Modelling Colour Appearance: Applications in Skin Image Perception

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Tushar Chauhan.

Defended: March 2016

Final submission: November 2016

Abstract

Humans are trichromatic, and yet their perception of colours is rich and complex. The research presented in this thesis explores the process of colour appearance of uniform patches and natural polychromatic stimuli. This is done through the measurement and analysis of the achromatic locus (Chapter 2), modelling of chromatic adaptation in a large dataset of unique hues settings (Chapter 3), and measurement of thresholds for uniform and polychromatic stimuli derived from simulated skin images (Chapter 4).

Chapter 2 proposes a novel navigation scheme based on unique hues for traversing colour space. The results show that when colour adjustments are made using this novel scheme, the variability of achromatic settings made by observers is reduced compared to the classical method of making colour adjustments along the cardinal axes of the $CIELUV$ colour space. This result holds across the tested luminance levels ($5, 20, 50\ cd/m^2$) in each of the three tested ambient illumination conditions – dark, simulated daylight and cool white fluorescent lighting. The analysis also shows that the direction of maximum variance of the achromatic settings lies along the daylight locus.

Chapter 3 evaluates models of chromatic adaptation by using unique hues settings measured under different ambient illumination conditions. It is shown that a simple diagonal model in cone excitation space is the most efficient in terms of the trade-off between accuracy and degrees of freedom. It is also found that diagonal and linear models show similar performances, reiterating their theoretical equivalence. Performances of these diagonalisable models are found to be worse for UR and UG unique hue planes compared to UY and UB planes.

Chapter 4 presents a set of three experiments reporting estimations of perceptual thresholds for polychromatic and uniform stimuli in a 3-D chromaticity-luminance colour space. The first experiment reports thresholds for simulated skin images and uniform stimuli of the corresponding mean $CIELAB$ colour. The second and third

experiments investigate the effect of ambient illumination and the location of the stimuli in colour space. The thresholds for the polychromatic stimuli are found to be consistently higher than those for the uniform patches, for both the chromatic, and the luminance projections. The area of the chromaticity ellipses shows a gradual increase with distance from the illuminant chromaticity. The orientations of these ellipses for simulated skin are found to align with the vector joining the mean patch chromaticity and the illuminant chromaticity.

(suddenly in sunlight

he will bow,

& the whole garden will bow)

- e. e. cummings

Dedicated to my mother.

Acknowledgements

I would like to thank my supervisor Prof. Sophie Wuerger for her constant guidance, support and encouragement throughout the course of this research. I am also grateful to Dr. Kaida Xiao for his invaluable advice, especially in matters regarding colorimetry and device calibration, and to Dr. Yannis Goulermas for his help and suggestions regarding optimisation techniques. I am also indebted to Prof. Anya Hurlbert for her insightful comments and recommendations which have greatly improved not only the quality, but also the presentability of the work.

Finally, I would like to thank my family for their unwavering, unconditional support.

Contents

Chapter 1

Introduction

The seeds of modern colour theory can be seen as early as the works of Newton (Newton, 1730), who explained colour based on light matching experiments. He proposed the concept of complementary colours, which he explained using a circular arrangement of colours akin to modern day colour-circles. In 1802 a model of colour vision based on three-primaries was proposed by Young (Young, 1802). Young's theory was further expounded by Helmholtz (Helmholtz, 1867), who explained trichromatic colour based on the laws of additivity of light which had earlier been proposed by Grassmann (Grassmann, 1853). Maxwell (1860), by proposing a quality and a magnitude to colour, essentially introduced the idea of separable chromatic and luminance components of colour. In doing so, he not only provided evidence in support of Young's trichromatic theory, but also introduced possibility of colour photography. It is these Grassmann and Maxwell laws (not to be confused with his laws of electromagnetism) which form the basis of modern colorimetry and colour theory.

With the discovery of the three cone photoreceptors (which had been predicted by Young in 1802) in the human retina, the trichromatic theory was given a solid physiological basis in terms of the quantum catches in the three photoreceptors. The photosensitive pigments in the three cones, called the L, M and S cones, are reactive to long, medium and short range of visible frequencies respectively. Figure 1.1b shows the plot of the sensitivities for the L, M and S cone classes (Smith & Pokorny, 1975) as a function of the wavelength of incident light.

In modern colorimetry, human trichromatic colour sensitivity is measured by what are known as colour matching functions. Given a set of three primary lights, these functions define the proportions in which these lights must be mixed in order to match the appearance of monochromatic reference lights. It must be noted that these

functions can also have negative values implying that the corresponding primary must be subtracted from the mixture, i.e., added to the reference light. These colour matching functions were first proposed (indirectly) by Maxwell (1860), and have traditionally been measured using either some form of bipartite field matching techniques (Stiles & Burch, 1955) or heterochromatic flicker photometry approaches (Stockman & Sharpe, 2000). In bipartite matching experiments the observer typically matches the colour between two separate fields, a test and a reference field, till they appear the same. In heterochromatic flicker photometry the observer views a stimulus that alternates between a test and a reference colour at a certain frequency, and adjusts the colour of the test stimulus till the flicker is no longer perceptible.

Thus, the modern formulation of the trichromatic theory describes the light reaching an observer's eyes in terms of a three dimensional vector of tristimulus values which describe the proportion in which a given set of primary lights must be mixed to illicit the same response as the stimulus. Since light has a continuous spectrum, this is achieved by projecting the spectrum of incident light on to a three-dimensional space defined by the three colour matching functions which operate on human-perceptible (visible) range of frequencies. If a surface is viewed normally under a light which is also incident normally, the tristimulus coordinates of the light reaching the observer's eye are given by

$$\boldsymbol{X}_i = \int_{visible} \boldsymbol{L}(\lambda) \cdot \boldsymbol{r}(\lambda) \cdot \bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda) \cdot d\lambda\,, \qquad i \in \{1,2,3\} \qquad \text{Eq. 1.1}$$

Here, λ is the wavelength in the visible spectrum, $\boldsymbol{L}(\lambda)$ is the spectrum of the illuminant, $\boldsymbol{r}(\lambda)$ is the reflectance spectrum at the point of incidence, $\bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda)$ is the i^{th} colour matching function (such as a cone sensitivity function or a CIE colour matching function) and $\boldsymbol{X}_i$ is the i^{th} tristimulus coordinate corresponding to $\bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda)$. The product $\boldsymbol{L}(\lambda) \cdot \boldsymbol{r}(\lambda)$ simply represents the amount of light reflected by the surface. The colour matching functions $\bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda)$ most commonly employed in human vision science are the cone sensitivities such as those reported by Smith & Pokorny (1975) or Stockman & Sharpe (2000). In applied colorimetry, the most widely used formulations of trichromatic colour spaces are those proposed by the Commission Internationale de

l'éclairage (CIE). A standard set of colour matching functions defined by the CIE for a 2° standard observer, called the CIE 1931 2° XYZ matching functions, is shown in Figure 1.1a. An important point to note here is that since these colour matching functions are each based on sets of three primaries (or reference lights), they are related to each other through linear transformations. Thus, any set of matching functions can, theoretically, be used to derive every other set of matching functions provided this transformation matrix is known.

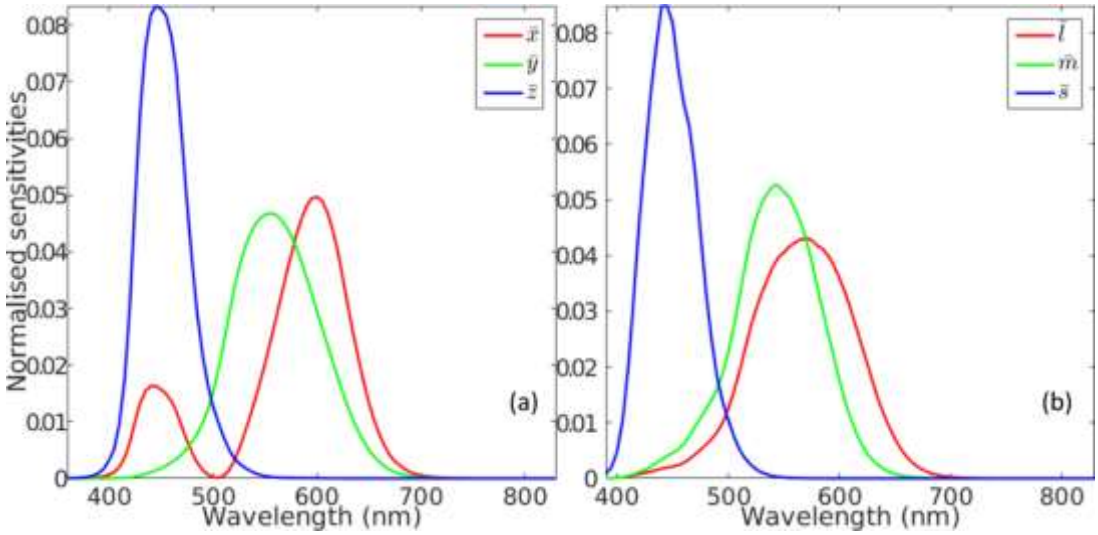


Figure 1.1: (a) CIE 1931 2° colour matching functions (b) Smith Pokorny 2° cone fundamentals, normalised. Data for both sets of curves downloaded from http://www.cvrl.org/.

While this trichromatic view of colour is sufficient to characterise a given light in terms of the photoreceptor quantum catches, it fails to explain the percept elicited by this light to any satisfactory degree. In other words, it cannot describe how an observer would experience the appearance of this light in terms of perceptual dimensions such as its apparent brightness, its hue or its saturation. A subtle but important corollary of this limitation is that trichromatic values can predict when two given lights will match (i.e., evoke equivalent photoreceptor responses); but, if the two lights do not match, trichromatic theory, by itself, is not sufficient to make any predictions about how different the two lights are. Thus, trichromatic theory does not offer a rigorous perceptual description of light beyond matching. Due to this, trichromatic theory also fails to offer explanations for more complex colour appearance phenomena such as

simultaneous contrast and the Bezold-Brücke effect. Simultaneous contrast (Blackwell & Buchsbaum, 1988) refers to a change in the appearance of a given stimulus with a change in background colour, while the Bezold-Brücke effect describes the observation that the perceived colour of a monochromatic stimulus changes as a function of its luminance (Purdy, 1931).

A more successful explanation of colours appearance was offered by the opponency mechanisms proposed by Hering (1920). Hering's colour opponency theory describes colour in terms of four basic percepts – red, green, yellow and blue, which operate in red-green and yellow-blue opponency mechanisms. According to this theory, it is impossible to experience two opponent colour percepts simultaneously, e.g., a reddish-green or a yellowish-blue colour. Although red and green lights can indeed be mixed, the resulting percept is not red-green, but rather, yellow. Based on the idea of opponency mechanisms, the colours perceived by humans were recognised to have certain quasi-invariants, such as the four unique hues and a percept of achromaticity. The percept of a unique hue is said to occur when one of the opponent mechanisms is silenced. For example, a unique red percept would imply the yellow-blue mechanism to be at a resting state. In a similar fashion, an achromatic percept is experienced when both the opponency channels are at equilibrium. In other words, an achromatic percept is neither red nor green, and at the same time, neither yellow nor blue.

Physiologically, while chromatic opponency pathways have been identified in the LGN (Lateral Geniculate Nucleus), their cardinal axes have been found to code opponency channels which do not correspond directly to these perceived unique hues (Derrington, Krauskopf, & Lennie, 1984; Tailby, Solomon, & Lennie, 2008). Plausible higher order representation of the unique hues in the glob cells of the inferior temporal cortex has been suggested by Stoughton & Conway (2008), but their results cannot be considered conclusive due to the limitations of the methodology used (Conway & Stoughton, 2009; Mollon, 2009). Thus, as it stands, no unequivocally clear neural basis of unique hues has so far been reported (Mollon & Jordan, 1997; Valberg, 2001), and they are often considered a putative third stage in the colour vision pathway (Kuehni, 2014; Smithson, 2014).

Despite offering a better explanation of certain observations, colour opponency is still far from sufficient as an accurate description of colour appearance. Phenomena such as chromatic adaptation (MacAdam, 1956) and non-linearity reported in these adaptation mechanisms (MacAdam, 1961, 1963) require more elaborate models which can reliably predict the appearance of stimuli across a number of illumination conditions. A more complete description of colour appearance is offered by modern day Colour Appearance Models (Fairchild, 2013) or CAMs. CAMs aim to achieve this by offering a colour space which can accurately describe perceptual colour order systems such as the Munsell colour system and the Natural colour system or NCS. Both the Munsell and the NCS are systems that organise colours along three perceptual dimensions which roughly correspond to an achromatic percept, and a set of four unique hues arranged in two opponent directions, red-green and yellow-blue. Thus, most modern CAMs (Hunt, 1991; Moroney et al., 2002; Nayatani, Takahama, & Sobagaki, 1981) rely heavily on quasi-invariants of human vision such as unique hues and the achromatic locus to formulate quantitative correlates of colour appearance. For instance, achromatic loci are used as a means to scale the cone fundamentals (Bompas, Powell, & Sumner, 2013; Walraven & Werner, 1991) and to define the null point of opponency mechanisms, while unique hues are used to formulate perceptually uniform partitions of the chromaticity plane (Xiao, Pointer, Cui, Chauhan, & Wuerger, 2015; Xiao, Wuerger, Fu, & Karatzas, 2011).

To illustrate the structure of a typical colour appearance model, let us briefly consider the transformations employed by the CIECAM02, one of the more widely accepted CAMs originally proposed by Moroney et al. (2002). The steps described here are a quick run-through of the forward model, adapted from a detailed description by Fairchild (2013). The main inputs to the model are the $CIE\ 1931\ XYZ$ tristimulus values for the stimulus $\boldsymbol{X} = \{x_i\}_{i=1}^{3}$ and the white point $\boldsymbol{X^W} = \{x_i^W\}_{i=1}^{3}$, the adapting luminance, and various parameters describing the nature of the surround luminance, the degree of adaptation, the strength of stimulus-surround interaction, and an exponential non-linearity factor. Starting from these inputs, the following steps are followed to arrive at correlates of colour appearance proposed by the CIECAM02.

1. The input tristimulus values are transformed to a spectrally sharpened space using a transform ($\boldsymbol{M_{CAT02}}$) called the CAT02 (Calabria & Fairchild, 2001). This transform is normalised such that an equal-energy illuminant produces equal cone responses. This transformation is also performed for the white point.

$$\boldsymbol{R} = \{r_i\}_{i=1}^{3} = \boldsymbol{M_{CAT02}}\boldsymbol{X}$$
$$\boldsymbol{R^W} = \{r_i^W\}_{i=1}^{3} = \boldsymbol{M_{CAT02}}\boldsymbol{X^W} \qquad \text{Eq. 1.2}$$

   Throughout this description, the symbol r with subscripts 1, 2 and 3 will refer to quantities related to the long, medium and short wavelength cone photoreceptors respectively.

2. Using the input parameter describing the degree of adaptation (D) and the sharpened coordinates of the white-point $\boldsymbol{R^W}$ calculated in the previous step, the sharpened responses $\boldsymbol{R}$ are then transformed to a space corresponding to adaptation to an equal-energy white illuminant. This transform uses a von Kries type model where each sharpened channel is scaled independent of the other channels. The adapted coordinates $\boldsymbol{R^a}$ are given by

$$\boldsymbol{R^a} = \{r_i^a\}_{i=1}^{3}$$
$$r_i^a = [(100D/r_i^W) + (1 - D)]r_i \qquad \text{Eq. 1.3}$$

3. A linear transform ($\boldsymbol{M_{HPE}}$) from the CAT02 primaries to the Hunter-Pointer-Estevez (HPE) primaries is then employed to convert the adapted tristimulus values $\boldsymbol{R^a}$ to coordinates $\boldsymbol{R^{HPE}}$ in the HPE space. The primaries of this space closely resemble human cone sensitivities, and lead to more accurate calculations of subsequent non-linearities.

$$\boldsymbol{R^{HPE}} = \{r_i^{HPE}\}_{i=1}^{3} = \boldsymbol{M_{HPE}}\boldsymbol{M_{CAT02}^{-1}}\boldsymbol{R^a} \qquad \text{Eq. 1.4}$$

4. A post-adaptation non-linearity based on Hunt (1991) is then applied in the HPE space to derive the compressed responses $\boldsymbol{R^c}$.

$$\boldsymbol{R^c} = \{r_i^c\}_{i=1}^3$$
$$r_i^c = \frac{400(F_L r_i^{HPE}/100)^{0.42}}{27.13 + (F_L r_i^{HPE}/100)^{0.42}} + 0.1 \qquad \text{Eq. 1.5}$$

Here, F_L is a luminance-level adaptation factor which is calculated using the adapting luminance (an input parameter).

5. The compressed responses are converted to opponency channels: the red-green channel a and the yellow-blue channel b.

$$a = r_1^c - (12r_2^c/11) + (r_3^c/11)$$
$$b = (1/9)(r_1^c + r_2^c - 2r_3^c) \qquad \text{Eq. 1.6}$$

6. From these opponency based coordinates, various perceptual correlates are calculated such as the hue angle, lightness, brightness, colourfulness, saturation and chroma.

As is evident from the above description of CIECAM02, quasi invariants of human vision like the unique-hues and the achromatic locus strongly underpin the transformations employed by modern CAMs. In addition, CAMs also use a von Kries type chromatic adaptation (CAT02 in case of the CIECAM02) to transform cone-excitations from the stimulus to a standard illumination condition (such as the equal-energy white for CIECAM02).

One of the most desirable features of any model of colour appearance is the definition of a perceptually relevant difference metric. Such a difference metric enables the interpretation of distances in the colour space proposed by the model as perceptual differences in the appearance of colours. Thus, two colours with coordinates that are equidistant from the coordinates of a given reference colour would have the same perceived difference in appearance from the reference colour. Helmholtz had first used a Riemannian line-element to define distances in colour space, and over the years there have been many proposed modifications to the line-element (Schrödinger, 1920; Stiles, 1946; Vos & Walraven, 1972). The basic idea of the line-element is to define a differential space with a distance metric based on perceptual differences between

colours. The smallest unit of distance described by this perceptual colour-difference metric is the just-noticeable-difference or JND. Given a reference stimulus in a certain colour space, a JND describes the distance one must move in this colour space in order to achieve a just-perceivable change in the colour of the stimulus. The first successful integration of a Riemannian model with measured psychophysical data was accomplished by MacAdam (1942), who used the theory in a two dimensional chromaticity plane to define unit standard-deviation steps for his data. Since then, many other studies (Melgosa, Hita, Poza, Alman, & Berns, 1997; Poirson & Wandell, 1990a; Poirson, Wandell, Varner, & Brainard, 1990) have measured and characterised colour discrimination surfaces, both in chromaticity as well as 3-D (i.e., luminance-chromaticity) colour spaces. Many colour spaces with uniform perceptual distances have been proposed by the CIE such as the $CIELAB$ and the CIE 1976 UCS spaces. The descriptions of these spaces can be found in Appendix 1.

A logical extension of colour appearance models is a model that describes appearance phenomena, not only for uniform patches, but also complex polychromatic stimuli. Although appearance models for images have been proposed in colour imaging, the most prominent being the iCAM.(Fairchild & Johnson, 2004), these models are geared towards assessment and description of image and video quality on the basis of metrics motivated by human vision. For instance, the iCAM model employs spatial convolution using low pass filters on the image as the first stage of processing, and then uses the output from these filters as the adapting stimulus in calculations very similar to CIECAM02 to arrive at correlates of image appearance. The focus of the model is to enable cross-media colour management and reproduction for colour images. Since it is very difficult to establish an equivalent of a colour-order system (such as the Munsell or NCS) for textures due to their complexity, the performance of these image-appearance models is harder to quantify in terms of human perception.

Despite employing certain elements inspired by properties of early colour vision such as low-pass spatio-temporal filtering and colour opponent processing, current image appearance models do not use higher order phenomena reported in the literature on perception of polychromatic stimuli. For example, Thomson & Foster (1997) showed

that in a discrimination task, human observers are preferentially sensitive to second and higher order statistics in natural images. In fact, natural scenes and textures have been reported to be important priors for our visual system (Geisler, 2008; Nascimento, Ferreira, & Foster, 2002; Webster & Mollon, 1994, 1997). They have been proposed as important factors in the evolution of human colour vision itself (Changizi, Zhang, & Shimojo, 2006; Regan et al., 2001), and constraints based on the reflectance spectra of natural surfaces have also been used to provide rigorous accounts of observed colour phenomena such as the asymmetry in the unique hues and frequencies of linguistic colour categories (Philipona & Regan, 2006). Thus, in order to put forward a convincing model of image appearance, it is perhaps necessary to characterise the discrimination performance of the human visual system for natural polychromatic stimuli. This becomes especially relevant if one is to define differences metrics in the appearance space proposed by the model. In particular, comparison of discrimination performances for natural textures and uniform patches could be very informative in understanding how the statistics of complex polychromatic stimuli change discrimination thresholds compared to static uniform colour patches. Montag & Berns (2000) compared luminance thresholds for textures and uniform patches and found the luminance thresholds for textures to be higher by a factor of two. Hansen, Giesel, & Gegenfurtner (2008) and Giesel, Hansen, & Gegenfurtner (2009) estimated chromatic thresholds in an isoluminant plane for uniform patches, natural objects and polychromatic textures with colour distributions similar to natural stimuli.

