp. 1171–1184, 2006.

[131] A. C. Chadwick and R. W. Kentridge, "The perception of gloss: A review", *Vision Research*, vol. 109, no. PB, pp. 221–235, 2015.

[132] R. W. Fleming, "Visual perception of materials and their properties", *Vision research*, vol. 94, no. 24, pp. 62–75, 2014.

[133] L. T. Maloney and D. H. Brainard, "Color and material perception: Achievements and challenges", *Journal of Vision*, vol. 10, no. 9, p. 19, 2010.

[134] J. Berzhanskaya, G. Swaminathan, J. Beck, and E. Mingolla, "Remote effects of highlights on gloss perception", *Perception*, vol. 34, no. 5, pp. 565–575, 2005.

[135] K. Doerschner, L. T. Maloney, and H. Boyaci, "Perceived glossiness in high dynamic range scenes", *Journal of Vision*, vol. 10, no. 9, pp. 1–11, 2010.

[136] M. Giesel and K. R. Gegenfurtner, "Color appearance of real objects varying in material, hue, and shape", *Journal of vision*, vol. 10, no. 9, pp. 1–21, 2010.

[137] J. J. M. Granzier, R. Vergne, and K. R. Gegenfurtner, "The effects of surface gloss and roughness on color constancy for real 3-D objects", *Journal of vision*, vol. 14, no. 2, pp. 1–20, 2014.

[138] Y. Ling and A. Hurlbert, "Color and size interactions in a real 3D object similarity task", *Journal of Vision*, vol. 4, no. 9, pp. 721–734, 2004.

[139] A. Radonjić, N. P. Cottaris, and D. H. Brainard, "Color constancy supports cross-illumination color selection", *Journal of Vision*, vol. 15, no. 6, pp. 1–19, 2015.

[140] A. Radonjic, N. P. Cottaris, and D. H. Brainard, "Color constancy in a naturalistic goal-directed task", *Journal of Vision*, vol. 15, no. 13, pp. 1–21, 2015.

[141] J. N. Yang and S. K. Shevell, "Surface color perception under two illuminants: The second illuminant reduces color constancy", *Journal of Vision*, vol. 3, no. 5, pp. 369–379, 2003.

[142] G. J. Ward, "Measuring and modeling anisotropic reflection", *ACM SIGGRAPH Computer Graphics*, vol. 26, no. 2, pp. 265–272, 1992.

[143] B. Xiao and D. H. Brainard, "Surface gloss and color perception of 3D objects", *Visual Neuroscience*, vol. 25, no. 3, pp. 371–385, 2008.

[144] K. Amano and D. H. Foster, "Colour constancy under simultaneous changes in surface position and illuminant", *Proceedings of the Royal Society B: Biological Sciences*, vol. 271, no. 1555, pp. 2319–2326, 2004.

[145] B. G. Khang and Q. Zaidi, "Illuminant color perception of spectrally filtered spotlights", *Journal of Vision*, vol. 4, no. 9, pp. 680–692, 2004.

[146] K. J. Linnell and D. H. Foster, "Scene articulation: Dependence of illuminant estimates on number of surfaces", *Perception*, vol. 31, no. 2, pp. 151–159, 2002.

[147] J. J. McCann, "Rules for colour constancy", *Ophthalmic and Physiological Optics*, vol. 12, no. 2, pp. 175–177, 1992.

[148] S. Tominaga, S. Ebisui, and B. A. Wandell, "Scene illuminant classification: brighter is better", *Journal of the Optical Society of America A*, vol. 18, no. 1, pp. 55–64, 2001.

[149] R. J. Lee and H. E. Smithson, "Motion of glossy objects does not promote separation of lighting and surface colour", *Royal Society Open Science*, vol. 4, pp. 1–16, 2017.

[150] M. Pharr, W. Jakob, and G. Humphreys, *Physically Based Rendering: From Theory to Implementation 3rd Edition*. Morgan Kaufmann, 2016, pp. 1–1266.