This thesis presents a series of experiments and analyses which can be divided into two parts. The first part examines two important quasi-invariants which underpin most modern CAMs – the achromatic locus (Chapter 2) and the unique hues (Chapter 3). The validity and robustness of these quasi-invariant percepts is investigated under varying illumination conditions. Chapter 3 also explores various models of chromatic adaptation for unique hues, and evaluates their accuracy and efficiency. The second part (Chapter 4) presents experiments aimed at extending CAMs to the appearance of a crucial natural texture – human skin. This is done by estimating discrimination thresholds for simulated colour-accurate skin patches, thereby providing JND measurements which can be used to define difference metrics in appearance space.

Chapter 2 reports measurements of achromatic settings made under three illumination conditions – dark, simulated daylight, and fluorescent lighting It introduces a novel algorithm for navigating colour space and evaluates its effect on the variability of the achromatic settings made by observers. Often, in experiments measuring achromatic settings, observers are asked to make adjustments in colour space such that each step falls along the cardinal axes of a colour or chromaticity space such as the MacLeod-Boynton space (MacLeod & Boynton, 1979) or the $CIELUV$ space (CIE, 2004). The axes of these colour spaces do not necessarily denote perceptually robust colour directions. The first chapter of this thesis proposes a method of navigation in colour space which employs movements along unique-hue directions, and reports two estimates of achromatic settings for each observer (in each condition), one made using the novel navigation method, and the other made using the traditional method of movement along cardinal colour-space axes. The results indicate that achromatic settings made using the novel navigation scheme do indeed show a reduced variability. The measurements also reveal that achromatic settings show maximum variance along the daylight locus.

Chapter 3 evaluates the performance of three models of chromatic adaptation (diagonal, linear and affine) using a large unique hues dataset (Wuerger, 2013; Xiao et al., 2011) collected in three different ambient illumination conditions. The model optimisation is done in three colour spaces – the cone excitation space, the differential cone excitation space, and the cone-contrast based DKL opponency space. The small-sample Akaike criterion AIC_c (Akaike, 1974; C. Hurvich & Tsai, 1989), which takes the trade-off between the accuracy of the model and the number of free parameters into account, is used to evaluate the models. The diagonal Von Kries model in the cone excitation space is found to perform the best in terms of AIC_c. Model performance is also calculated for three metrics – the $CIELAB$ ΔE_{LAB} colour difference (CIE, 2004), the $CIELAB$ Δh_{ab} hue-angle difference (CIE, 2004), and the angle $\Delta\theta_{UH}$ between unique hue planes (Wuerger, Atkinson, & Cropper, 2005) fitted to the measured and predicted settings. The model with the highest number of parameters, i.e. the Affine model, is found to perform the best on all the metrics. The diagonal and linear models show similar performances across colour spaces, reiterating their theoretical equivalence through decorrelation and sharpening transforms. In addition, these

models also show deviations from unimodality in the prediction errors for luminance dependent metrics ΔE_{LAB} and $\Delta \theta_{UH}$ across observers.

Although not a part of the corpus of this thesis, work was also done (Xiao et al., 2015) on evaluating whether CIECAM02 is indeed capable of matching NCS unique-hue data. The results of the study indicate significant differences between the CIECAM02 defined unique hue lines and unique-hue lines corresponding to actual NCS data.

Chapter 4, the main chapter in this thesis, presents three experiments which investigate the discrimination performance of human observers on simulated skin images and polychromatic images derived from simulated skin patches. These measurements are performed in three ambient illumination conditions – dark, simulated daylight and cool white fluorescent light. For each polychromatic stimulus, thresholds for the corresponding uniform patch of the same mean $CIELAB$ colour are also measured. The first experiment compares the thresholds for simulated skin and corresponding mean uniform patches. The thresholds for the polychromatic stimuli are consistently higher, for both the chromatic, and the luminance projections. The second and third experiments investigate the effect of ambient illumination and the relative location of the stimuli in colour space. The observed trends match those reported by Hansen et al. (2008) very closely, with the area of the chromaticity ellipses showing a gradual increase with distance from the illumination chromaticity. The orientations of the chromatic ellipses for simulated skin are found to align with the vector joining the mean patch chromaticity with the illuminant chromaticity.

Chapter 2

The achromatic locus: effect of navigation direction in colour space

Abstract

The achromatic locus is defined as a light that is devoid of any hue. This is usually achieved by asking observers to adjust the stimulus such that it looks neither red nor green and at the same time neither yellow nor blue. Despite the theoretical and practical importance of the achromatic locus, little is known about the variability in these settings.

The main purpose of the current study was to evaluate whether achromatic settings were dependant on the task of the observers, namely the navigation direction in colour space. Observers could adjust the test patch either along the two chromatic axes in the CIE u^*v^* diagram, or alternatively, navigate along the unique hue lines.

The main result of the study is that the navigation method affects the reliability of these achromatic settings. Observers are able to make more reliable achromatic settings when adjusting the test patch along the directions defined by the four unique hues as opposed to navigating along the main axes in the commonly used CIE u^*v^* chromaticity plane. This result holds across different ambient viewing conditions (DARK; D65; CWF) and different test luminance levels ($5,\ 20,\ 50\ cd/m^2$). The reduced variability in the achromatic settings is consistent with the idea that internal colour representations are more aligned with the unique hue lines than the u^* and v^* axes.

NOTE: This chapter is a modified version of the original article published in the Journal of Vision (Chauhan et al., 2014).

2.1 Introduction

An achromatic stimulus is defined as patch of light that is devoid of any hue. This is usually achieved by asking observers to adjust the stimulus such that it looks neither red nor green and at the same time, neither yellow nor blue. In terms of opponent-colour theory, both chromatic opponent mechanisms, the red-green and the yellow-blue mechanisms, are at equilibrium if a colour-normal observer views such an achromatic stimulus. The output of the chromatic channels is hypothesised to be at zero since no hue is perceived in such a stimulus. The idea that the chromatic system is at a 'resting state' at the achromatic locus has led to the use of the achromatic point as a mean to scale the cone fundamentals (Bompas et al., 2013; Walraven & Werner, 1991). Experimentally, settings of the achromatic loci are widely used to establish the effect of illumination or to assess spatial or temporal context effects (e.g. Doerschner, Boyaci, & Maloney, 2004; Lee, Dawson, & Smithson, 2012). The extent of invariance of these achromatic settings depends on the spatial configuration, the chromaticity of the background and the ambient illumination (Bäuml, 2002; Brainard, 1998; Delahunt & Brainard, 2004; Helson & Michels, 1948), on eye fixation (Granzier, Toscani, & Gegenfurtner, 2012), on viewing pattern (Golz, 2010) and on the precise task instructions (Ekroll, Faul, Niederee, & Richter, 2002). The purpose of the current study was to evaluate whether achromatic settings were dependent on one particular task aspect, namely the navigation direction in colour space. It focuses on the reliability of the achromatic settings across and within observers; virtually all experiments using achromatic settings make the implicit assumption that observers are able to consistently navigate in a two-dimensional (or even three-dimensional) colour space. Here this assumption was tested directly by manipulating the directions in colour space along which observers can adjust the stimuli to obtain an achromatic setting. In the first method, observers were able to adjust the stimuli along the two main axes in an approximately uniform CIELUV colour space (Wyszecki & Stiles, 2000). Here, this navigation method is referred to as the u^*v^* method. In the second method, referred to as UH method, stimuli could be adjusted along the unique hue lines (Hering, 1920). A secondary aim was to evaluate whether these achromatic settings are invariant under changes in ambient illumination. The main result of the study is that the reliability in the achromatic settings is higher when observers are able to adjust along the unique

hue directions; this result generalises over all three ambient viewing conditions and test luminance levels.

2.2 Methodology

2.2.1 Apparatus

Stimuli were generated with the CRS MATLAB toolbox on a 14-bit ViSaGe system (Cambridge Research System, Kent, UK) and displayed on the CRT monitor (Mitsubishi DiamondPro 2070). Observer responses were collected using a CB6 response box (Cambridge Research System, Kent, UK). The monitor was calibrated using the ColourCal calibration device (Cambridge Research System, Kent, UK). The CIE chromaticity coordinates and luminance of the phosphors at peak output were as follows: red = $(0.629,\ 0.3421,\ 25.4\ cd/m^2)$; green = $(0.290,\ 0.605,\ 65.5\ cd/m^2)$; blue = $(0.148,\ 0.070,\ 11.2\ cd/m^2)$. The monitor was switched on at least 30 minutes before the start of the experiment to ensure a stable luminance output.

A GTI ColorMatcher GLE M5/25 installed on the ceiling in the centre of a booth was used to provide two lighting conditions (Xiao, Fu, Mylonas, Karatzas, & Wuerger, 2013): a D65 simulator for daylight, and CWF for typical white fluorescent office light; in addition, a DARK condition was also included where the only source of illumination was the test patch on the CRT screen. The inside of the booth was painted dark grey, reflecting very little light. Table 2.1 reports the luminance and CIE xy chromaticity coordinates of the illuminants as given by a white tile placed underneath the GTI ColorMatcher and measured by a PhotoResearch PR-650 spectroradiometer.

The peak output of the monitor was measured under all three ambient viewing conditions (DARK, D65 and CWF). Due to the small amount of ambient light reflected from the CRT display, the transformation matrices from RGB to XYZ vary slightly under the different viewing conditions. These small differences have been taken into account in the computation of the $u'v'$ co-ordinates.

	CIE xy coordinates	Luminance	Correlated Colour Temperature
D65	$(0.312, 0.334)$	$41\ cd/m^2$	$6376\ K$
CWF	$(0.394, 0.387)$	$136\ cd/m^2$	$3747\ K$

Table 2.1: Illuminant colour coordinates and Correlated Colour Temperatures

2.2.2 Observers

Thirty subjects (18 females and 12 males; mean age: 24.24 years; age range: 18–60 years) participated in the experiment. All participants had their colour vision assessed using the Cambridge Colour Test (Regan, Reffin, & Mollon, 1994) and were found to be colour normal.

2.2.3 Experimental Procedure

The effect of three variables on the achromatic locus were investigated: the navigation direction in colour space (u^*v^* method or UH method), which was the main manipulation; the luminance level of the stimuli ($5,\ 20,$ and $50\ cd/m^2$), and the ambient illumination condition (DARK, D65, CWF). Each participant repeated each achromatic setting three times; in the course of the complete experiment, each of the 30 observers therefore made 54 achromatic settings (3 illumination conditions x 3 luminance levels x 2 navigation methods x 3 repetitions).

Achromatic settings under the different ambient illumination conditions were obtained in separate blocks which were run in the order: DARK, D65 and CWF. The background of the monitor was set to black, throughout all experiments. The observer adapted for at least 5 minutes to the prevailing illumination to ensure steady adaptation. Within each block, the order of the navigation direction (which is the main variable in the study) was balanced: half the observers first conducted the u^*v^* task followed by the UH task; the other half ran the experiments in the reverse order. This ensured that

learning effects would not differentially bias the achromatic settings in one of the two tasks.

Each block (i.e. fixed ambient illumination condition) lasted about 15-20 minutes; hence the entire set of experiments lasted between 45 minutes and one hour for each observer, including pre-experiment adaptation and a short break between the different illumination conditions. Observers were compensated for their time with a small fee.

Each trial started with an initial colour which was chosen randomly from a pre-defined radius of eight colours at distances of ~25 u^*v^* units around the point $(0, -5)$ in the u^*v^* plane. The choice of this central point is to some extent arbitrary, and was chosen at ($x = 0.289$, $y = 0.298$) since this was the background chromaticity used in previous unique hue experiments (Wuerger, 2013). The CIE xy coordinates of the eight starting values were as follows: (0.3908, 0.3006); (0.4079, 0.3587); (0.3947, 0.4148); (0.3458, 0.4433); (0.2765, 0.4273); (0.1869, 0.3159); (0.2018, 0.2244); (0.2936, 0.2110). On each trial the observer was asked to adjust the colour of a central circular ~2.6° patch such that it contained neither red nor green and neither yellow nor blue. The participant could make these adjustments along four directions in colour space by pressing one of four buttons on the response box. These buttons roughly controlled movement along the red-green and yellow-blue directions (see details below), i.e., pressing the red button reduced the level of red in the test patch, pressing the green button reduced the level of green and so on. There was no time limit and observers could switch back and forth between the two chromatic axes. When observer was satisfied with his/her choice s/he pressed a fifth button and the next trial started.

To statistically evaluate the effects of all three variables, a multivariate analysis of variance (MANOVA; SPSS Version 20) was performed with the dependent variable being the two-dimensional co-ordinates in the $u'v'$ chromaticity diagram. A MANOVA is most efficient for moderately correlated dependent variables (Stevens, 2009); in the data reported below the correlations varied from 0.2 to 0.8, which is in line with the range of recommended correlations (0.3-07). Subsequently post-hoc comparisons were performed with correction for multiple comparisons.

2.2.4 Navigation Directions in Colour space.

The main manipulation was the navigation direction in colour space: to find an achromatic setting, the observers could adjust the stimuli either along the two axes (u^* and v^*) of the standard CIE LUV colour space (denoted as u^*v^* task; see Fig 1a) or along the unique-hue directions (denoted as UH task; see Fig 1c). The unique hue lines used were based on the unique hue settings of 185 colour-normal observers (Wuerger, 2013) and obtained using a hue selection task (Wuerger et al., 2005). In this large sample, the inter-observer variability exceeded the intra-observer variability by a factor of 2 (details on observer variability are in Table II, Xiao, Wuerger, Fu, & Karatzas, 2011). Inter-observer perceptual errors (expressed in ΔE_{00}) for the four unique hues range from 1.17 to 2.3. The inter-observer variability was therefore considered sufficiently small to use the same directions for all observers. If anything, unique hue directions not optimised for each individual observer should result in less pronounced task differences, and thereby underestimating the effect of navigation direction.

Both tasks in the current experiment were cancellation tasks, i.e., the colour directions chosen by the observers were used to cancel colour from the stimulus to obtain the subsequent stimulus. Figures 2.1b and 2.1d describe how the stimulus was updated after each observer's response. For both navigation directions, the step size in u^*v^* space was fixed to 5 units, which corresponds roughly to a distance of 0.007 in the $u'v'$ diagram. This step-size has been derived from extensive preliminary studies with the authors as experienced observers; it is well below the discrimination threshold around a typical CRT white point obtained for colour-normal observers (Regan et al., 1994). Within each trial, the luminance level of the test patch was fixed and observers could only adjust the chromaticity of the patch.

For the u^*v^* method (Figure 2.1b) the starting value for the chromaticity of the test patch might be at (u_o, v_o), for example. The next colour stimulus (u_n, v_n) was calculated by moving 5 units from the initial point (u_o, v_o) along one of the four directions determined by the axes of the u^*v^* plane. The direction was chosen by the

participants by pressing one of four buttons. In response to the button press, the test patch was updated accordingly (as shown in Figure 2.1b).

In the UH method (Figure 2.1d), the step-size was also 5 u^*v^* units, but now the direction of change was calculated as follows. First the directional unit vector $\hat{\boldsymbol{e}}$ from the chosen unique hue $\boldsymbol{x_{uh}} = [u_{uh}\ v_{uh}]^T$ to the initial colour $\boldsymbol{x_0} = [u_0\ v_0]^T$ was computed as:

$$\hat{\boldsymbol{e}} = -\frac{\boldsymbol{x_{uh}} - \boldsymbol{x_0}}{\|\boldsymbol{x_{uh}} - \boldsymbol{x_0}\|} = -\frac{1}{\sqrt{(u_{uh} - u_o)^2 + (v_{uh} - v_o)^2}}\begin{bmatrix} u_{uh} - u_o \\ v_{uh} - v_o \end{bmatrix} \qquad \text{Eq. 2.1}$$

The new stimulus at iteration n, $\boldsymbol{x_n} = [u_n\ v_n]^T$, was then defined as the point lying at a distance of $s = 5$ units from the initial colour in direction $\hat{\boldsymbol{e}}$. Here s denotes the step size in the u^*v^* space.

$$\boldsymbol{x_n} = \begin{bmatrix} u_n \\ v_n \end{bmatrix} = \boldsymbol{x_0} + s \cdot \hat{\boldsymbol{e}} \qquad \text{Eq. 2.2}$$

An example is shown in Figure 2.1d, where the initial point is assumed to be at (u_o, v_o) and the participant cancels unique yellow, i.e. yellow is reduced in the stimulus. Since the L^* values always remained constant (at L_o) in this process, the new colour stimulus is at (L_o, u_n, v_n). This new colour is then the starting point for the next iteration. This procedure was repeated until the participant confirmed that according to him/her the stimulus displayed on the screen was achromatic, i.e. contained neither red nor green and neither yellow nor blue.

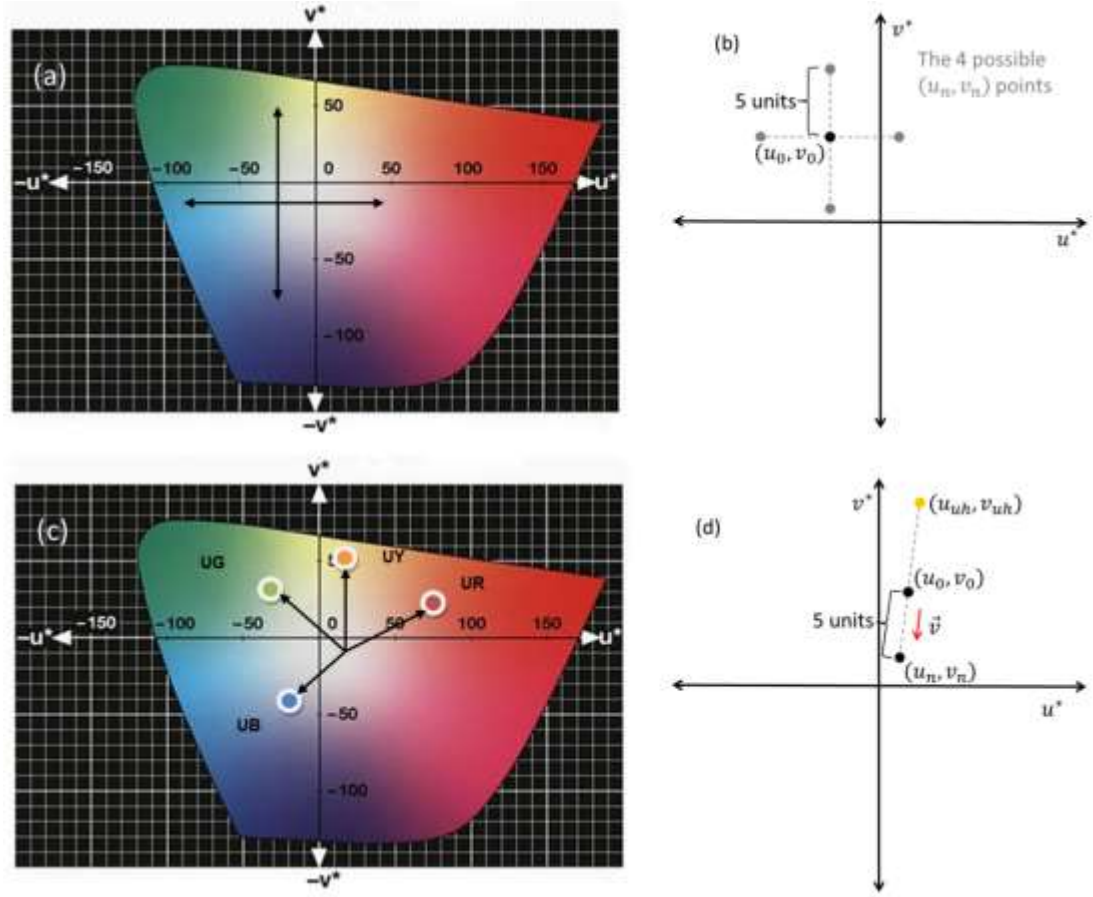


Figure 2.1: Navigation direction in colour space: (a) u^*v^* method: observers navigate along the axes in the u^*v^* chromaticity diagram. (b) The step size is 5 units in u^*v^* space. After each response the colour is updated along the relevant direction and the updated setting is used as the new starting value for the next iteration. (c) UH method: observers navigate along the directions of the unique hues. (d) The step size for the UH task is again 5 u^*v^* units but now the adjustment is made in one of the directions defined by the unique-hues. The updated setting is used as the new starting value for the next iteration.

During the experiment, the display RGB values of the final achromatic settings selected by the participants were automatically saved. Subsequently, they were transformed to CIE XYZ tristimulus values using the calibrated display profile that was generated based on colour measurements results by a spectroradiometer. To confirm the accuracy of the calibration, the CIE tristimulus values derived from the display profile were compared with the actual measurements; the differences were found to lie within measurement error.

2.3 Results

All achromatic settings will be presented in the $u'v'$ chromaticity diagram rather than in the CIE xy chromaticity diagram, since the former is known to be approximately uniform, that is, equal Euclidean distances in $u'v'$ reflect approximately similar perceptual distances. Approximate uniformity is of relevance for the interpretation of these results since comparisons of the intra- and inter-observer variability across different navigation methods and illumination conditions (Wuerger, Maloney, & Krauskopf, 1995) will be made. In what follows, the mean achromatic settings are reported first, followed by the effect of the navigation method on the reliability of the achromatic settings.

2.3.1 Mean achromatic settings: Effect of navigation method, ambient illumination and test luminance level

Figure 2.2 shows the achromatic settings averaged over all observers; error bars denote ± standard error of the mean. This section will, in turn, discuss the effect of the navigation method, the effect of illumination and test luminance level on the achromatic loci. As depicted in Figure 2.2, the achromatic settings depend on the navigation method; achromatic loci obtained with the UH method (▲) differ from the settings obtained with the u^*v^* method (■), $F(2, 521) = 23.1,\ p < 0.0001$.

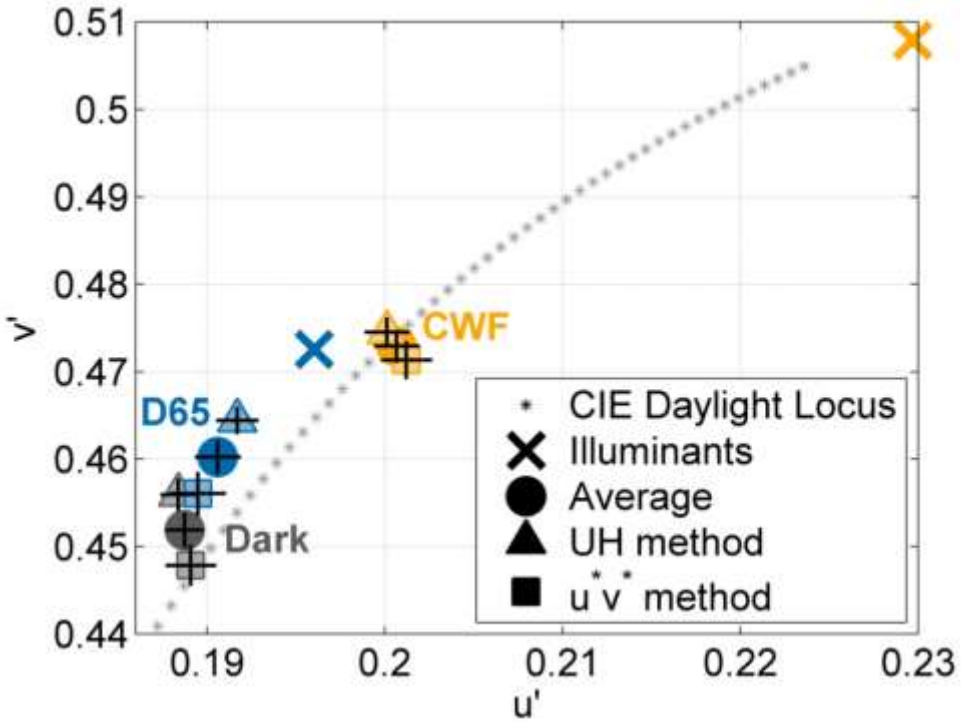


Figure 2.2: Achromatic loci averaged over all 30 observers and all test luminance levels. Error bars denote ±1 standard error of the mean. Grey symbols denote the mean settings under the dark viewing condition; blue symbols under D65 and yellow symbols under CWF. The settings for the UH method and the u^*v^* method are indicated by ▲ and ■ respectively, while ● denotes the mean setting under each illuminant.

Secondly, ambient illumination affects the achromatic locus by shifting the settings towards the prevailing illumination, $F(4, 1042) = 65.7;\ p < 0.0001$. The filled circles indicate the grand mean settings (averaged over all observers, all test luminance levels and both navigation methods) for the dark viewing condition (●), under D65 (●), and under CWF (●). When compared to achromatic settings obtained under dark viewing conditions (grey symbols), the settings are shifted towards the chromaticity of the prevailing illumination (bluish symbols for D65; yellowish symbols for CWF). When viewed in the D65 condition (×) achromatic settings (■ and ▲) move along the daylight locus towards D65; similarly, when viewed in the CWF condition (×), the achromatic settings (■ and ▲) move towards the chromaticity of the ambient illumination. Post-hoc comparisons revealed that all pairwise differences in the achromatic settings (DARK vs. D65; D65 vs. CWF; CWF vs. DARK) were statistically significant (ps corrected < 0.05).

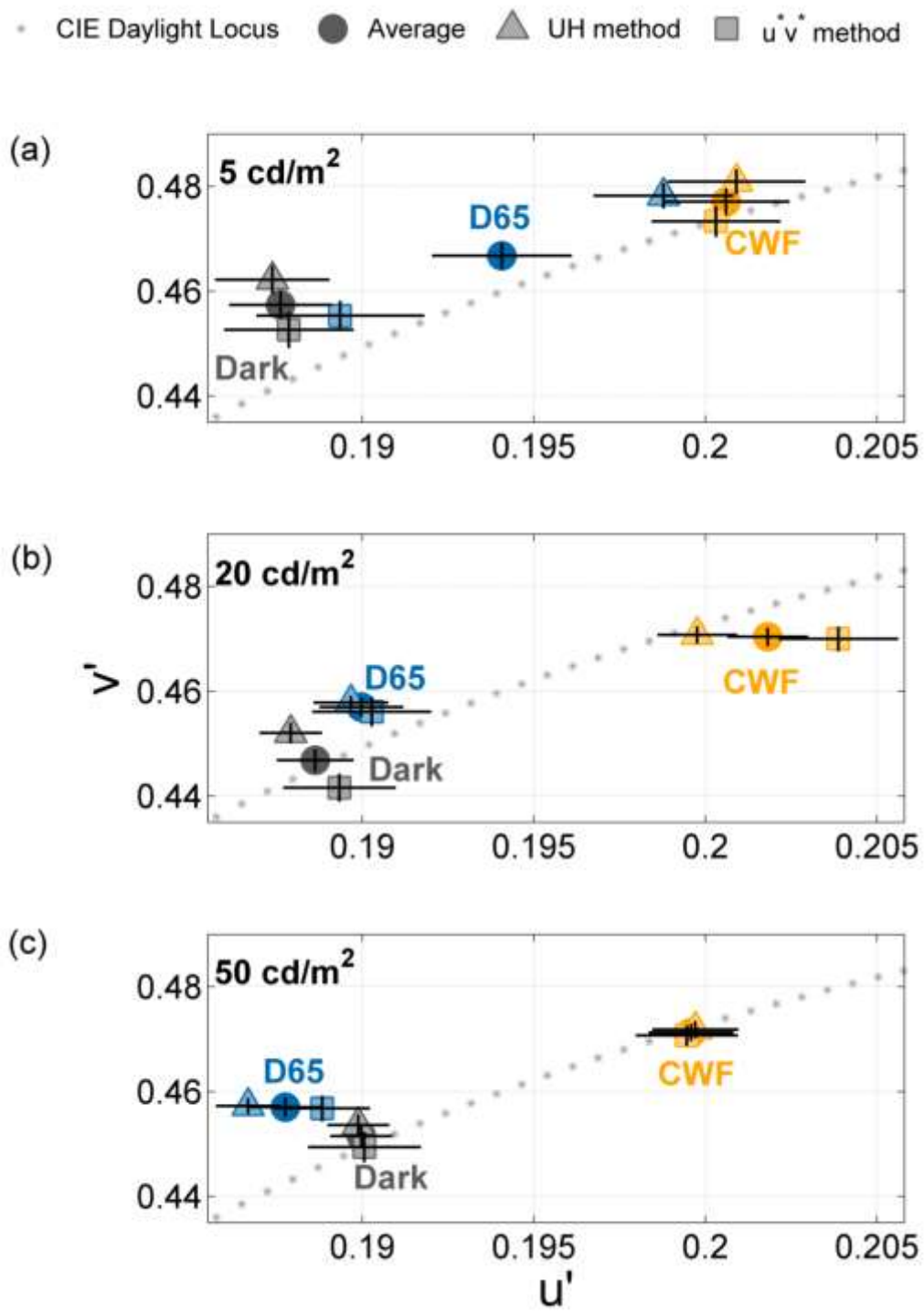


Figure 2.3: Achromatic settings under stimulus luminance level of (a) 5 cd/m^2, (b) 20 cd/m^2 and (c) 50 cd/m^2. The settings are averaged over the 30 observers and the same symbols as Figure 2.2 are used. At the highest luminance level both navigation methods converge on very similar achromatic settings.

Thirdly, the luminance level of the test patch affects the achromatic settings, $F(4,2042) = 14.6;\ p < 0.0001$. Figure 2.3 shows the achromatic settings for all

three luminance levels: (a) $5\ cd/m^2$, (b) $20\ cd/m^2$ and (c) $50\ cd/m^2$. Higher luminance level yield less variable results (cf Figure 2.3c with Figure 2.3a), both in terms of observer consistency as well as the discrepancy between the two navigation methods. At the highest luminance level (Figure 2.3c; $50\ cd/m^2$) both navigation methods converge on very similar achromatic settings. The effect of test luminance level on the achromatic settings is driven by the differences in v' under low luminance ($5\ cd/m^2$ vs. $20\ cd/m^2$: $p < 0.001$; $5\ cd/m^2$ vs. $50\ cd/m^2$: $p < 0.0001$). The larger variability for the achromatic settings at the $5\ cd/m^2$ luminance levels is likely to be related to the rod-cone interactions at mesopic light ranges (Stockman & Sharpe, 2006; Yebra, García, Nieves, & Romero, 2001). Since these interactions are non-linear and depend on a number of factors such as the distribution of the rod and cone receptors in the observer retina (Stockman & Sharpe, 2006), individual differences in receptor distribution are likely to play a role in the achromatic settings at $5\ cd/m^2$.

It must be noted that no effects of learning were found in the results. To evaluate this possibility, conducted further MANOVAs with session number as a factor were conducted. Each neutral grey setting was obtained three times; if learning took place one would expect to find a difference between these three settings. To test for such a difference, the data were split in three different groups for each of the 18 conditions (3 illuminants x 3 test luminance levels x 2 tasks) before running the MANOVAs. The results showed no statistically significant dependence on the session number: the 18 MANOVA p-values varying between 0.15 and 0.88, none of them significant.

2.3.2 Effect of navigation method on the variability of the achromatic settings

Figure 2.4 shows in more detail how the navigation method affects the variability of the achromatic settings. Each data point reflects the mean setting for an individual observer, for a particular test luminance level (Figure 2.4a: $5\ cd/m^2$; b: $20\ cd/m^2$; c: $50\ cd/m^2$) and for both navigation methods (on the left: UH method; on the right: u^*v^* method). Details of the ambient viewing conditions are as before (cf Figure 2.1).

The spread in the data points is larger for the u^*v^* method compared to the UH method (compare left and right panels in Figure 2.4); the spread in the settings is also reduced when the test luminance level is increased (compare Figure 2.4a with Figure 2.4c).

To quantify the effect of navigation method and test luminance level on the variability in the settings, both the intra-observer and inter-observer variability in the approximately uniform $u'v'$ chromaticity diagram (Xiao et al., 2013, 2011) were computed. Inter-observer variability indicates the extent to which individual observers agree with the average observer, whereas intra-observer variability indicates how consistent the individual observer is across several repetitions. Intra-observer variability is defined as:

$$OV_{INTRA} = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(u_i' - \overline{u'})^2 + (v_i' - \overline{v'})^2} \qquad \text{Eq. 2.3}$$

where n denotes the number of observations for a particular stimulus, (u_i', v_i') is the i^{th} observation, and $(\overline{u'}, \overline{v'})$ is the mean of these n observations. Inter-observer variability is calculated similarly:

$$OV_{INTER} = \frac{1}{m}\sum_{i=1}^{m}\sqrt{(\overline{u_i'} - \overline{\overline{u'}})^2 + (\overline{v_i'} - \overline{\overline{v'}})^2} \qquad \text{Eq. 2.4}$$

where m denotes the number of participants, $(\overline{u_i'}, \overline{v_i'})$ is the average setting for the i-th participant, and $(\overline{\overline{u'}}, \overline{\overline{v'}})$ is the mean of the average settings for all m observers.

The ratio between inter- and intra-observer variability is an indicator of how consistent the settings are across the sample in relation to the consistency within each observer (Kuehni, 2005). There is a large effect of navigation method on the variability (Mean variabilities: Table 2.2; Range of variabilities: Table 2.3): both inter- and intra-observer variability are much smaller when settings are obtained using the UH method, in comparison to the u^*v^* method; variability in the UH method is, on an average, just

above 70% of the variability in the u^*v^* method (73% for the intra- and 74% for the inter-observer variability). This reduction in variability in the UH method is seen for all test luminance levels, but is more pronounced under the dark viewing condition. The ratio between inter- and intra-observer variability is on average 1.74, and does not depend on the navigation method. This ratio ranges from about 1.5 to 2, indicating a good consistency in the achromatic settings across participants, and is comparable to the ratios found for unique hue settings (Xiao et al., 2013, 2011).

To quantify the reduction in variability the ratio between the variability in the UH method and the u^*v^* method for all conditions (Figure 2.5) were plotted. This ratio is always smaller than unity, for both intra- (light bars) and inter-observer variability (dark bars) indicating that the settings are less variable when observers adjust the achromatic locus along the UH directions. This result is not contingent on the choice of $u'v'$ space; almost identical variability ratios are obtained when using LAB space (Figure 2.6).

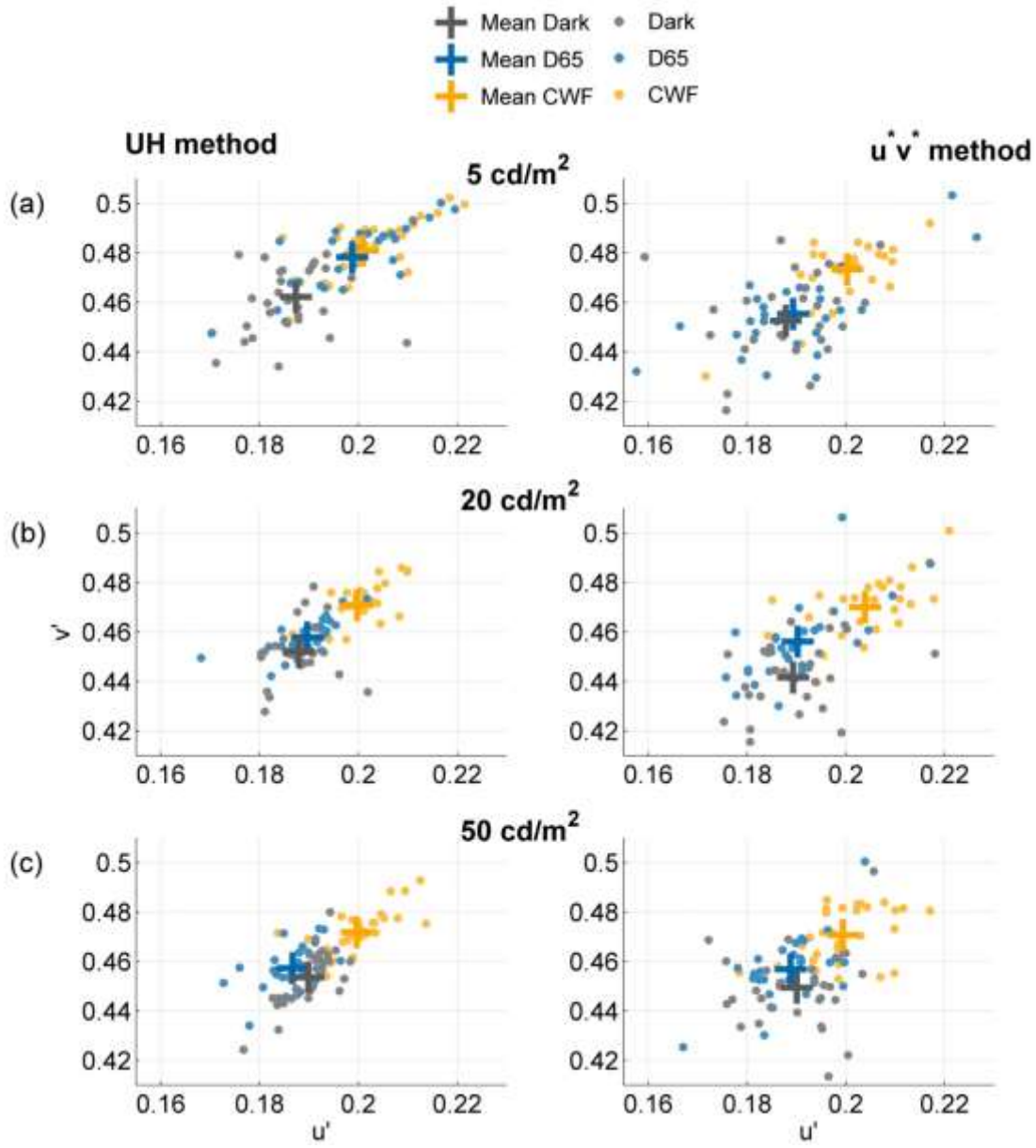


Figure 2.4: Mean achromatic settings for each observer under a stimulus luminance level of (a) 5 cd/m^2, (b) 20 cd/m^2, (c) 50 cd/m^2. Grey points denote the dark condition; blue points denote D65, and the yellow points denote CWF. The observed spread in the settings is larger for the u^*v^* method (right panel) compared to the UH method (left panel).

	DARK			D65			CWF			Avg. over Illuminants		
	Intra	Inter	Ratio	Intra	Inter	Ratio	Intra	Inter	Ratio	Intra	Inter	Ratio
UH												
5 cd/m^2	9.6	14.9	1.55	6	14.7	2.45	6.2	13.7	2.21	7.3	14.4	1.99
20 cd/m^2	6.4	9.1	1.42	4.7	7.7	1.64	5.2	8.8	1.69	5.4	8.5	1.57
50 cd/m^2	5.3	10.2	1.92	5.6	8	1.43	5.4	9.3	1.72	5.4	9.2	1.69
Lum. avg.	7.1	11.4	1.61	5.4	10.1	1.87	5.6	10.6	1.89	6	10.6	1.77
u^*v^*												
5 cd/m^2	11.5	18.3	1.59	8.2	15.9	1.94	7.2	13.7	1.90	9	16	1.78
20 cd/m^2	10.2	14.7	1.44	6.7	14.1	2.10	7	13.3	1.90	8	14	1.76
50 cd/m^2	8.8	15	1.70	7.2	11.8	1.64	7.7	12.2	1.58	7.9	13	1.65
Lum. avg.	10.2	16	1.57	7.4	13.9	1.89	7.3	13.1	1.79	8.2	14.2	1.73
Avg.	8.6	13.5	1.58	6.3	11.9	1.87	6.4	11.8	1.84	7.1	12.4	1.74

Table 2.2: Mean intra- and inter-observer variability (expressed in e-3 units) for the achromatic settings for all three illumination conditions (DARK, D65, CWF), test luminance levels (5, 20, 50 cd/m2) and both navigation methods (UH method, u^*v^* method). The ratio between inter- and intra-observer variability is independent of the navigation method.

	UH		u^*v^*	
	Intra	Inter	Intra	Inter
DARK				
5 cd/m^2	(0.58,69.30)	(2.22,39.56)	(0.57,93.36)	(3.63,45.97)
20 cd/m^2	(0.67,19.81)	(0.53,26.37)	(0.95,74.01)	(0.65,34.57)
50 cd/m^2	(0.11,17.25)	(3.25,32.94)	(1.30,65.26)	(3.70,49.74)
D65				
5 cd/m^2	(0.75,19.65)	(2.43,42.75)	(0.53,52.56)	(1.61,58.02)
20 cd/m^2	(0.55,18.17)	(0.38,23.92)	(0.00,22.78)	(2.76,50.99)
50 cd/m^2	(0.13,19.54)	(0.21,25.04)	(0.24,21.56)	(0.38,46.28)
CWF				
5 cd/m^2	(0.78,20.15)	(0.89,45.54)	(0.5421.43)	(1.60,57.81)
20 cd/m^2	(0.23,20.23)	(0.64,21.81)	(0.34,40.49)	(3.37,35.08)
50 cd/m^2	(0.47,20.51)	(0.61,24.46)	(0.83,35.58)	(2.21,26.46)

Table 2.3: Minimum and maximum values of Intra- and inter-observer variability (expressed in e-3 units) for the achromatic settings for all three illumination conditions (DARK, D65, CWF), test luminance levels (5, 20, 50 cd/m2) and both navigation methods (UH method, u^*v^* method). Each entry is in the form (Min,Max) such that Min is the minimum and Max is the maximum value of the intra- or inter- observer variability in the corresponding condition.

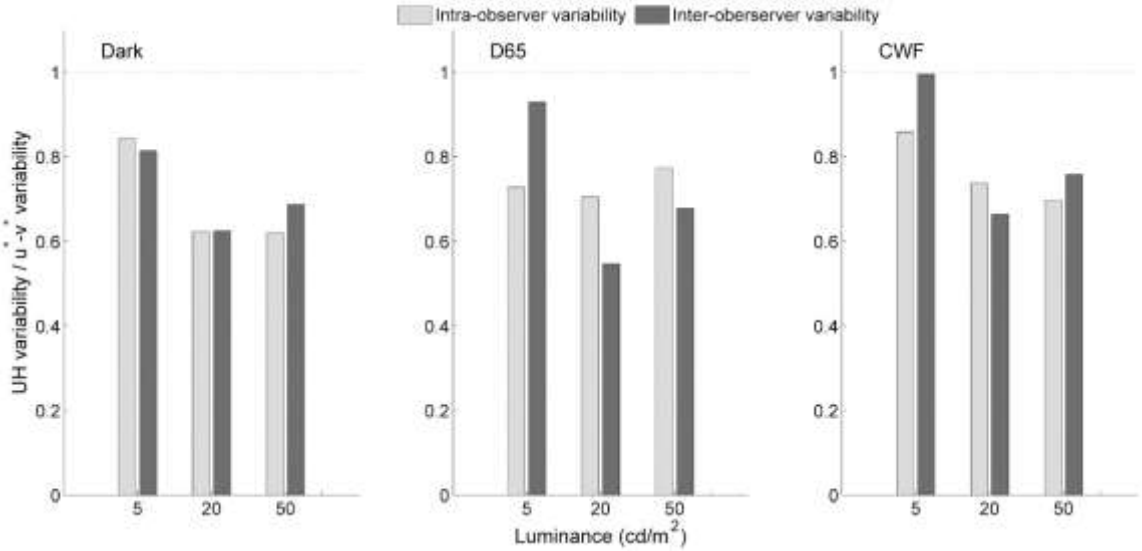


Figure 2.5: The main finding of the study is that the variability in the UH task is always smaller than in the u^*v^* task. The ratio between the variability in the UH method and the u^*v^* method, plotted against the stimulus luminance level (5, 20 and 50 cd/m2). The intra-observer variability ratios are in light colour, while the inter-observer variability ratios are in dark. Each subplot denotes a separate ambient viewing condition (DARK, D65 and CWF). The calculations for these plots are performed in the $u'v'$ space, but the same effects are found in other colour spaces too (See Figure 2.6).