[151] B. S. Heasly, N. P. Cottaris, D. P. Lichtman, B. Xiao, and D. H. Brainard, "Rendertoolbox3: MATLAB tools that facilitate physically based stimulus rendering for vision research", *Journal of Vision*, vol. 14, no. 2, pp. 1–22, 2014.

[152] K. Amano, D. H. Foster, and S. M. C. Nascimento, "Color constancy in natural scenes with and without an explicit illuminant cue.", *Visual Neuroscience*, vol. 23, no. 3-4, pp. 351–356, 2006.

[153] L. Chittka, A. Shmida, N. Troje, and R. Menzel, "Ultraviolet as a component of flower reflections, and the colour perception of hymenoptera", *Vision Research*, vol. 34, no. 11, pp. 1489–1508, 1994.

[154] M. J. Vrhel, R. Gershon, and L. S. Iwan, "Measurement and Analysis of Object Reflectance Spectra", *Color Research & Application*, vol. 19, no. 1, pp. 4–9, 1994.

[155] N. J. Marshall, "Communication and camouflage with the same 'bright' colours in reef fishes", *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 355, no. 1401, pp. 1243–1248, 2000.

[156] P. J. Marlow, J. Kim, and B. L. Anderson, "The perception and misperception of specular surface reflectance", *Current Biology*, vol. 22, no. 20, pp. 1909–1913, 2012.

[157] J. Kim, P. Marlow, and B. L. Anderson, "The perception of gloss depends on highlight congruence with surface shading", *Journal of Vision*, vol. 11, no. 9, pp. 1–19, 2011.

[158] I. Motoyoshi, S. Nishida, L. Sharan, and E. H. Adelson, "Image statistics and the perception of surface qualities", *Nature*, vol. 447, no. 7141, pp. 206–209, 2007.

[159] S. Nishida and M. Shinya, "Use of image-based information in judgments of surface-reflectance properties", *Journal of the Optical Society of America A*, vol. 15, no. 12, pp. 2951–2965, 1998.

[160] L. Sharan, Y. Li, I. Motoyoshi, S. Nishida, and E. H. Adelson, "Image statistics for surface reflectance perception.", *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 25, no. 4, pp. 846–865, 2008.

[161] J. T. Todd, J. F. Norman, and E. Mingolla, "Lightness constancy in the presence of specular highlights", *Psychological science*, vol. 15, no. 1, pp. 33–39, 2004.

[162] G. Wendt, F. Faul, V. Ekroll, and R. Mausfeld, "Disparity, motion, and color information improve gloss constancy performance.", *Journal of vision*, vol. 10, no. 9, pp. 1–17, 2010.

[163] J. N. Yang and L. T. Maloney, "Illuminant cues in surface color perception: Tests of three candidate cues", *Vision Research*, vol. 41, no. 20, pp. 2581–2600, 2001.

[164] R. J. Lee and H. E. Smithson, "Context-dependent judgments of color that might allow color constancy in scenes with multiple regions of illumination", *Journal of the Optical Society of America A*, vol. 29, no. 2, A247–A257, 2012.

[165] H. Boyaci, K. Doerschner, and L. T. Maloney, "Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity", *Journal of vision*, vol. 4, no. 9, pp. 664–679, 2004.

[166] K. Doerschner, H. Boyaci, and L. T. Maloney, "Human observers compensate for secondary illumination originating in nearby chromatic surfaces", *Journal of Vision*, vol. 4, no. 2, pp. 92–105, 2004.

[167] R. O. Dror, A. S. Willsky, and E. H. Adelson, "Statistical characterization of real-world illumination", *Journal of Vision*, vol. 4, no. 9, pp. 821–837, 2004.

[168] M. Olkkonen and D. H. Brainard, "Perceived glossiness and lightness under real-world illumination", *Journal of Vision*, vol. 10, no. 9, pp. 1–19, 2010.

[169] I. Motoyoshi and H. Matoba, "Variability in constancy of the perceived surface reflectance across different illumination statistics", *Vision Research*, vol. 53, no. 1, pp. 30–39, 2012.

[170] Y. Morgenstern, W. S. Geisler, and R. F. Murray, "Human vision is attuned to the diffuseness of natural light", *Journal of Vision*, vol. 14, no. 9, pp. 1–18, 2014.

[171] R. W. Fleming, R. O. Dror, and E. H. Adelson, "Real-world illumination and the perception of surface reflectance properties", *Journal of Vision*, vol. 3, no. 3, pp. 347–368, 2003.

[172] K. Doerschner, H. Boyaci, and L. T. Maloney, "Testing limits on matte surface color perception in three-dimensional scenes with complex light fields", *Vision Research*, vol. 47, no. 28, pp. 3409–3423, 2007.

[173] K. R. Gegenfurtner and D. C. Kiper, "Contrast detection in luminance and chromatic noise", *J Opt. Soc. Am. A.*, vol. 9, no. 11, pp. 1880–1888, 1992.

[174] B. Pearce, S. Crichton, M. Mackiewicz, G. D. Finlayson, and A. Hurlbert, "Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations", *PLoS ONE*, vol. 9, no. 2, pp. 1–10, 2014.

[175] A. Radonjic, B. Pearce, S. Aston, A. Krieger, H. Dubin, N. P. Cottaris, D. H. Brainard, and A. C. Hurlbert, "Illumination discrimination in real and simulated scenes", *Journal of Vision*, vol. 16, no. 11, pp. 1–18, 2016.

[176] W. Jakob, *Mitsuba renderer*. [Online]. Available: `https://www.mitsuba-renderer.org`.

[177] V. B., *Light probes*. [Online]. Available: `http://dativ.at/lightprobes/`.

[178] B. Smits, "An RGB to Spectrum Conversion for Reflectances", *J. Graph. Tools*, pp. 11–22, 2000.

[179] F. A. A. Kingdom and N. Prins, *Palamedes: Matlab routines for analyzing psychophysical data*. [Online]. Available: `http://www.palamedestoolbox.org.`.

[180] S. M. C. Nascimento, K. Amano, and D. H. Foster, "Spatial distributions of local illumination color in natural scenes", *Vision Research*, vol. 120, pp. 39–44, 2016.

[181] M. A. Webster and J. D. Mollon, "Adaptation and the Color Statistics of Natural Images", *Vision Research*, vol. 37, no. 23, pp. 3283–3298, 1997.

[182] M. A. Webster, "Adaptation and visual coding", *Journal of Vision*, vol. 11, no. 3, pp. 1–23, 2011.

[183] L. E. Welbourne, A. B. Morland, and A. R. Wade, "Human colour perception changes between seasons", *Current Biology*, vol. 25, no. 15, R646–R647, 2015.

[184] J. Neitz, J. Carroll, Y. Yamauchi, M. Neitz, and D. R. Williams, "Color perception is mediated by a plastic neural mechanism that is adjustable in adults", *Neuron*, vol. 35, no. 4, pp. 783–792, 2002.

[185] D. L. MacAdam, "Visual Sensitivities to Color Differences in Daylight", *J Opt. Soc. Am. A.*, vol. 32, no. 5, pp. 247–274, 1942.

[186] T. Yeh, J. Pokorny, and V. C. Smith, "Chromatic discrimination with variation in chromaticity and luminance: Data and theory", *Vision Research*, vol. 33, no. 13, pp. 1835–1845, 1993.

[187] J. Hernández-Andrés, J. Romero, J. L. Nieves, and R. L. Lee, "Color and spectral analysis of daylight in southern Europe", *Journal of the Optical Society of America A*, vol. 18, no. 6, pp. 1325–1335, 2001.

[188] J. M. Bosten, R. D. Beer, and D. I. A. MacLeod, "What is white?", *Journal of Vision*, vol. 15, no. 16, pp. 1–19, 2015.