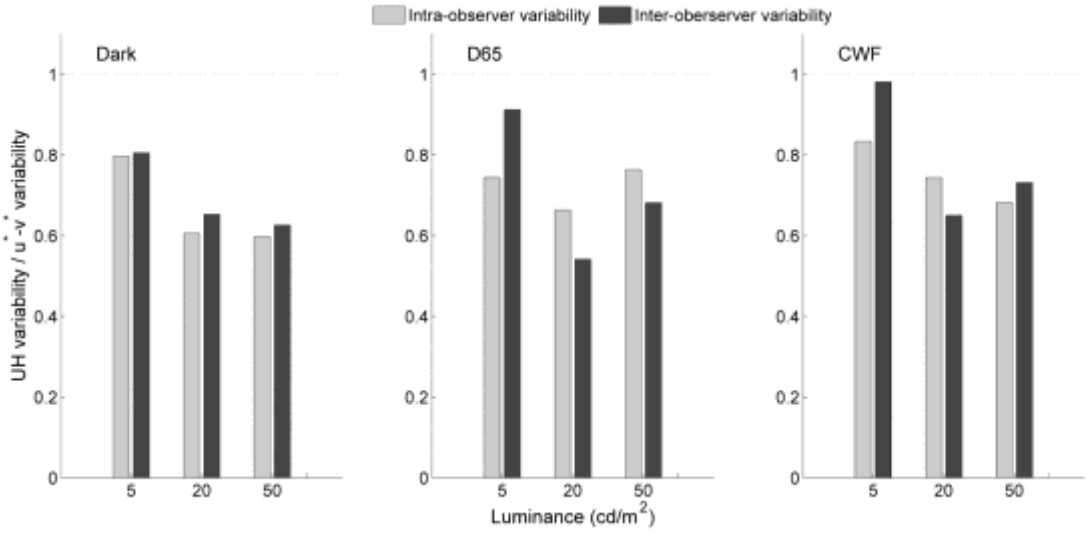


Figure 2.6: $CIELAB$ space calculation of the ratio between the variability in the UH method and the u^*v^* method. The exact same axes and colours as Figure 2.5 are used.

2.4 Discussion

The purpose of this study was to evaluate the reliability with which colour-normal human observer can perform achromatic settings, that is, to adjust a patch of light such that it appears void of any hue. Two different navigation methods in chromaticity planes of constant luminance were evaluated. The main finding is that observers' achromatic settings are more reliable (in terms of inter- and intra-observer variability) when they are asked to adjust the light along the unique hue lines compared to adjustment along the main axes of the commonly used CIE u^*v^* chromaticity diagram (Figure 2.5).

This effect of navigation method on the variability in the achromatic settings holds across all test luminance levels and ambient illumination conditions. While the stimuli used in the experiment are of limited complexity, there is no reason to believe that this effect of navigation method on the reliability of the achromatic setting depends on the particular spatial structure or viewing condition. Since the UH method relies on unique hues which are approximate pseudo-invariants of human colour vision, it is more likely that the effect of navigation method is related to the internal representation of the different directions in colour space.

A unique hue percept theoretically corresponds to an equilibrium state in one of the colour opponent channels (either red-green or yellow-blue), while an achromatic percept corresponds to an equilibrium in both the opponent channels. In the UH method, the navigation scheme constantly aims to balance one of the opponent channels by restricting movement in colour space to be always towards one of the four unique hues. Thus, the reduced variability of achromatic settings obtained by the UH navigation scheme could be a result of this iterated asymptotic balancing of colour opponency channels.

2.4.1 Shifts in the achromatic locus under changes in ambient illumination

In comparison to the achromatic settings obtained under the dark viewing condition, the achromatic loci move towards the chromaticity of the ambient illumination (Figure 2.7: ✕ indicates a light source). The observed shift (from D65 to CWF) is indicated by a ●; the predicted shift by ▲. A simple cone adaptation model fails to predict the achromatic shifts on several accounts: Firstly, while the direction of the shift is accounted for by cone adaptation without assuming any cross-talk between the different cone classes (Wuerger, 1996), the observed magnitude of the shift is much smaller, only about 30% of the predicted shift. This is consistent with previous studies showing poor colour constancy – on a dark background (Hansen, Walter, & Gegenfurtner, 2007), for increments relative to the background instead of decrements (Helson & Michels, 1948) and when the immediate surround is not at the chromaticity of the Illuminant (Delahunt & Brainard, 2004; Kraft & Brainard, 1999).

Secondly, consistent with Kuriki's study (Kuriki, 2006), the achromatic settings vary with stimulus intensity (Figure 2.3) which precludes modelling the effect of ambient illumination with cone adaptation where the scaling factors depend only on the cone absorptions of the illuminations. To compare the test luminance dependency of the achromatic settings with Kuriki's, the same analysis was performed: the relative cone weights (M/L and M/S) were plotted against the luminance in a log-log plot (Figure 2.8). It was found that the log-relative cone ratios depend both on the luminance of the test patch and on the ambient illumination. Under D65 viewing, the achromatic settings converge for high luminance values to the settings obtained under the dark viewing condition, which is in agreement with Kuriki's findings (Fig. 4; Kuriki, 2006). This suggests that the achromatic settings become independent of the illumination when the luminance level increases. However, no convergence under CWF (comparable to Kuriki's orange illumination) is found; when observers are adapted to CWF, the test luminance has little effect on the cone ratios but the cone ratios depend on the ambient illumination. There are important differences between these two studies: the area surrounding the test patch in Kuriki's study was about 10 cd/m^2 whereas in this study

the background was black; in this experiment no attempt was made to induce an 'object mode' and the observers were aware that the test patch was a self-luminous source.

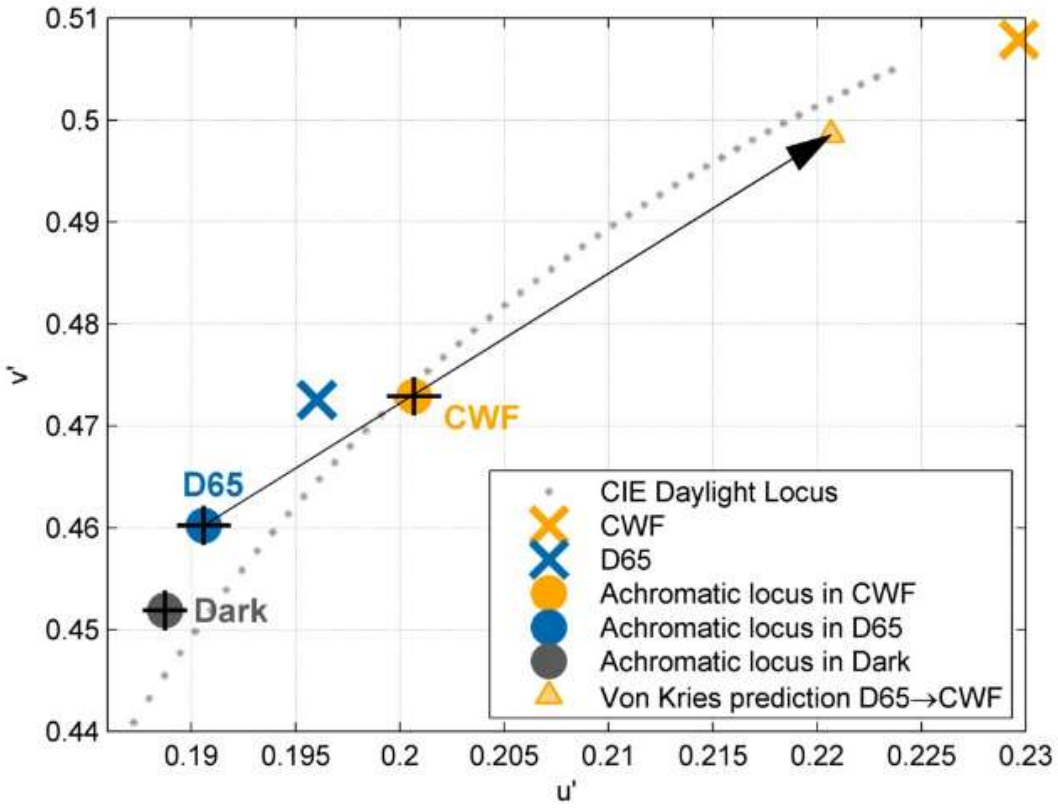


Figure 2.7: Von Kries prediction for the mean achromatic setting when the ambient illumination changes from D65 to CWF. The observed shift is in the same direction as the predicted shift.

Qualitatively similar shifts have been reported by Brainard and colleagues (Brainard, 1998; Brainard & Ishigami, 1995). Figure 2.9 shows a comparison of the achromatic loci obtained in the current study (●,● and ●) with previous experiments. Although there is some variability in the previously reported achromatic loci, generally, the points for all viewing conditions lie close to the daylight locus. The achromatic points under dark viewing condition are close to the equal energy white point, and shifts introduced by a change in illumination along the daylight locus generally follow the daylight locus (Brainard, 1998; Schefrin & Werner, 1993; Valberg, 1971).

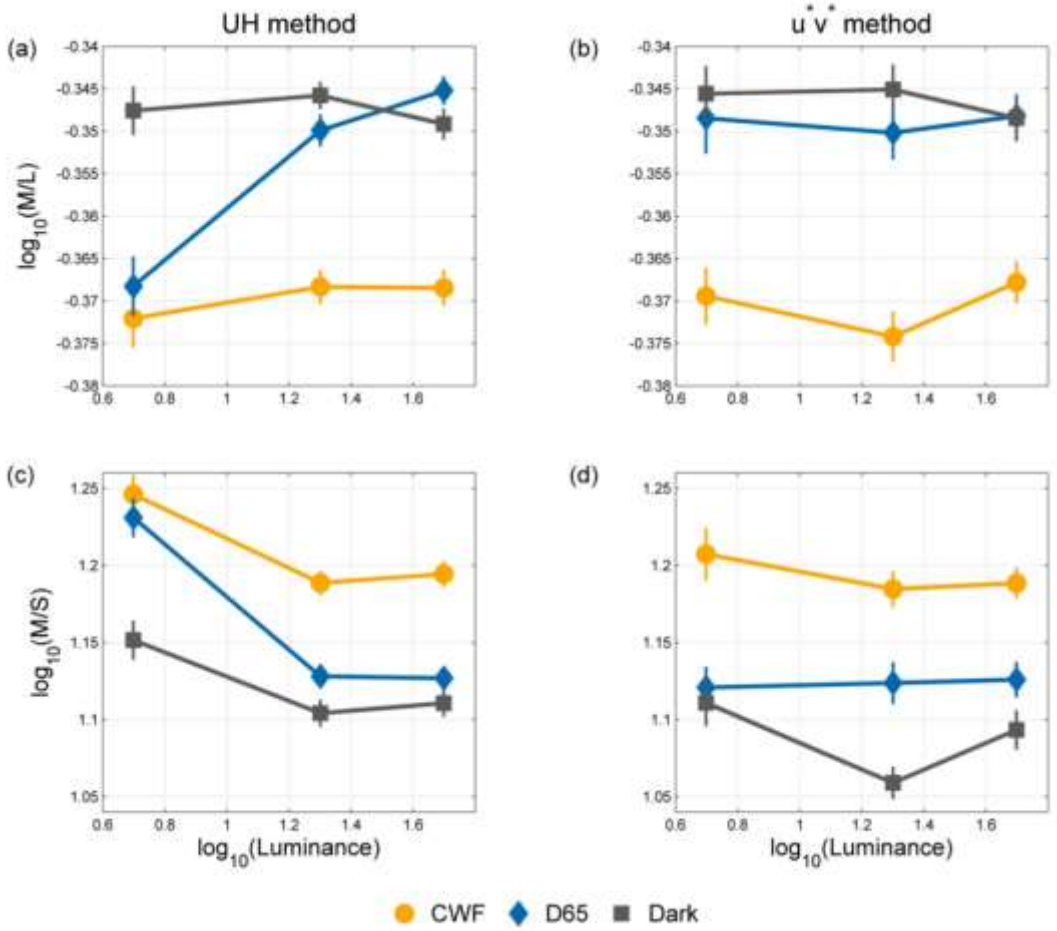


Figure 2.8: Comparison with Kuriki (2006). Effect of stimulus luminance on achromatic settings, in terms of relative weights for M and L cones (a, b) and M and S cones (c, d). The vertical axis shows the relative cone weights for the achromatic settings; different symbols indicate different ambient illumination conditions.

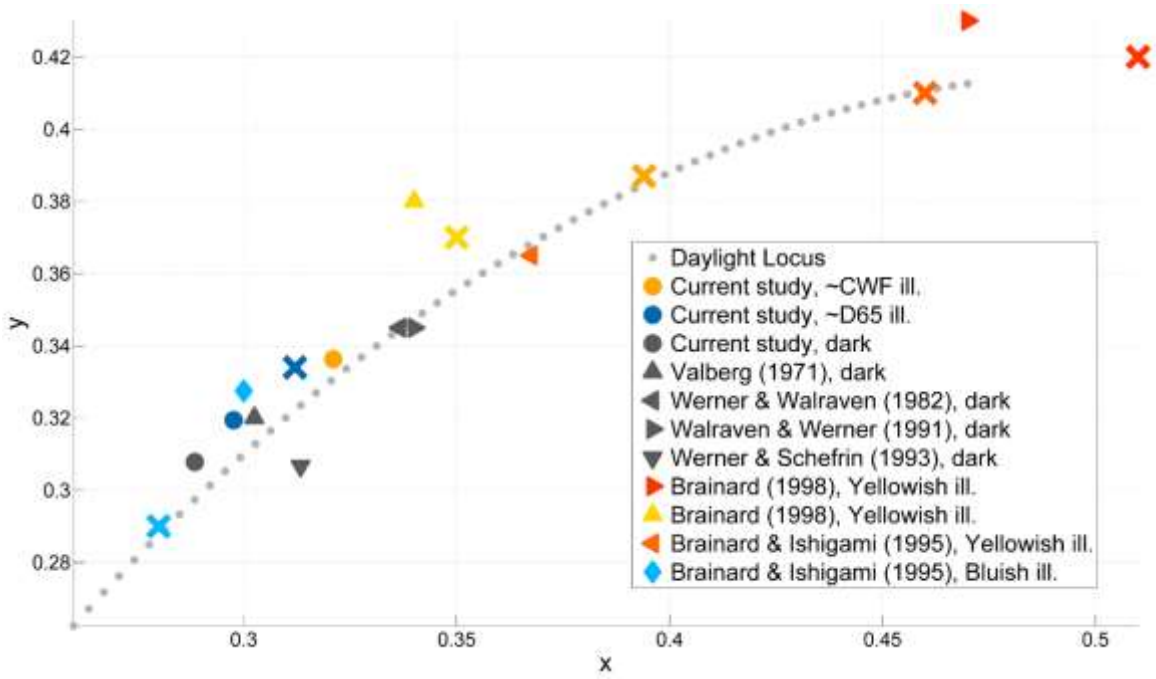


Figure 2.9: Mean achromatic settings reported by previous studies, along with those for the current study. The achromatic settings are denoted by symbols as listed in the legend, while the corresponding ambient illuminations are denoted by a ✖ of the same colour. The grey dotted line indicates the daylight locus. The achromatic loci for the dark condition lie close to the equal energy white point and illuminant shifts along the daylight locus induce achromatic shifts in the same direction.

2.4.2 Covariation along the daylight locus.

Within a particular illumination and test luminance condition, the $u'v'$ settings are not independent (see Figure 2.4), but are aligned with the daylight locus, as reported by Witzel and colleagues (Witzel, Valkova, Hansen, & Gegenfurtner, 2011). A principal component analysis (PCA) of the achromatic settings confirms their results; in Figure 2.10, the direction of the main co-variation (1st principal component; solid grey line) is aligned with the daylight locus (dotted line).

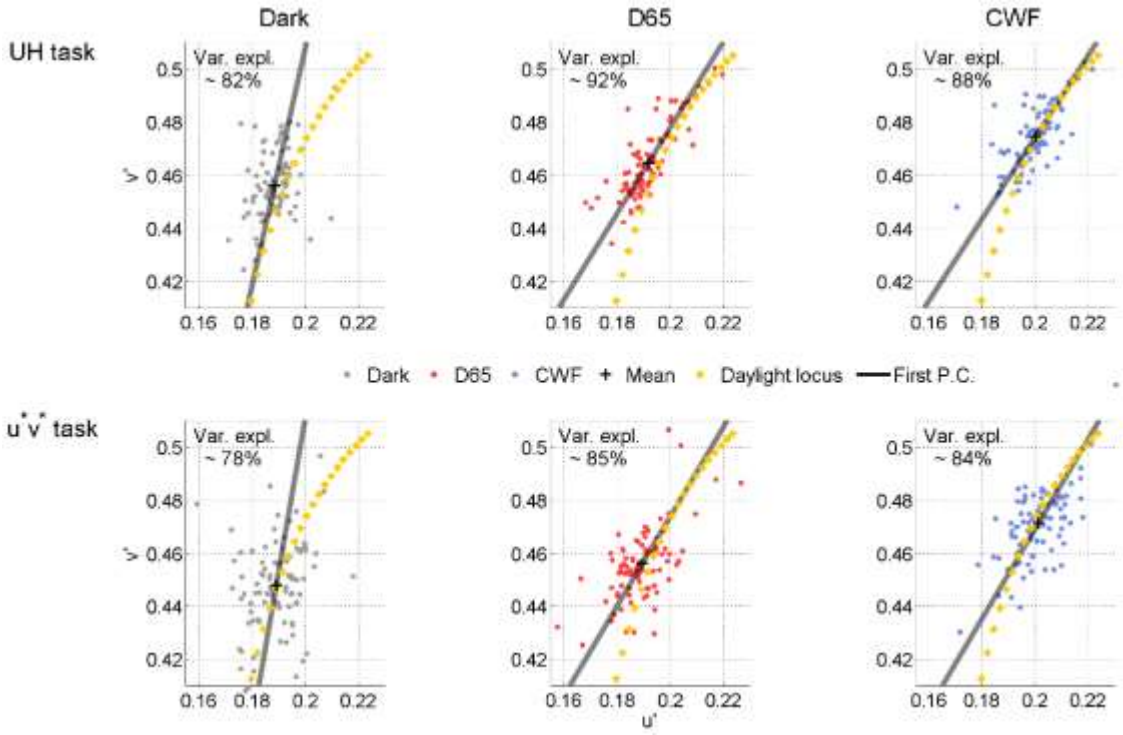


Figure 2.10: The direction of the first principal component (axis of maximum covariance; thick black line) of the achromatic settings in the $u'v'$ plane, shown together with the daylight locus (yellow dotted line). The axis of maximum covariance closely follows the daylight locus.

The illuminations used in this experiment are close to the daylight locus (Figure 2.9); however, it is important to note that the lower variability in the achromatic settings with the UH method is not simply a consequence of the illumination change coinciding with the yellow-blue UH line. Observers are never asked to match achromatic points between illumination conditions; their task is to reach an achromatic point from a randomly chosen starting point for a fixed illumination condition.

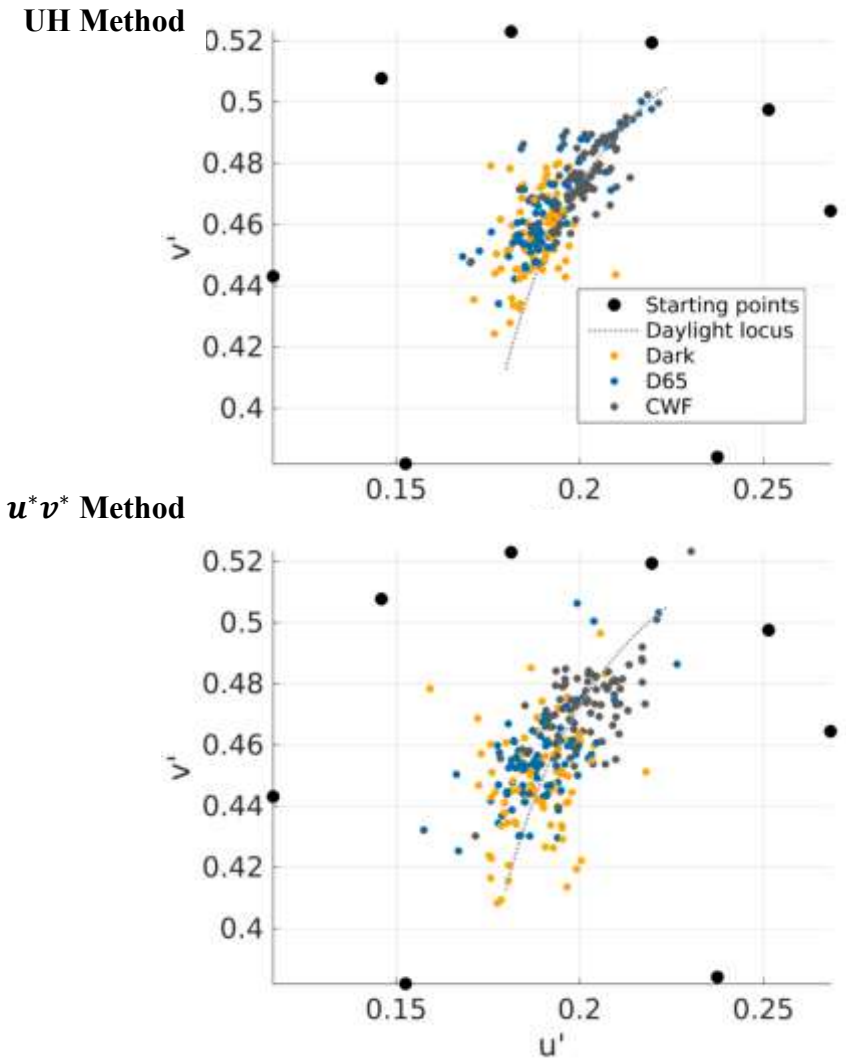


Figure 2.11: Starting points (black filled circles), along with the data for each navigation method (UH and u^*v^*) collapsed across the luminance levels. The starting points were chosen randomly.

These randomly chosen starting points are shown in Figure 2.11, along with the observer settings collapsed across luminance conditions. In most cases adjusting the test patch along the daylight locus will not suffice to obtain an achromatic stimulus since the starting points are not restricted to lie on the daylight locus but vary randomly along both dimensions. It is however conceivable that the unique hue lines, at least the line connecting yellow and blue, are more salient perceptual colour directions, in comparison to u^*v^*, since the unique yellow-blue line reflects a naturally occurring change in illumination, the 'cerulean line' (Mollon, 2006). Whether unique hue lines have any special status in the perceptual colour space is still an open question and not much substantial evidence exists supporting this claim, with the exception of

Danilova's work demonstrating improved discrimination performance for lights close to the unique yellow-blue hue line (Danilova & Mollon, 2010).

In conclusion, this study shows that observers' ability to make reliable achromatic settings is increased when observers are asked to adjust the test patch along the directions defined by the four unique hues as opposed to navigating along the main axes in the commonly used CIE u^*v^* chromaticity plane. While this reduced variability in the achromatic settings does not prove that observers use particular salient mechanisms in the UH method, it is consistent with the idea that internal colour representations are more aligned with the unique hues than the u^* and v^* axes.

The UH task in this experiment was designed using previous measurements of unique hue directions (Wuerger, 2013). An interesting extension to this study would be to measure the unique hue directions of each observer at each luminance level, and to then use these directions in the UH method of navigation for achromatic settings. Under these conditions, the conclusions from the present study would predict a further decrease in the variability of the achromatic settings.

Chapter 3

Modelling Chromatic Adaptation in Unique Hues

Abstract

A large ($N = 185$) set of unique hue settings collected in three different ambient illumination conditions was used to model chromatic adaption using three forms of the transform – diagonal, linear and affine. This study critically examines the accuracy of these models in predicting both, the individual unique hue settings, and the unique hue planes fitted to the data. The structure of the data itself is analysed, especially with respect to the bimodality of the unique hues. The results show that a simple diagonal model in the LMS cone excitation space is the most efficient in terms of the trade-off between accuracy and degrees of freedom. They also show that diagonal and linear models show similar performances, reiterating their theoretical equivalence. The performance of these diagonalisable models: 1. Shows a tendency to deviate from unimodality when evaluated against luminance-dependent metrics 2. Is worse for UR and UG planes compared to UY and UB planes.

3.1 Introduction

Humans are trichromatic, yet their perception of colours is defined by four basic colours – red, green, yellow and blue (Helmholtz, 1867). The photoreceptors in the human retina which absorb long, medium and short wavelengths of light (L, M and S cones respectively) combine in opponency mechanisms (Hering, 1920) to generate these sensations. These perceived sensations of unique colours are referred to as unique hues and have been measured time and time again (Rubin, 1961; Webster, Miyahara, Malkoc, & Raker, 2000; Wuerger et al., 2005). Although chromatic cone-opponency pathways have been identified in the LGN, their cardinal axes have been found to code opponency channels which do not correspond to perceived unique hues (Derrington et al., 1984; Tailby et al., 2008). Stoughton & Conway (2008) suggested a plausible higher-order neural representation of unique hues in the glob cells of the posterior inferior temporal cortex. Due to methodological limitations of their work (Conway & Stoughton, 2009; Mollon, 2009), as it stands, no unequivocally clear neural basis of unique hues has so far been reported (Mollon & Jordan, 1997; Valberg, 2001).

Perceived unique hues have been shown to be fairly robust to diverse factors such as age (Schefrin & Werner, 1990; Wuerger, 2013), gender (Hinks, Cárdenas, Kuehni, & Shamey, 2007) and individual differences in spectral sensitivities (Webster et al., 2000). Thus, in some sense, they could be called pseudo-invariants of perceived colour sensation. It then follows that transforms which can predict the appearance of these unique hues under different illumination conditions are central to the understanding of colour vision, and to the development of a theory of colour appearance in general (Brainard & Wandell, 1992; Fairchild, 2013; Moroney et al., 2002; Xiao et al., 2015).

In the following analyses, a large dataset of unique hues settings (Wuerger, 2013; Xiao et al., 2011) was used to calculate optimised chromatic adaptation transforms. Three forms of the transform were considered – diagonal, linear and affine; and the optimisations were performed in three colour spaces – the cone excitation space LMS, the differential cone excitation space with respect to the background ΔLMS, and the

higher-order contrast opponency space DKL (Derrington et al., 1984; the current implementation was based on Appendix IV of Kaiser & Boynton (1996), which was contributed by David Brainard). The aim of the modelling was to evaluate the best-performing model for adaptation under the given illumination conditions, the performance being defined by an information-theoretic metric. Since unique hues represent the putative third stage in human colour vision (Smithson, 2014), a better understanding of chromatic adaptation models operating on these colours could also suggest possible sites for the adaptations (Dunn, Lankheet, & Rieke, 2007).

3.2 The experiment and the dataset

This section gives an overview of the experimental procedure and the task used to collect the data, ending with a short summary of the data itself. Since the experimental details have previously been described in Xiao et al. (2011) and Wuerger (2013), the aim of this section is a brief recapitulation of the experimental methods used to collect the data.

3.2.1 Experimental set-up

The experiments were carried out in a light-proof anechoic chamber fitted with a Graphic Technology Inc. GLE-M5/32 overhead luminaire. The light reaching the screen in each luminaire mode was measured using a Photo Research© PR-650 spectroradiometer. The measured SPDs (Spectral Power Distributions) of the two illuminants are shown in Figure 3.1.

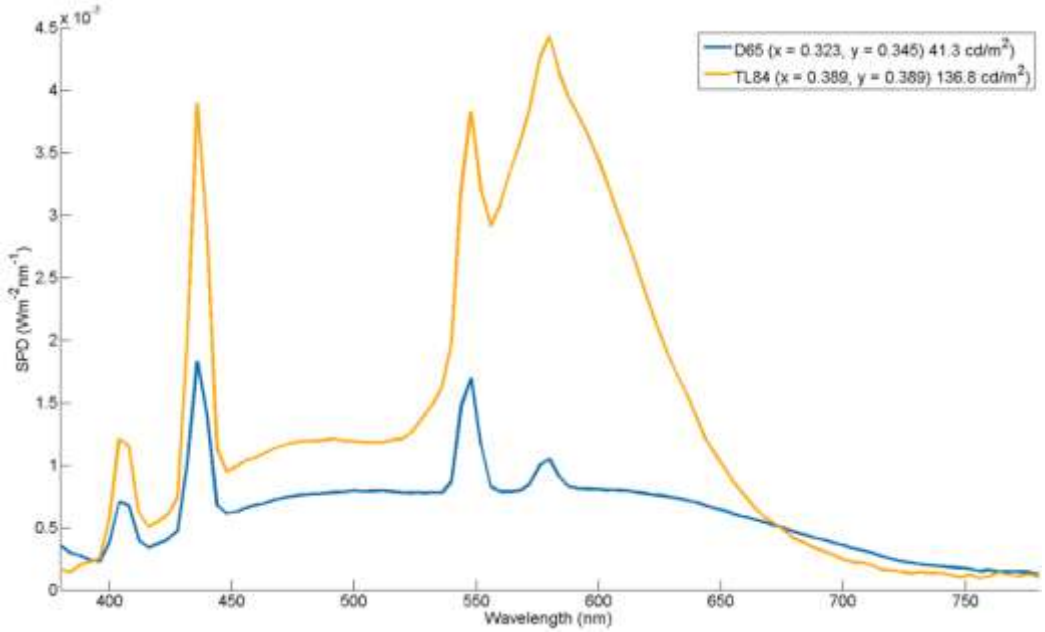


Figure 3.1: Spectral Power Distributions of illuminants used.

The two sources were approximately metameric with standard D65 and F11 (Philips TL84) illuminants and the measured illuminant chromaticities and those of the closest CIE standard illuminants are shown in Table 3.1 below.

	D65		TL84	
	Luminaire	Standard	Luminaire	Standard
Correlated Colour Temperature (K)	~6100	~6500	~3900	~4000
x-y chromaticity	(0.32, 0.34)	(0.31, 0.33)	(0.39, 0.39)	(0.38, 0.38)

Table 3.1: Illuminant chromaticities

The stimuli were displayed on a 21-inch Sony GDM-F520 monitor. A ViSaGe stimulus generator and a ColorCAL colorimeter (both manufactured by Cambridge Research Systems, CRS©) were used to calibrate the screen. The chromaticities and luminance of the screen primaries are presented in Table 3.2. The White Point of the monitor was

theoretically at $\approx 8800K$.

	Red	Green	Blue
x-y chromaticity	(0.63,34)	(0.30,0.61)	(0.15,0.08)
Luminance (cd/m^2)	26.36	73.96	14.82

Table 3.2: CRT Primaries

The stimuli were generated using the CRS© toolbox for MATLAB (Mathworks Inc.) and displayed using a CRS© ViSaGe system. Observer responses were collected using a CRS© CT6 response-box.

3.2.2 Task, stimuli and observers

A modified hue selection task (Wuerger et al., 2005) was used to determine the unique hue settings. The stimuli consisted of ten circular uniform colour patches, spatially arranged along the circumference of a large circle (Figure 3.2). Each circular patch subtended a visual angle of 2° while the annulus eccentricity was about 4°. The background throughout the experiment was a mid-grey (CIE $L^* = 50$ with respect to the monitor white point) with CIE coordinates of $(0.29, 0.30, 24\, cd/m^2)$. For determining the settings for a given unique hue, the instructions given to the observers emphasised the silencing of the alternate colour-opponent mechanism. For instance, to identify unique-red, the observers were asked to choose a patch that contained neither yellow nor blue.

The experiment was conducted in a lightproof anechoic chamber. In each session, observers adapted to the ambient illumination for about 5 minutes before starting the experiment. In each trial, a stimulus with ten circular uniform colour patches arranged in an annulus (Figure 3.2) was shown to the observer. The observer used the response

pad to select one of the ten uniform patches being shown as the required unique hue. After the button press the experiment moved on to the next stimulus. The stimulus presentation was not time-limited and the observers could take as long as they needed to complete the task. Each session involved 108 stimuli presentations and (4 unique hues × 9 lightness-saturation combinations × 3 repetitions) and lasted for about 20 minutes on average.

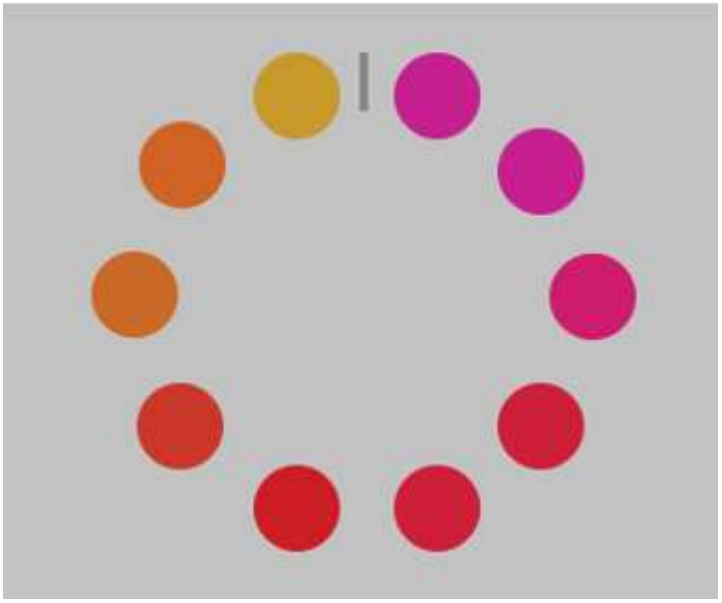


Figure 3.2: A schematic of the hue selection task (not to actual scale or colour accuracy). The circular patches were arranged along an annulus. The patches within a given stimulus were isoluminant and of equal saturation, with a constant hue difference between two adjacent patches. The circular patches subtended 2° of visual angle on the observer's retina while the annulus eccentricity was 4°.

In each trial of the experiment, all ten circular patches on a given stimulus had the same luminance and saturation, and were equally spaced in terms of the CIELUV hue angle. The hue angle steps between the patches (and consequently the range of hues displayed) for each trial was based on hue ranges obtained in earlier experiments by Wuerger et al. (2005). For each unique hue, the observers were tested at nine saturation-brightness combinations, with the brightness ranging from 8.5 cd/m^2 to 60 cd/m^2. Care was taken to ensure that all tested colours were within the gamut of the monitor, resulting in a slightly different set of luminance levels for each unique

hue. Further details, including the exact colour coordinates of the stimuli are available in Xiao et al. (2011).

The three ambient illumination conditions were tested in separate sessions (due to the stabilisation period of the luminaire), with the observers adapting to the ambient illumination for about five minutes before the start of each session. Each observer made each setting three times. In total, 185 observers (age range: 18-75 years, mean age = 32.2 years) were tested during the course of the study. Each observer was screened for colour-normal vision using the Cambridge Colour Test (Regan et al., 1994). The participants were compensated for their time with a small fee.

3.2.3 The dataset

During the course of the experiment, each of the 185 observers made 324 judgements (4 unique hues × 9 saturation-brightness combinations × 3 illumination conditions × 3 repetitions). Throughout the article, the measured data are presented in the colour-opponency based DKL space (Derrington et al., 1984), which is described in detail in Section 3.3.2. The three panels in Figure 3.3 show the average settings for all four unique hues for each observer under the three illuminants. The two axes of the graph represent the S-Luminance and the L-M mechanisms, the plane thereby being what could be called a chromaticity space. The unique hues are colour coded with Unique Red (UR) being in red, Unique Green (UG) being in green and so forth. Averages across observers for each of the 9 saturation-brightness combinations for each unique hue are shown as larger filled circles. The projections of planes passing through these 9 average settings for each unique hue are also shown. The planes were calculated using least-square fitting. Due to the nature of the DKL space, the relative thickness of these planes on the chromaticity diagramme indicates the amount of interaction between the luminance and the chromatic opponency channels. The UG and UR in TL84 show marked interactions.

Figure 3.4 shows a rotation of the same plots such that the luminance variation of these

settings is apparent. For the Dark condition, Xiao et al. (2011) proved that there is indeed an interaction between the luminance, saturation and the hue angle of the unique hues. Later, they also went on to report (Xiao et al., 2015) a difference between Unique hues settings measured using different media such as NCS chips and a cathode-ray-tube monitor. Thus, it must be borne in mind that the settings in the dataset were obtained on a CRT monitor with a constant, grey background.

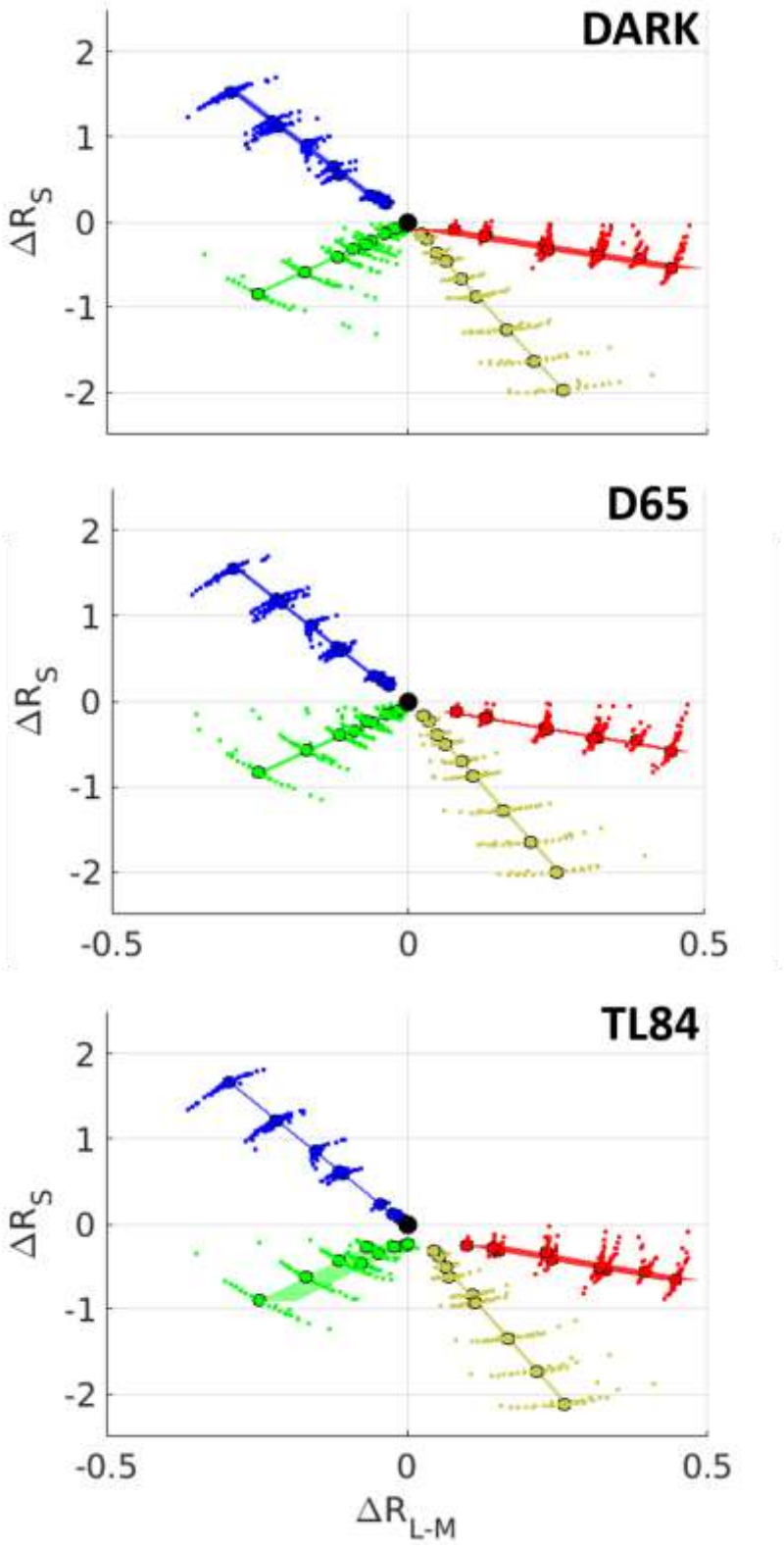


Figure 3.3: Unique Hue settings. Observer settings are shown as colour coded points of the colour corresponding to the Unique Hue. This colour coding is maintained for all points and lines in the graph. Larger colour coded, filled circles represent average settings at a given saturation-brightness combination (9 per Unique Hue). Projections of planes fitted to these average settings are also shown. The filled black circle represents the background.