[189] R. W. Fleming, A. Torralba, and E. H. Adelson, "Specular reflections and the perception of shape", *Journal of Vision*, vol. 4, no. 9, pp. 798–820, 2004.

[190] A. Hurlbert, "Colour constancy", *Current Biology*, vol. 17, no. 21, pp. 906–907, 2007.

[191] R. J. Lee, K. A. Dawson, and H. E. Smithson, "Slow updating of the achromatic point after a change in illumination", *Journal of Vision*, vol. 12, no. 1, pp. 1–22, 2012.

[192] L. K.Doerschner, H.Boyaci, "Estimating the glossiness transfer function induced by illumination change and testing its transitivity", *Journal of Vision*, vol. 10, no. 4, pp. 1–9, 2010.

[193] T. Morimoto and H. E. Smithson, "Discrimination of spectral reflectance under environmental illumination", *Journal of the Optical Society of America A*, vol. 35, no. 4, B244–B255, 2018.

[194] T. Morimoto, S. Kishigami, M. M. L. Joao, S. M. C. Nascimento, and H. E. Smithson, "Hyperspectral environmental illumination maps : characterizing directional spectral variation in natural environments", *Optics Express*, vol. 27, no. 22, pp. 32 277–32 293, 2019.

[195] D. H. Brainard, "Color constancy in the nearly natural image 2 Achromatic loci", *Journal of the Optical Society of America A*, vol. 15, no. 2, pp. 307–325, 1998.

[196] D. H. Brainard, N. P. Cottaris, and A. Radonjic, "The perception of colour and material in naturalistic tasks", *Interface Focus*, vol. 8, no. 4, pp. 1–12, 2018.

[197] A. Radonjić, N. P. Cottaris, and D. H. Brainard, "The relative contribution of color and material in object selection", *PLoS computational biology*, vol. 15, no. 4, pp. 1–27, 2019.

[198] R. Robilotto and Q. Zaidi, "Lightness identification of patterned three-dimensional, real objects", *Journal of Vision*, vol. 6, no. 1, pp. 18–36, 2006.

[199] Q. Zaidi and M. Bostic, "Color strategies for object identification", *Vision Research*, vol. 48, no. 26, pp. 2673–2681, 2008.

[200] R. Robilotto and Q. Zaidi, "Limits of lightness identification for real objects under natural viewing conditions", *Journal of Vision*, vol. 4, no. 9, pp. 779–797, 2004.

[201] B. Smits, "An rgb-to-spectrum conversion for reflectances", *Journal of Graphics Tools*, vol. 4, no. 4, pp. 11–22, 1999.

[202] Y. Ling and A. Hurlbert, "Role of color memory in successive color constancy: erratum", *Journal of the Optical Society of America A*, vol. 25, no. 6, pp. 1215–1226, 2008.

[203] K. Uchikawa, I. Kuriki, and Y. Tone, "Measurement of color constancy by color memory matching", *Optical Review*, vol. 5, no. 1, pp. 59–63, 1998.

[204] M. Olkkonen and S. R. Allred, "Short-term memory affects color perception in context", *PLoS ONE*, vol. 9, no. 1, pp. 1–11, 2014.

[205] N. Kriegeskorte, "Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing", *Annual Review of Vision Science*, vol. 1, no. 1, pp. 417–446, 2015.

[206] R. M. Cichy, A. Khosla, D. Pantazis, A. Torralba, and A. Oliva, "Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence", *Scientific Reports*, vol. 6, pp. 1–13, 2016.

[207] T. S. A. Wallis, C. M. Funke, A. S. Ecker, L. A. Gatys, F. A. Wichmann, and M. Bethge, "A parametric texture model based on deep convolutional features closely matches texture appearance for humans", *Journal of Vision*, vol. 17, no. 12, pp. 1–29, 2017.

[208] F. A. Wichmann, D. H. J. Janssen, R. Geirhos, G. Aguilar, H. H. Schütt, M. Maertens, and M. Bethge, "Methods and measurements to compare men against machines", *IS and T International Symposium on Electronic Imaging Science and Technology*, pp. 36–45, 2017.

[209] N. J. Majaj and D. G. Pelli, "Deep learning-Using machine learning to study biological vision", *Journal of Vision*, vol. 18, no. 13, pp. 1–13, 2018.

[210] K. W. Houser, M. Wei, A. David, M. R. Krames, and X. S. Shen, "Review of measures for light-source color rendition and considerations for a two-measure system for characterizing color rendition", *Optics Express*, vol. 21, no. 8, pp. 10 393–10 411, 2013.

[211] D. B. Judd, D. L. MacAdam, G. Wyszecki, H. W. Budde, H. R. Condit, S. T. Henderson, and J. L. Simonds, "Spectral distribution of typical daylight as a function of correlated color temperature", *Journal of the Optical Society of America A*, vol. 54, no. 8, pp. 1031–1040, 1964.

[212] J. Hernández-Andrés, J. Romero, J. L. Nieves, and R. L. Lee, "Color and spectral analysis of daylight in southern Europe", *Journal of the Optical Society of America A*, vol. 18, no. 6, pp. 1325–1335, 2001.

[213] J. Tian, Z. Duan, W. Ren, Z. Han, and Y. Tang, "Simple and effective calculations about spectral power distributions of outdoor light sources for computer vision", *Optics Express*, vol. 24, no. 7, pp. 7266–7286, 2016.

[214] M. Spitschan, G. K. Aguirre, D. H. Brainard, and A. M. Sweeney, "Variation of outdoor illumination as a function of solar elevation and light pollution", *Scientific Reports*, vol. 6, pp. 1–14, 2016.

[215] D. Rada, V. W. Joost, M. Damien, H. Mathieu, and T. Alain, "Deep spectral reflectance and illuminant estimation from self-interreflections", *J. Opt. Soc. Am. A*, vol. 36, no. 1, pp. 105–114, 2019.

[216] T. Machida, N. Yokoya, and H. Takemura, "Surface reflectance modeling of real objects with interreflections", in *Proceedings of IEEE International Conference on Computer Vision*, IEEE, 2003, pp. 170–177.

[217] R. D., D. M., M. H., and A. T., "Interreflections in computer vision: A survey and an introduction to spectral infinite-bounce model", *Journal of Mathematical Imaging and Vision*, vol. 60, no. 5, pp. 661–680, 2018.

[218] A. A. Mury, S. C. Pont, and J. J. Koenderink, "Structure of light fields in natural scenes", *Applied Optics*, vol. 48, no. 28, pp. 5386–5395, 2009.

[219] V. S. Ramachandran, "Perception of shape from shading", *Nature*, vol. 331, no. 4, pp. 163–166, 1988.

[220] P. Mamassian and R. Goutcher, "Prior knowledge on the illumination position", *Cognition*, vol. 81, no. 1, B1–B9, 2001.

[221] Y. Morgenstern, R. F. Murray, and L. R. Harris, "The human visual system's assumption that light comes from above is weak", *Proceedings of the National Academy of Sciences*, vol. 108, no. 30, pp. 12 551–12 553, 2011.

[222] P. Debevec, "Rendering synthetic objects into real scenes", in *Proceedings of the ACM SIGGRAPH*, ACM, 1998, pp. 19–24.

[223] S. M. C. Nascimento, F. P. Ferreira, and D. H. Foster, "Statistics of spatial cone-excitation ratios in natural scenes", *Journal of the Optical Society of America A*, vol. 19, no. 8, pp. 1484–1490, 2002.

[224] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes", *Journal of the Optical Society of America A*, vol. 23, no. 10, pp. 2359–2372, 2006.

[225] A. M. Molinaro, R. Simon, and R. M. Pfeiffer, "Prediction error estimation: A comparison of resampling methods", *Bioinformatics*, vol. 21, no. 15, pp. 3301–3307, 2005.