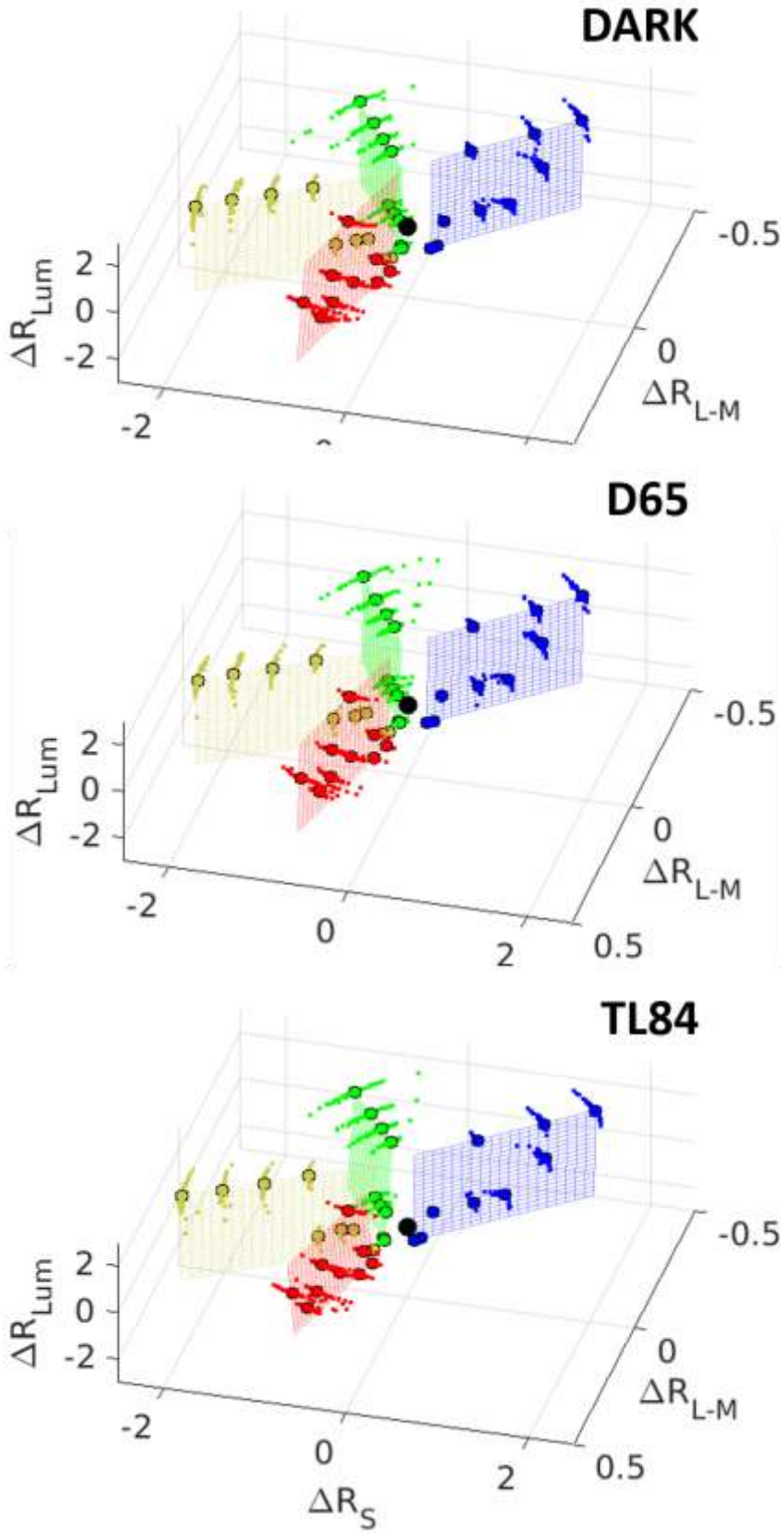


Figure 3.4: Unique Hue settings rotated in colour space to show the luminance variation of the measurements. The luminance settings differ between the unique hues due to gamut constraints imposed by the display. The aesthetics of the graph are identical to Figure 3.3.

3.3 Chromatic adaptation models

3.3.1 Previous work

Colour constancy and the description of colour appearance under changing illuminants has traditionally employed chromatic adaptation transforms based on the Von Kries model (Helmholtz, 1867; von Kries, 1905). Several variations of chromatic transforms optimised for various datasets have been reported (Bianco & Schettini, 2010; MacAdam, 1956; Nayatani et al., 1981; Vazquez-Corral, O'Regan, Vanrell, & Finlayson, 2012). At some stage or the other, all of these transforms use an illumination dependent gain control to adjust the cone excitations for each channel. The limitations of a channel-wise gain control mechanism have been analysed and often criticised theoretically (West, 1979; West & Brill, 1982; Worthey & Brill, 1986) and through experimental data (MacAdam, 1961, 1963; and more recently Kulikowski et al., 2012). Despite this, the diagonal model of gain control has been supported time and time again by studies such as those by Walraven (1976) and Brainard & Wandell (1992). Theoretical arguments like the spectral sharpening hypothesis by Finlayson, Drew, & Funt (1993), and the rank-constraint theory by Chong, Gortler, & Zickler (2007) have established the mathematical equivalence of methods involving linear channel interactions and the traditional diagonal von Kries transform.

3.3.2 Colour spaces

In the present analysis the unique hues dataset described in Section 3.2.3 was used to test different models of chromatic adaptation. These models were tested in three colour spaces: the cone excitation space LMS (see Eq. 1.1), the differential cone excitation space ΔLMS which represents the cone excitations with respect to the background, and the higher order cone-contrast based opponency space DKL (Derrington et al., 1984).

In the LMS space, let $\boldsymbol{x}$ represent the vector of cone excitations elicited by a certain

stimulus displayed on a background with coordinates $\boldsymbol{x_0}$ (also in LMS space). In other words, $\boldsymbol{x}$ denotes a measure of the photonic absorption at the three cone photoreceptors due to the stimulus. If P_L, P_M and P_S are used to denote the absorptions at the long, medium and short wavelength cones, the cone excitation vector can be written as

$$\boldsymbol{x} = \begin{bmatrix} P_L \\ P_M \\ P_S \end{bmatrix} \qquad \text{Eq. 3.1}$$

Similarly, let $\boldsymbol{x_0} = \begin{bmatrix} P_{L0} \\ P_{M0} \\ P_{S0} \end{bmatrix}$ denote the LMS coordinates of the background.

Next, let $\boldsymbol{\Delta x}$ denote the differential coordinates of the same stimulus in the ΔLMS space. Using Eq. 3.1 we can now write

$$\boldsymbol{\Delta x} = \begin{bmatrix} \Delta P_L \\ \Delta P_M \\ \Delta P_S \end{bmatrix} = \boldsymbol{x} - \boldsymbol{x_0} \qquad \text{Eq. 3.2}$$

Here, the symbol $\boldsymbol{\Delta}$ in ΔP_L etc. simply denotes the differential nature of the coordinates. Coordinates in this space represent very early retinal processing such as the contrast coding carried out by horizontal cells, albeit without feedback processing.

The third set of colour coordinates used in this analysis was the DKL system. The DKL coordinate system describes early visual processing based on contrast-opponency mechanisms. As mentioned earlier, the current implementation was based on Appendix IV of Kaiser & Boynton (1996), contributed by David Brainard. This model proposes opponent coding channels based on luminance, L - M cone opponency, and the S cone - luminance opponency. Throughout this chapter, these channels are denoted by the symbols ΔR_{Lum}, ΔR_{L-M} and ΔR_S respectively, the ΔR in these symbols simply

implying a differential channel response.

The calculation of DKL coordinates involves the computation of cone contrast signals, their appropriate combination and weighting based on opponency constraints, and most crucially, the normalisation of channel isolating stimuli such that a channel isolating stimulus with a unit pooled-contrast produces a unit response in the intended DKL channel.

Let $\boldsymbol{\Delta r}$ denote the signal in DKL space such that

$$\boldsymbol{\Delta r} = \begin{bmatrix} \Delta R_{Lum} \\ \Delta R_{L-M} \\ \Delta R_S \end{bmatrix} \quad \text{Eq. 3.3}$$

where the opponent channels are represented by the corresponding subscripts: Lum (luminance), $L-M$ (red-green opponency) or S (yellow-blue opponency). Furthermore, using previously introduced symbols (Eq. 3.1 and Eq. 3.2), one could write the contrast signal as

$$\boldsymbol{c} = \begin{bmatrix} \mathrm{C}_L \\ \mathrm{C}_M \\ \mathrm{C}_S \end{bmatrix} = \begin{bmatrix} \Delta P_L / P_{L0} \\ \Delta P_M / P_{M0} \\ \Delta P_S / P_{S0} \end{bmatrix}, \quad \text{Eq. 3.4}$$

where C_L, C_M and C_S are the contrasts defined for the three cone classes. The pooled-contrast can now be defined as the magnitude of this contrast signal

$$\|\boldsymbol{c}\|_2 = \sqrt{C_L^2 + C_M^2 + C_S^2} \quad \text{Eq. 3.5}$$

After the combination, weighting and normalisation of the differential signal using opponency and unit-pooled contrast constraints, the final form of the transform for the conversion of differential cone signals to DKL signals can be calculated to be as follows

$$\boldsymbol{\Delta r} = \begin{bmatrix} \frac{\sqrt{3}}{L_0 + M_0} & \frac{\sqrt{3}}{L_0 + M_0} & 0 \\ \frac{\sqrt{L_0^2 + M_0^2}}{L_0(L_0 + M_0)} & -\frac{\sqrt{L_0^2 + M_0^2}}{M_0(L_0 + M_0)} & 0 \\ -\frac{1}{L_0 + M_0} & -\frac{1}{L_0 + M_0} & \frac{1}{S_0} \end{bmatrix} \boldsymbol{\Delta x} \qquad \text{Eq. 3.6}$$

Eq. 3.6 is a closed form representation of the unit-normalisation procedure for pooled contrasts outlined by David Brainard, and does not appear directly in the reference (Appendix IV, Kaiser & Boynton, 1996).

3.3.3 Proposed adaptation models

In each of these colour spaces, models similar to those tested by Brainard & Wandell (1992) were used to model chromatic adaptation in the data. The models used in the analysis were the diagonal model, the simple linear model and the affine model. A summary of these models is provided in Table 3.3, with $\boldsymbol{x}$ and $\boldsymbol{y}$ representing the input and output colour coordinates in a particular colour space. $\boldsymbol{\varepsilon}$ denotes the channel noise vector.

The diagonal model is a simple channel gain mechanism. In the LMS space, it corresponds to what is often referred to as an ideal von Kries transform (von Kries, 1905). In this model, adaptation is represented by channel-specific gains which lie along the diagonal of the transform matrix.

The linear model extends the idea of the diagonal model by allowing for channel cross-talk. This cross-talk allows each channel to influence adaptation in the other channels. One possible hypothesis explaining this cross-talk is the spectral sharpening of cone sensitivities (Finlayson et al., 1993), and a number of adaptation models using linear sharpened transforms have been proposed (Finlayson & Süsstrunk, 2000; Vazquez-Corral et al., 2012). It must be noted here though, that all linear transforms can

essentially be *whitened*, i.e., the channels can be decorrelated using diagonalisation methods such as Singular Value Decomposition, Principal Component Analysis and rank-constraint procedures on measurement tensors (Chong et al., 2007).

Model	Form	Parameters
Diagonal	$\boldsymbol{y} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \boldsymbol{x} + \boldsymbol{\varepsilon}$	6
Linear	$\boldsymbol{y} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \boldsymbol{x} + \boldsymbol{\varepsilon}$	12
Affine	$\boldsymbol{y} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \boldsymbol{\varepsilon}$	15

Table 3.3: The models tested. The adaptation stimuli are represented as $\boldsymbol{x}$ and the test stimuli are represented as $\boldsymbol{y}$. The linear matrix is denoted by $\boldsymbol{A} = [a]_{ij}$ and the translation component in the affine transform by $\boldsymbol{b} = [b]_i$. The channel noise is denoted by the vector $\boldsymbol{\varepsilon}$. The third column lists the number of parameters for each model, including the channel-errors.

The third and final model considered here was the affine model. It differs from the diagonal and linear models in that it cannot be explained solely in terms of three diagonalisable gain controlled channels. Brainard & Wandell (1992) described this model as a generalisation of the two-process model proposed by Jameson & Hurvich (1964). This two-step structure of adaptation has classically been explained in various ways. Walraven (1976) and Shevell (1978) interpreted the additive term as a result of simultaneous contrast mechanisms. Jameson, Hurvich, & Varner (1979) later proposed that the multiplicative term of the transform could denote the fast receptoral or early post-receptoral gain mechanisms, while the additive term could explain the slower changes in the biases of the neural opponency-based mechanisms.

Finally, it must be noted that the parameters of each model depend on the specific

illumination change being modelled.

3.3.4 Model fitting and evaluation

Transforms between 6 pairs of illumination conditions were modelled, corresponding to pairwise permutations of the D65, TL84 and Dark conditions in the dataset. In each of the three colour spaces (LMS, ΔLMS and DKL), for each permuted pair of illumination conditions, three mappings (corresponding to the three models being tested) of the settings were calculated for each observer. These mappings or model-predictions were then compared to the actual measured settings in order to derive error and efficiency criteria of the models for each observer. This process is expressed in detail below.

Let

1. $f_{i,j}^{n}$ denote one of the transforms (Diagonal, Linear or Affine) which models settings of the n^{th} observer from illumination condition i to j. Here i and j can be any two non-equal illumination conditions out of Dark, D65 and TL84. 6 such pairings of illumination conditions are possible in total.
2. $\boldsymbol{r_i^n}$, $\boldsymbol{g_i^n}$, $\boldsymbol{y_i^n}$ and $\boldsymbol{b_i^n}$ be the average (over the three repetitions) unique hue settings for unique red, green, yellow and blue respectively, made by the n^{th} observer in illumination condition i. These settings are assumed to be in one of the three colour spaces: LMS, ΔLMS or DKL. For a given observer and illumination condition, each of these sets contains 9 individual settings corresponding to nine luminance-saturation combinations tested.
3. $\boldsymbol{x_i^n} = \cup\{\boldsymbol{r_i^n}, \boldsymbol{g_i^n}, \boldsymbol{y_i^n}, \boldsymbol{b_i^n}\}$ be the set of all unique hue settings made by the n^{th} observer in illumination condition i. Thus, for each observer in a given illumination condition, this set contains 36 settings, 9 per unique hue.

The optimal transform $\tilde{f}_{i,j}^{n}$ which optimally transforms all settings made by the n^{th} observer in illumination condition i to illumination condition j can now be written as

$$\tilde{f}_{i,j}^{n} = \min_{f_{i,j}^{n}} \left\| \boldsymbol{x_j^n} - f_{i,j}^{n}(\boldsymbol{x_i^n}) \right\|_2 \qquad \text{Eq. 3.7}$$

There are three points to note here. First, each optimised f denotes an optimal transform of a certain form (Diagonal, Linear or Affine) in a certain colour space (LMS, ΔLMS or DKL). Thus, any given f denotes one of nine possible model-colour space combinations. The current study models all nine combinations.

Second, $\tilde{f}_{i,j}^{n}$ is optimised separately for each of the n observers, and each of the six ordered pairs of illumination conditions (i, j). Thus, any perceptual error or information-efficiency metric must be calculated for each observer and each pair of illumination conditions. In this study, no pair of illumination conditions was preferred and all metrics were averaged over all six possible pairs of illumination conditions, thus giving a unique error or efficiency index for each observer. Let $\Lambda_{i,j}^{n}(\tilde{f}_{i,j}^{n})$ denote a metric (such as the $CIELAB$ colour difference or the Akaike information criterion) calculated over $\tilde{f}_{i,j}^{n}$. Since no particular pair of illumination conditions is preferred, an adaptation-averaged value of the metric is calculated such that

$$\widehat{\Lambda}^{n} = \frac{1}{N(N-1)} \sum_{i=1}^{3} \sum_{\substack{j=1 \\ j \neq i}}^{3} \Lambda_{i,j}^{n} \qquad \text{Eq. 3.8}$$

where $N = 3$ is the total number of illumination conditions. All analyses in this chapter were performed on this adaptation-averaged form of the metrics.

Third, $\tilde{f}_{i,j}^{n}$ and $\tilde{f}_{j,i}^{n}$ must be invertible for a symmetric model. Under constraints of optimality, one does not expect this requirement to be difficult to satisfy.

The optimisation calculations were performed in MATLAB (Mathworks Inc.). The diagonal model was optimised using the pre-implemented Simulated Annealing algorithm in the Global Minimisation Toolbox. For faster computation in case of the

linear and affine models, an anisotropic scaling extension of the Procrustes analysis developed by Paláncz, Zaletnyik, Awange, & Heck (2010) was implemented and used. Generalised Procrustes matching traditionally solves the problem of finding a similarity transformation (i.e., assuming rotation, translation and isotropic scaling of the three axes, with no anisotropy) for two sets of data points, and the algorithm used here is an extension of the ABC algorithm (Awange, Bae, & Claessens, 2008).

Since the optimisation process in the three different colour spaces optimised Euclidean distance within that space, to compare results across these models, the data and the model-predictions were converted to the $CIELAB$ colour space (CIE, 2004) – a perceptual colour space with a widely used and industrially accepted colour-difference metric. The calculation of $CIELAB$ coordinates is further explained in Appendix 1.1. The traditional ΔE_{LAB} colour-difference and Δh_{ab} hue-angle difference metrics were used to evaluate the models. If $\boldsymbol{l_i^n} = [L_i^n\ a_i^n\ b_i^n]^T$ and $\boldsymbol{l_j^n} = \left[L_j^n\ a_j^n\ b_j^n\right]^T$ are the $CIELAB$ colour coordinates of the settings made by the n^{th} observer for a given unique hue under illumination conditions i and j respectively, these metrics are defined as

$$\Delta E_{LAB} = \left\| \boldsymbol{l_i^n} - \boldsymbol{l_j^n} \right\|_2 \qquad \text{Eq. 3.9}$$

$$\Delta h_{ab} = \cos^{-1}\left(\frac{\left| a_i^n a_j^n + b_i^n b_j^n \right|}{\sqrt{\{(a_i^n)^2 + (b_i^n)^2\} \cdot \{\left(a_j^n\right)^2 + \left(b_j^n\right)^2\}}} \right) \qquad \text{Eq. 3.10}$$

In addition, the angle between the least-squares fitted planes for measured and predicted unique hue settings was also calculated. If $\widehat{\boldsymbol{\eta}}_{\boldsymbol{1}}$ and $\widehat{\boldsymbol{\eta}}_{\boldsymbol{2}}$ are unit vectors normal to the measured and predicted planes for a given unique hue UH, the angle $\Delta\theta_{UH}$ between them is calculated as

$$\Delta\theta_{UH} = \cos^{-1}(\widehat{\boldsymbol{\eta}}_{\mathbf{1}} \cdot \widehat{\boldsymbol{\eta}}_{\mathbf{2}}) \quad \text{Eq. 3.11}$$

The above metrics give us a method of comparing the results of the models, but do not offer any information on the model performance in terms of sufficiency of parameters and over- and under-fitting. Since the models have a large difference in their degrees of freedom, the small-sample correction of the Akaike Information Criterion (Akaike, 1974) recommended by Burnham & Anderson (2004), denoted here by AIC_c, was used to evaluate the performance of the models.

The Akaike Information Criteria or AIC is an information theoretic measure which evaluates models on the basis of a trade-off between the goodness of fit and the complexity of the model. It is given by

$$AIC = 2k - 2\log(\mathcal{L}(\widehat{\boldsymbol{\theta}}|\boldsymbol{x})) \quad \text{Eq. 3.12}$$

where k is the number of estimated parameters, and $\mathcal{L}(\widehat{\boldsymbol{\theta}}|\boldsymbol{x})$ represents the maximised likelihood of the model described by estimated parameters $\widehat{\boldsymbol{\theta}}$, given the data $\boldsymbol{x}$. The maximised likelihood of a model is a measure of the goodness of fit of the model given a particular set of data.

Two observations about the nature of the AIC can be made through an examination of Eq. 3.12. First, as the AIC varies inversely with the goodness of fit $\mathcal{L}(\widehat{\boldsymbol{\theta}}|\boldsymbol{x})$, one can say that a lower AIC indicates a better fit. Second, if two models fit a given dataset equally well, AIC penalises the model with the higher number of free parameters (as AIC increases with an increase in k). In other words, AIC is a measure of model performance which incorporates a trade-off between the goodness of fit and model complexity, with lower values indicating a better performance.

In cases where the number of parameters is comparable to the number of samples n

(recommended strongly for $k/n < 40$, Burnham & Anderson, 2004), a corrected version, often denoted by AIC_c (C. Hurvich & Tsai, 1989; Sugiura, 1978) is used.

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1} \qquad \text{Eq. 3.13}$$

For linear least-square estimations under the assumption of normally distributed errors, this definition reduces to

$$AIC_c = n \log\left(\frac{\sum \epsilon_i^2}{n}\right) + 2k + \frac{2k(k+1)}{n-k-1} \qquad \text{Eq. 3.14}$$

where $\sum \epsilon_i^2$ simply denotes the residual-sum-of-squares from the fitted model.

3.4 Results

The AIC_c values averaged over all observers and all 6 permutations of the adaptation and test illumination conditions are shown in Figure 3.5. Each panel shows the performance of the models for one of the three colour spaces. The error-bars mark the standard error across observers.

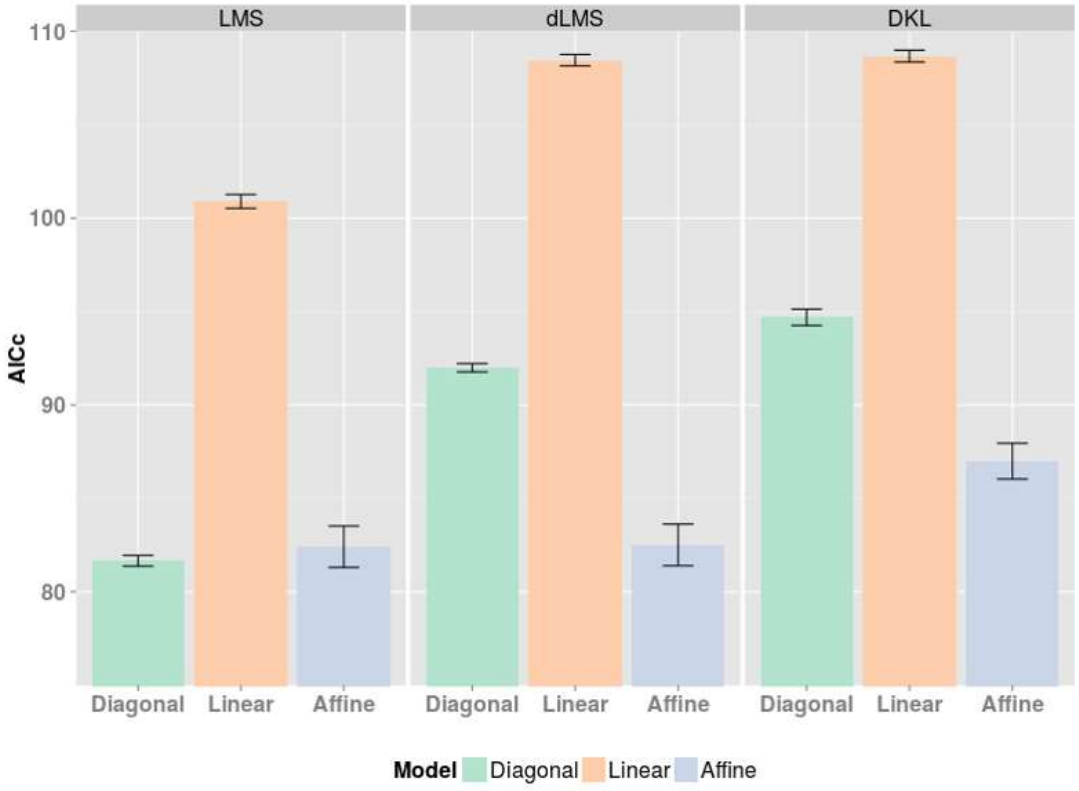


Figure 3.5: AIC_c for the models in LMS, ΔLMS and DKL spaces.