[226] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann Publishers Inc., 2005.

[227] B. Zitová and J. Flusser, "Image registration methods: A survey", *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[228] Y. Li, Y. Wang, W. Huang, and Z. Zhang, "Automatic image stitching using SIFT", in *Proceedings of IEEE Conference on Audio, Language and Image Processing*, IEEE, 2008, pp. 568–571.

[229] D. H. Foster and K. Amano, "Hyperspectral imaging in color vision research: tutorial", *Journal of the Optical Society of America A*, vol. 36, no. 4, pp. 606–627, 2019.

[230] N. Ranganath and B. R. Mandyam, "The science behind archaeological signatures from space", *Current Science*, vol. 113, no. 10, pp. 105–114, 2017.

[231] R. Ramamoorthi, "An Efficient Representation for Irradiance Environment Maps", in *Proceedings of SIGGRAPH '01 on Computer graphics and interactive techniques*, ACM, 2001, pp. 497–500.

[232] T. Morimoto and H. E. Smithson, "Identifying surface colours across different environmental illuminations", *Perception*, vol. 48, p. 47, 2019.

[233] S. Tominaga, A. Matsuura, and T. Horiuchi, "Spectral Analysis of Omnidirectional Illumination in a Natural Scene", *Journal of Imaging Science and Technology*, vol. 54, no. 4, pp. 405 021–405 029, 2010.

[234] K. Hirai, N. Osawa, M. Hori, T. Horiuchi, and S. Tominaga, "High-Dynamic-Range Spectral Imaging System for Omnidirectional Scene Capture", *Journal of Imaging*, vol. 4, no. 53, pp. 1–20, 2018.

[235] K. R. Gegenfurtner and J. Rieger, "Sensory and cognitive contributions of color to the recognition of natural scenes", *Current Biology*, vol. 10, no. 13, pp. 805–808, 2000.

[236] W. S. Stiles and G. W. Wyszecki, "Counting Metameric Object Colors", *Journal of the Optical Society of America*, vol. 52, no. 3, pp. 313–328, 1962.

[237] A. D. Logvinenko, B. Funt, and C. Godau, "Metamer mismatching", *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 34–43, 2014.

[238] A. D. Logvinenko, B. Funt, H. Mirzaei, and R. Tokunaga, "Rethinking colour constancy", *PLoS ONE*, vol. 10, no. 9, pp. 1–29, 2015.

[239] J. Broackes, *Color Ontology and Color Science*. MIT Press, 2010, pp. 291–406.

[240] H. Brettel, F. Vienot, and J. D. Mollon, "Computerized simulation of color appearance for dichromats", vol. 14, no. 10, pp. 2647–2655, 1997.

[241] D. H. Foster, K. Amano, S. M. Nascimento, and M. J. Foster, "The frequency of metamerism in natural scenes", *J Opt. Soc. Am. A.*, vol. 23, no. 10, pp. 2359–2372, 2006.

[242] A. Akbarinia and K. R. Gegenfurtner, "Color metamerism and the structure of illuminant space", *Journal of the Optical Society of America A*, vol. 35, no. 4, B231–B238, 2018.

[243] D. N. Dövencioğlu, M. W. Wijntjes, O. Ben-Shahar, and K. Doerschner, "Effects of surface reflectance on local second order shape estimation in dynamic scenes", *Vision Research*, vol. 115, pp. 218–230, 2015.

[244] D. N. Dövencioğlu, O. Ben-Shahar, P. Barla, and K. Doerschner, "Specular motion and 3D shape estimation", *Journal of Vision*, vol. 17, no. 6, pp. 1–15, 2017.

[245] K. Doerschner, R. W. Fleming, O. Yilmaz, P. R. Schrater, B. Hartung, and D. Kersten, "Visual motion and the perception of surface material", *Current Biology*, vol. 21, no. 23, pp. 2010–2016, 2011.