A 2-way repeated-measures ANOVA confirms a clear effect of both the colour space and the model ($p < 0.001$ in both cases) with a strong interaction (also $p < 0.001$). The best performance (corresponding to the lowest AIC_c) is given by the Diagonal model in the LMS space. It must be noted though, that the Affine model outperforms the other two models in both the DKL and the ΔLMS spaces, and is very close in mean performance to the Diagonal model in the LMS space as well.

Since reciprocal pairs of illumination conditions were modelled, in each case, for each observer, the invertibility of the optimised models between corresponding pairs was also investigated, and reciprocal pairs of transforms were indeed found to be nearly invertible to one another.

Although the AIC_c gives us a means of choosing a model, it is also interesting to see how each model optimises the mappings in terms of the colour (ΔE_{LAB}) and hue-angle

(Δh_{ab}) differences. Figure 3.6 shows boxplots and superimposed violin plots for ΔE_{LAB} colour differences between the predicted and measured settings for each model. The box edges in the plot denote the first and third quartiles while the central line denotes the median. The extents of the upper and lower whiskers mark ± 1.5 times the inter-quartile range, while the notch around the median line (often found to be extremely small in the present study) indicates 95% confidence-intervals. The superimposed violin plot shows the kernel density estimation of the data. The lowest ΔE_{LAB} values were found for the Affine model in all colour-spaces. Thus, an increased number of parameters does, as expected, lead to better fitting. It is also interesting to note that the performances of the Diagonal and Linear models are quite similar, once again an expected result if one considers their theoretical equivalence through *sharpening* transforms.

Analogous to the previous figure, Figure 3.7 shows hue-angle differences between the predicted and measured unique hue settings. Here too, the Affine model was found to show the smallest hue-angle deviation between the predicted and measured values, although the three models have very similar performances.

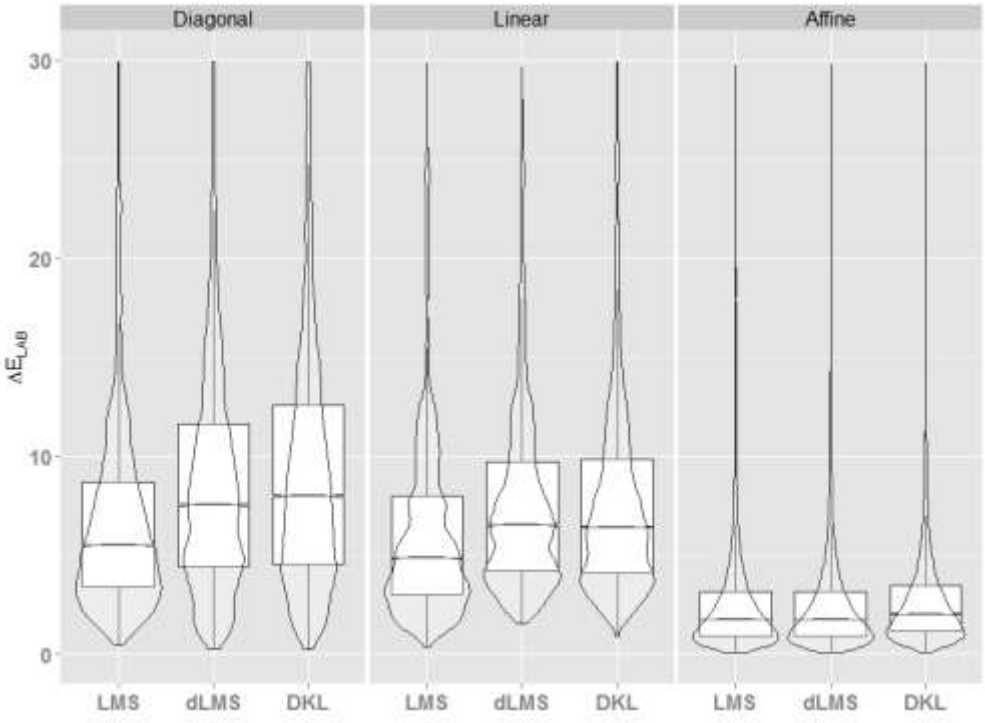


Figure 3.6: ΔE_{LAB} colour difference between the model-predictions and the measured settings. The box edges denote first and third quartiles while the central line denotes the median. Upper and lower whiskers mark ± 1.5 times the inter-quartile range.

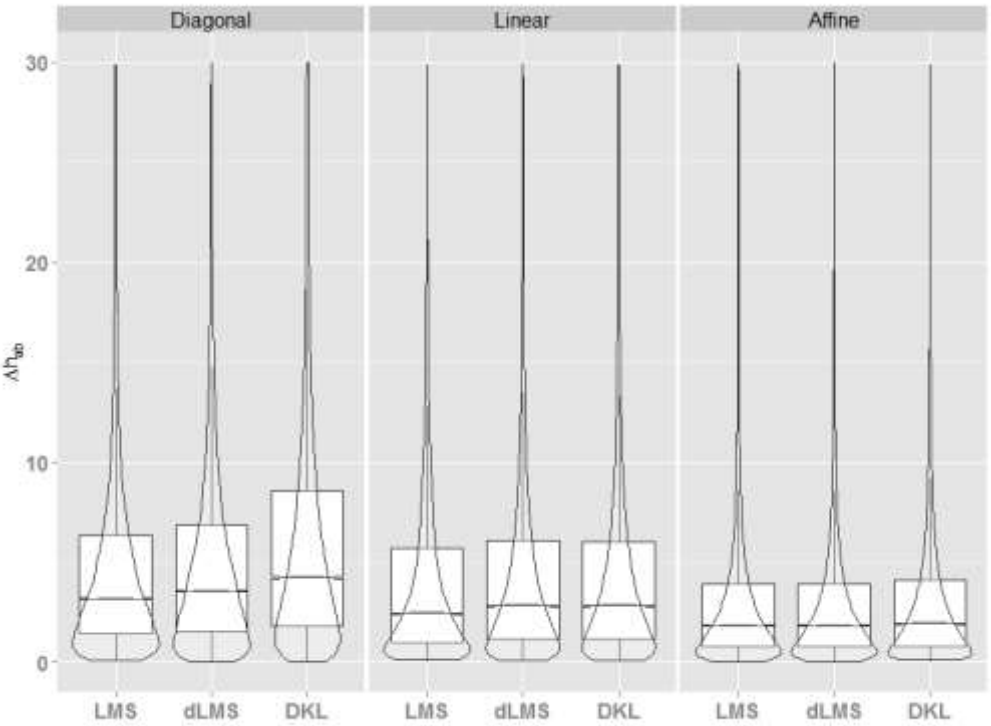


Figure 3.7: Δh_{ab} hue-angle difference between the model-predictions and the measured settings. The boxplots denote the same statistics as in Figure 3.6.

The colour and hue difference measures compare the unique hue settings predicted by the models and the actual settings. Here, one must note that even though the $CIELUV$ colours of the stimuli in different illumination conditions were matched, there is no reason to expect a one-to-one correspondence between the settings made by the observers due to added noise from the surroundings as well as the observer's own visual system. Assuming that the underlying unique-hue mechanisms are relatively robust, it is perhaps more interesting to compare the deviation of the predicted unique hue mechanism planes from the planes fitted to the measured data. The angles between the normal vectors of the predicted and the measured unique-hue planes are shown in Figure 3.8. Once again, one finds that the Affine models outperforms the other two in all three colour spaces.

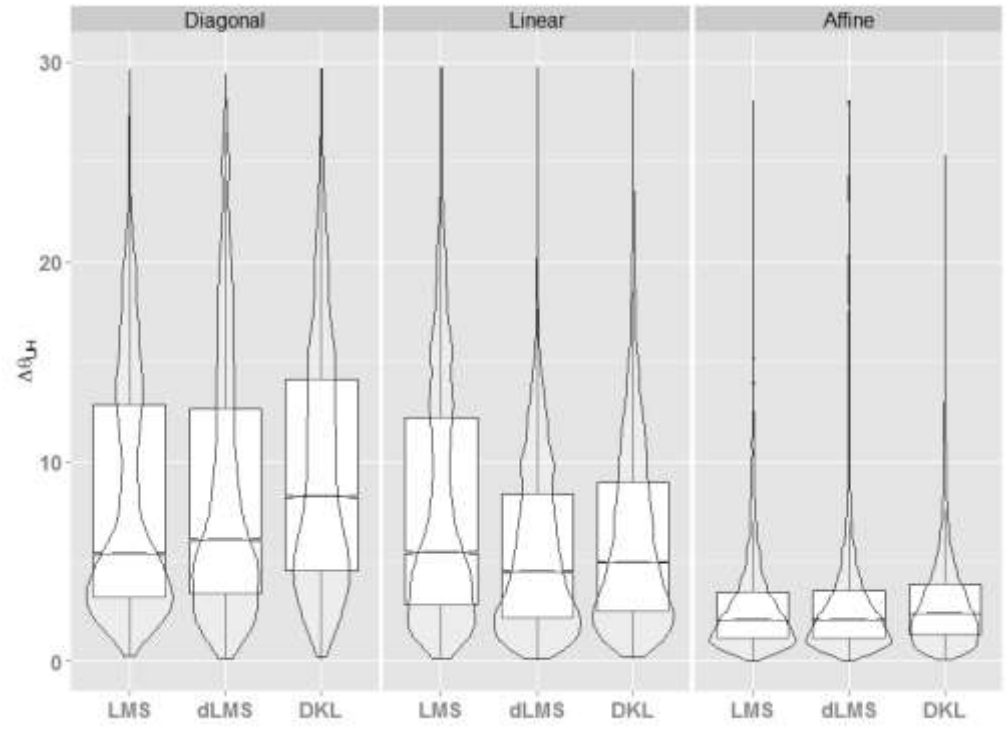


Figure 3.8: Angle $\Delta\theta_{UH}$ between the normal vectors of unique hue planes predicted by the models and the best-fit planes for the measured settings.

3.5 Discussion

This study evaluated the performance of three models with different degrees of freedom in predicting illuminant-dependent changes in unique hue settings in three physiologically relevant colour-spaces. The Diagonal model is a direct form of the classical Von Kries law, the Linear model can be viewed as a sharpened (Finlayson et al., 1993; Finlayson & Süsstrunk, 2000) form of the diagonal model, and the Affine model describes a two-step theory involving not only scaling, but a slower, induced lateral interaction (Brainard & Wandell, 1992; Jameson & Hurvich, 1964).

Figure 3.5 shows that the best performance in terms of the AIC_c criterion was achieved in the LMS space by the Diagonal model. It was almost equalled by the Affine model, albeit with a higher inter-observer variance. This suggests possible over-fitting in case of the Affine model, a possibility that is further supported by Figures 3.6, 3.7 and 3.8, where the Affine model shows nearly perfect median fits for all three metrics – the ΔE_{LAB}, the Δh_{ab} and $\Delta\theta_{UH}$. Thus, in terms of purely fitting *efficiency*, the Diagonal model delivers the best accuracy-complexity trade-off.

Furthermore, it is also interesting to note that for metrics which are functionally dependent on luminance differences, i.e. ΔE_{LAB} (Figure 3.6) and $\Delta\theta_{UH}$ (Figure 3.8), the Diagonal and Linear models showed a distinct trend towards deviation from a compact unimodal distribution, and secondary optima can be observed. This is not the case for hue-angle differences (Figure 3.7) which showed a strictly unimodal distribution for each model. This reaffirms the relevance of the issues raised by Kulikowski et al. (2012) concerning the bias of Von Kries adaptation transforms towards chromatic optimisation, and the consequent sub-optimality in their estimates of lightness constancy.

The plausibility of bimodality holds a special significance in unique hue literature, especially for unique green UG (L. Hurvich, Jameson, & Cohen, 1968; Rubin, 1961; Volbrecht, Nerger, & Harlow, 1997; Welbourne, Thompson, Wade, & Morland, 2013).

To further investigate how the distributions of these metrics were affected by the local structure of the unique hue data, the figures were re-plotted with separate boxplots and violin plots for each unique hue. The plots for ΔE_{LAB} and Δh_{ab} (Figure 3.9 and Figure 3.10 respectively) do not show noticeable differences between individual unique hues, except slightly higher values of ΔE_{LAB} for UR in the Diagonal and Linear models. For $\Delta\theta_{UH}$ (Figure 3.11), on the other hand, one observes that in the LMS and ΔLMS spaces, the Diagonal and Linear model predictions are significantly worse for the UR and UG compared to the UY and UB.

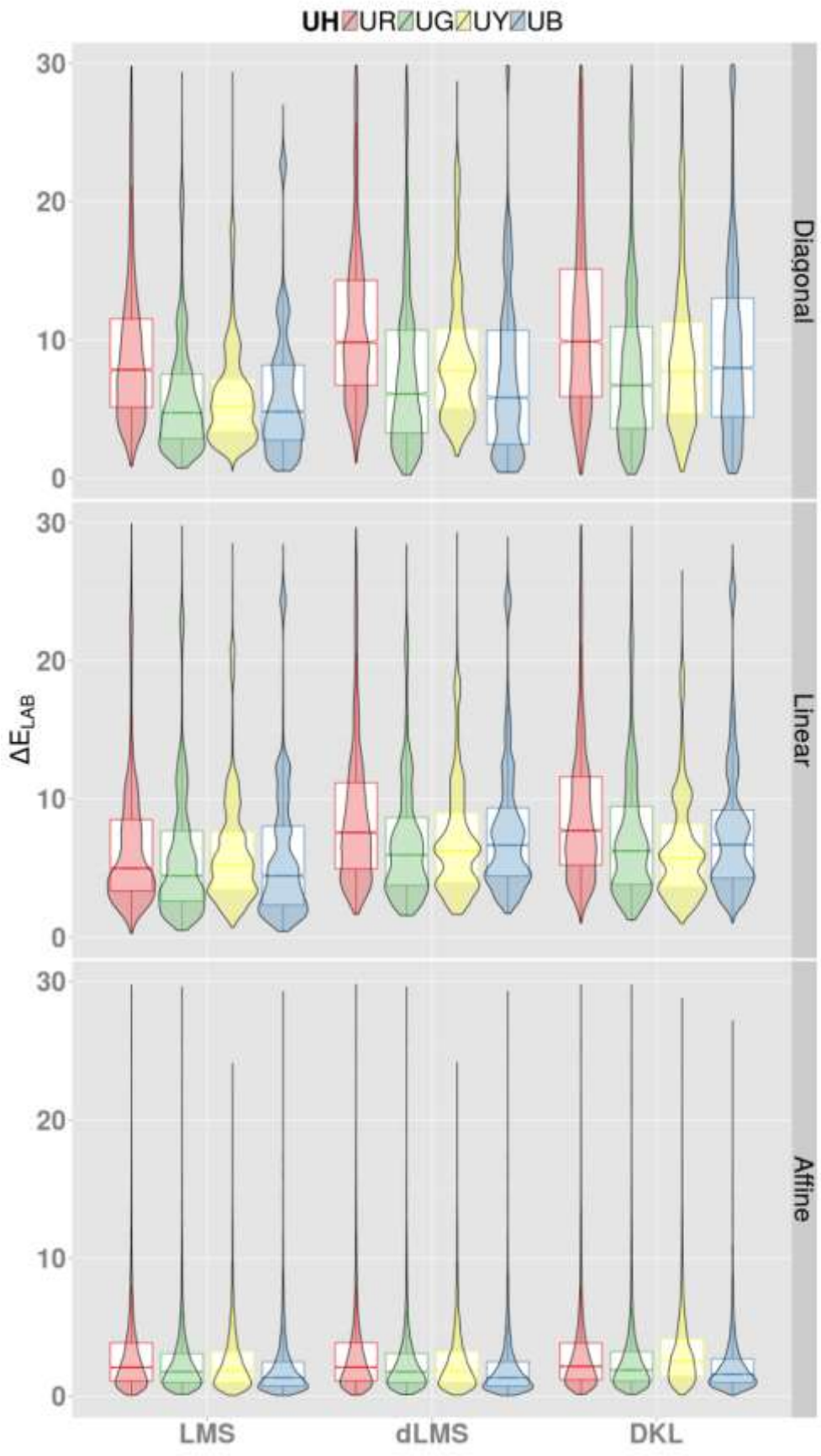


Figure 3.9: ΔE_{LAB} colour difference between the model-predictions and the measured settings.

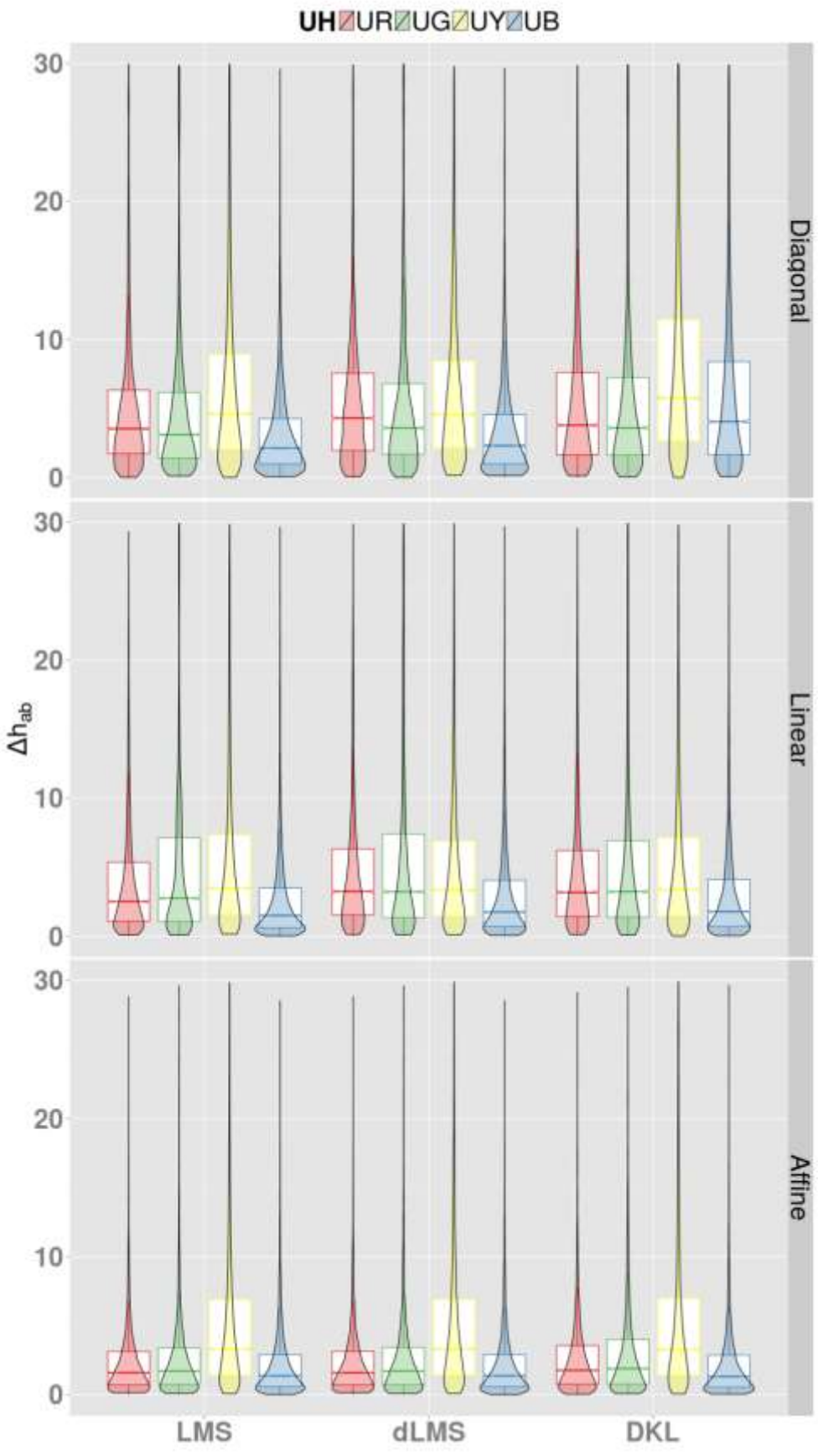


Figure 3.10: Δh_{ab} hue-angle difference between the model-predictions and the measured settings.

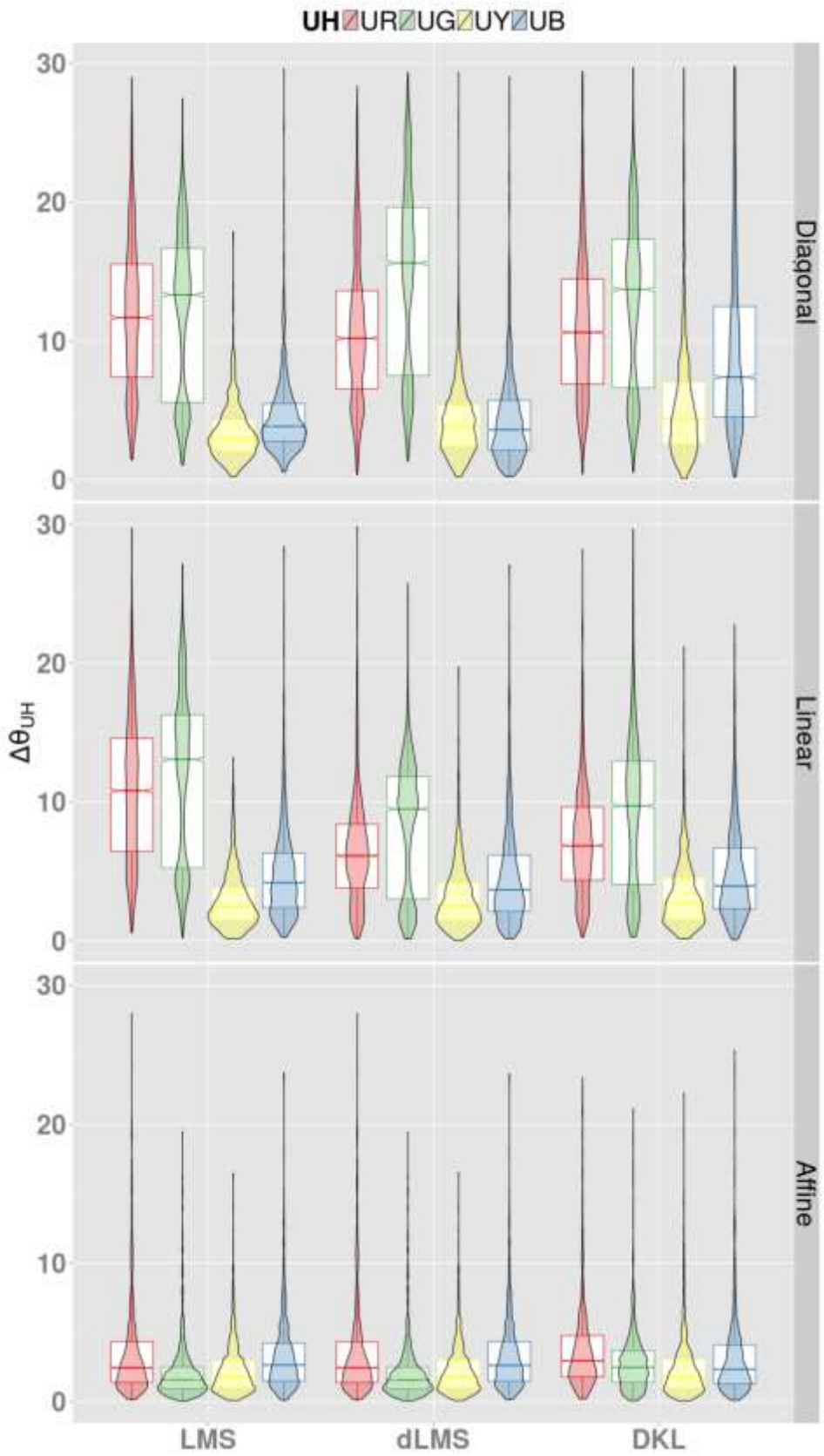


Figure 3.11: Angle $\Delta\theta_{UH}$ between the normal vectors of unique hue planes predicted by the models and the best-fit planes for the measured settings.

In other words, using purely diagonal or diagonalisable transforms in linear cone excitation spaces it is possible to relatively accurately model adaptations in the UY and UB mechanisms, but not the UR and UG mechanism planes. To investigate this further, the measured unique hues settings were plotted as a function of hue angle in the chromaticity plane of the DKL space. This is shown in Figure 3.12 where the colour of the points code for the unique hue. It is to be noted that the aim of this figure is to show the hue-angle spread, and the plotted points correspond to settings across all nine saturation-luminance combinations for each unique hue.

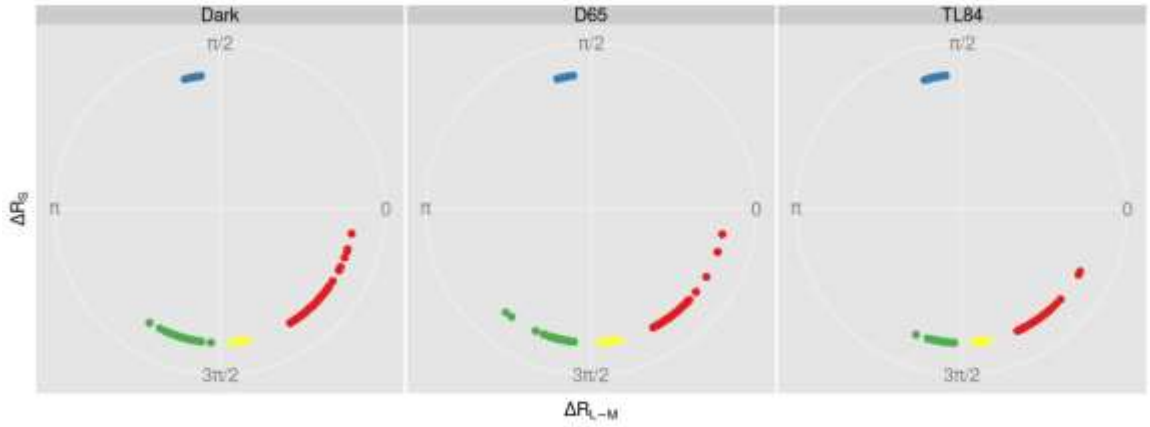


Figure 3.12: Measured unique hue settings as a function of phase in the chromaticity plane of the DKL space. This aim of this diagram is to merely illustrate the extent of the hue-angle variation for each unique hue. It does convey the kernel densities.

.

The UR and UG do indeed display a higher spread if one were to consider only hue angles in a colour-opponent space. While the UY and UB show a total spread of about $5°$, UR and UG show spreads of almost $20° - 25°$. Kuehni (2014), in his review article analyses several datasets of unique hues and reports a similar difference in spreads. To further see how these angles are distributed, and to investigate the often debated bimodality of UG, the kernel density functions for these spreads are plotted in Figure 3.13. The ten highest and lowest hue-angles for each set have also been labelled with the corresponding observer-number. The y-axis across the 4 unique hue panels have different scales as the figure is aimed solely at illustrating the densities of the spreads.

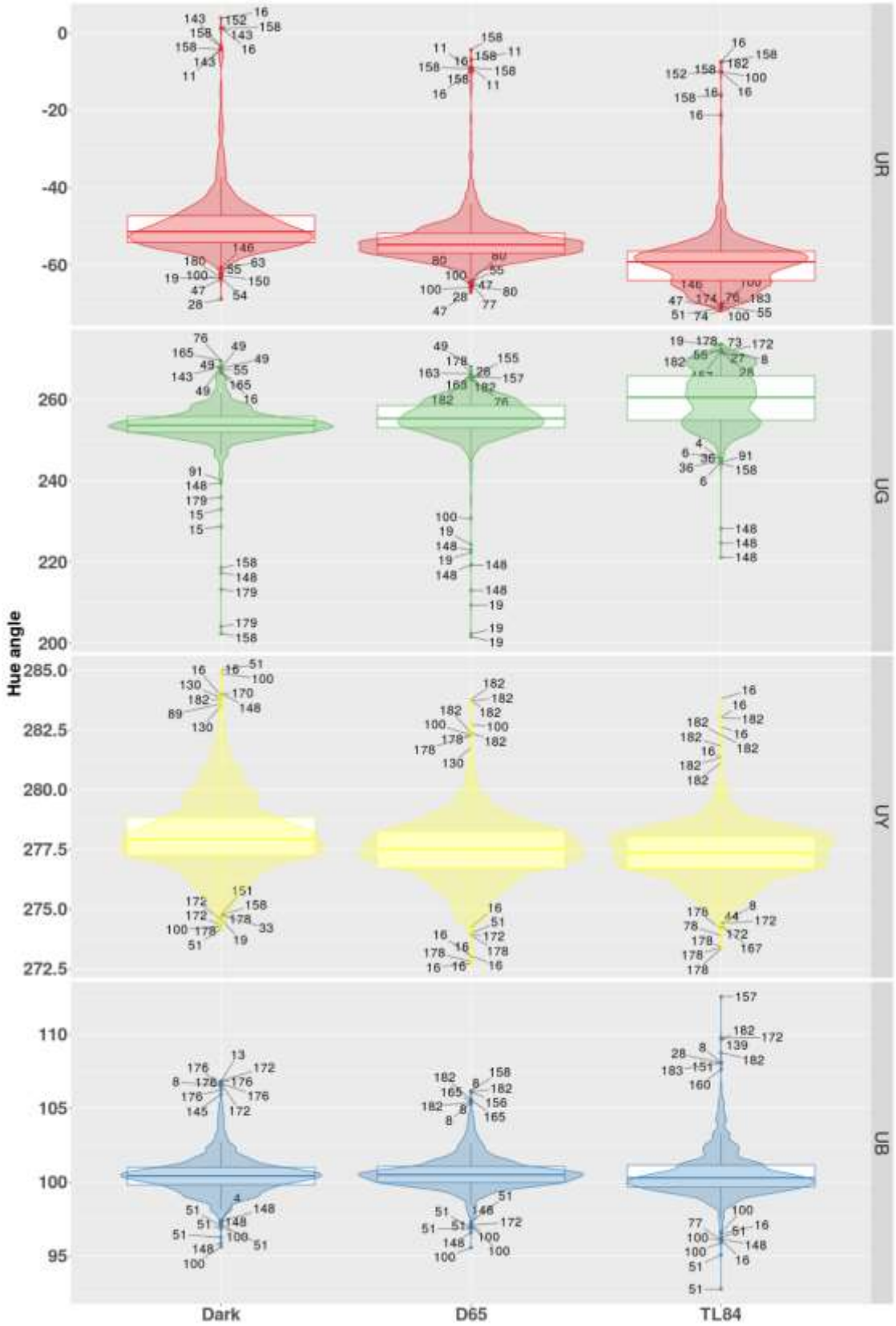


Figure 3.13: Kernel density functions and boxplots for the spreads in Figure 3.12. Note that in order to clearly show the shapes of the density functions, the scales for the 4-subpanels are not matched.

No bimodality is observed for the UG in the Dark and D65 conditions, but the TL84 condition does indeed show some signs of a multimodal distribution. Furthermore, if one examines the outliers, it can be seen that within a given unique hue, some observers consistently show settings which deviate considerably from the median, e.g., Observer 158 for UR and Observer 148 for UG. Since the plots show all 9 settings made by each observer in each condition, some of the observers show multiple settings which are outliers.

The data and the fitted models offer two, not necessarily mutually exclusive, explanations of the previous reports of bimodality. First, as pointed out by Volbrecht et al. (1997), the difference in distributions might arise out of differences in experimental setups, in particular, luminance differences and the control of lightness adaptation. One would expect a well-controlled isoluminant study to produce primarily unimodal distributions. Second, what could perhaps be misconstrued as bimodality in very large datasets (Rubin's dataset had 278 participants) are the exceptionally long tails of UG densities, which in the present dataset extend to about $10° - 20°$ of DKL hue angle under simulated daylight. Although these tails signify what would be classed as outliers, it is interesting to note that the number of these outliers is much higher for UR and UG when compared to UY and UB, which have relatively compact density functions.

The non-unimodal performance of the Diagonal and Linear models (in both linear and contrast based colour spaces) on the UR and UG planes, combined with a long-tailed but primarily unimodal density function for these unique hues (except UG in TL84) suggests three possibilities. First, that the adaptation does not follow a diagonalisable mechanism i.e., it is non-linear – a view which is not wholly uncommon in applied colour science (MacAdam, 1961, 1963; Moroney et al., 2002; Nayatani et al., 1981). Second, that the adaptation does not follow a three-channel optimisation, but is derived from adaptations along multiple selectively tuneable channels (Webster & Mollon, 1994, 1997). Third, that the UR and UG surfaces themselves are non-linear. Although the first and the third possibilities are mathematically very similar, the origin of the non-linearity in the former case lies in the adaptation mechanism, while in the latter

case it is in the unique hue mechanisms. To further examine the current dataset in the light of the above possibilities, Figure 3.14 shows the hue-angles for the settings in each condition as a function of the luminance in DKL coordinates. The plot also shows the best fit line in each case. The hue angles of the unique hue settings show a clear variation with the luminance of the settings, with steeper slopes for UR and UG compared to UY and UB. While this does not directly clarify whether the adaptation mechanisms involve multiple chromatic tunings or are non-linear, it strongly supports the possibility that the unique hue mechanisms themselves are non-linear, especially for UR and UG.

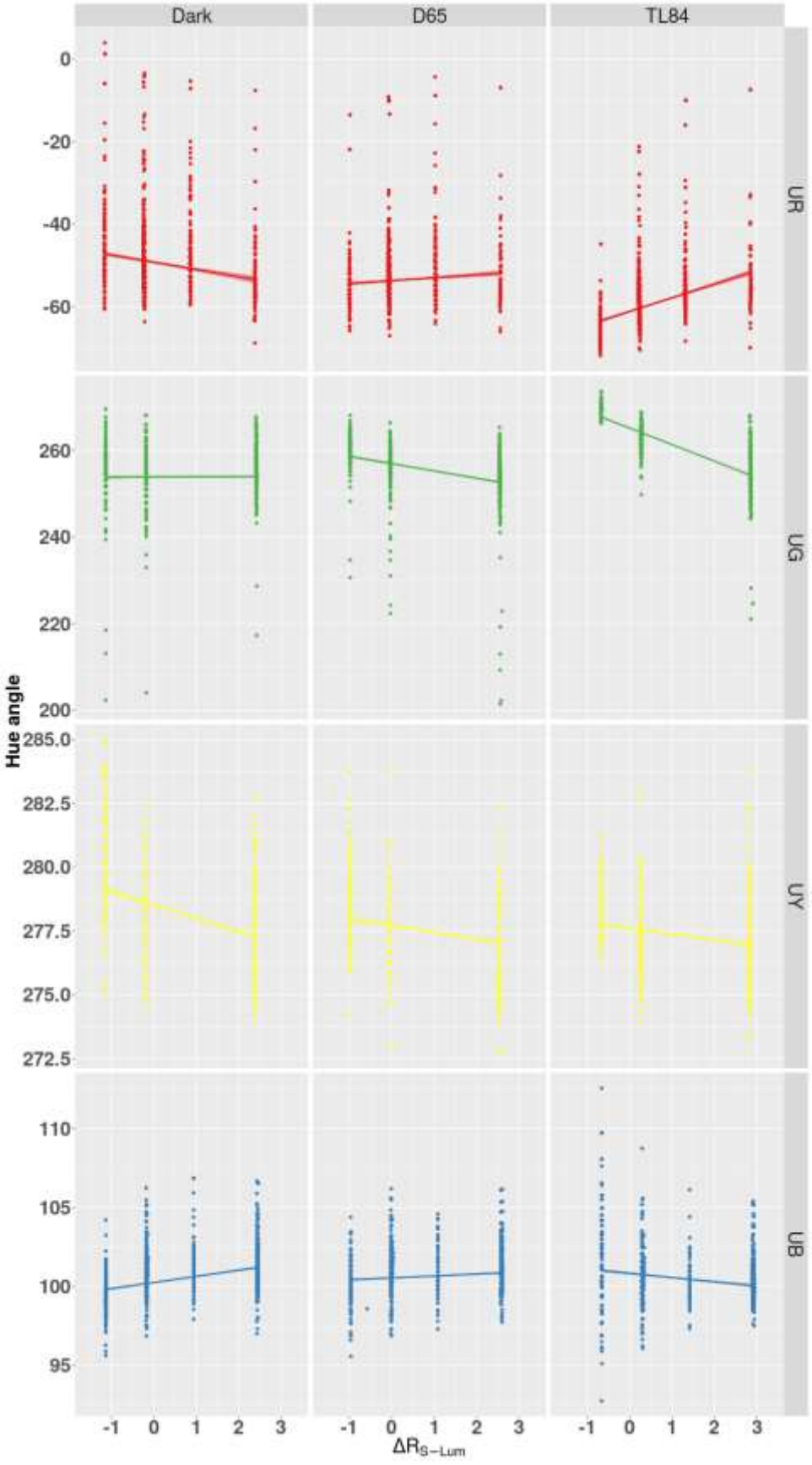


Figure 3.14: Hue angles of the unique hue settings plotted as a function of their luminance in DKL space. The best-fit line in each condition is also shown.

3.6 Conclusion

Chromatic adaptation was modelled using a large ($N = 185$) unique hue dataset which was collected in three illumination conditions – simulated daylight, cool white fluorescent and dark. The simple diagonal form of the Von Kries model in the LMS cone excitation space was found to be the most efficient in terms of the AIC_c criterion which balances the accuracy and complexity of the models. In terms of prediction accuracy, the model with the highest number of parameters, i.e. the Affine model, gave the best performance for all tested metrics. The performances of the Diagonal and the Linear models were found to be very similar to each other, reiterating their theoretical equivalence. A deviation from strict unimodality in the density functions for the ΔE_{LAB} and $\Delta\theta_{UH}$ metrics, both of which take luminance mechanisms into account, was also observed. No such deviations from unimodality were noted for the Δh_{ab} metric which operates only along chromaticity planes. Poorer fits were also found for UR and UG planes using Diagonal and Linear transforms. The analysis of the dataset in terms of the hue-angles for individual unique hues shows a much larger spread for UG and UR compared to UY and UB, with the UG showing possible multimodality under TL84 lighting. The hue angles of the settings also show a dependence on the luminance, suggesting possible non-linearities in the unique hue mechanisms.

Ideal data to tease apart some of the questions raised by this analysis would be isoluminant sets of unique hue settings collected over a large number of observers.

Chapter 4

Discrimination thresholds for skin images

Abstract

In the past, colour discrimination thresholds have been measured under various conditions for uniform colours (MacAdam, 1942; Melgosa et al., 1997; Poirson et al., 1990). To a lesser extent, attempts have also been made to estimate similar thresholds for natural textures (Giesel et al., 2009; Hansen et al., 2008). The first experiment in this set of studies was aimed at extending these measurements to polychromatic skin images in a 3-D chromaticity-luminance colour-space. Thresholds for uniform patches were also measured and the thresholds for the polychromatic stimuli were found to be consistently higher, for both the chromatic, and the luminance projections. In the second and third experiments, the effects of ambient illumination and the location of the stimuli in colour space were investigated. It was found that the observed trends match those reported by Hansen et al. (2008) very closely, with the area of the chromaticity ellipses showing a gradual increase with distance from the illumination chromaticity. The orientations of these chromatic ellipses for simulated skin were found to align with the vector joining the mean chromaticity of the patch and the illuminant chromaticity. This effect is not observed for the other textures, even though they have the same second and higher order statistics as simulated skin.

4.1 Introduction

Skin colour and texture is used by humans in processing and accomplishing a variety of tasks such as face recognition (Bar-Haim, Saidel, & Yovel, 2009), judgements of health (Stephen, Law Smith, Stirrat, & Perrett, 2009) and evaluation of attractiveness (Fink et al., 2008; Fink, Grammer, & Thornhill, 2001; Stephen et al., 2009). Communication of skin colour has also been proposed as one of the main driving factors in the evolution of the human colour vision (Changizi et al., 2006). Besides its biological and evolutionary importance for humans, it is also regarded as an important component of realistic machine vision algorithms (Cula, Dana, Murphy, & Rao, 2004; Jones & Rehg, 2002), computer graphics (Crichton, Pichat, Mackiewicz, Tian, & Hurlbert, 2012; Giard & Guitton, 2010) and animation (Jimenez, Sundstedt, & Gutierrez, 2009). Skin appearance also plays a very important role in the evaluation of dermatological treatments (Khemis, Kaiafa, Queille-Roussel, Duteil, & Ortonne, 2007) and prosthetics (Al-Harbi, Ayad, Saber, ArRejaie, & Morgano, 2015; Fine & Dent, 1978; Hungerford, Beatty, Marx, Simetich, & Wee, 2013). To these ends, various models of skin appearance have been proposed: models based on machine learning and image processing (Tsumura et al., 2003; Vezhnevets, Sazonov, & Andreeva, 2003), physics-based multi-layered models (Chen, Baranoski, Kimmel, & Miranda, 2015; Cula et al., 2004; Donner, Weyrich, D'Eon, Ramamoorthi, & Rusinkiewicz, 2008; Krishnaswamy & Baranoski, 2004), and, most recently, a model that uses imaging of skin microbiomes to predict appearance (Kaur, Dana, & Cula, 2015).

Due to its importance as a natural polychromatic texture which could very well have been a key factor in the evolution of human colour vision (Changizi et al., 2006), an investigation of how humans perceive skin under various conditions could contribute greatly to how human vision, in general, processes and is constrained by properties of natural textures. Thomson & Foster (1997) showed that in a discrimination task, human observers were found to be preferentially sensitive to second and higher order statistics in natural images. In fact, natural scenes and textures have been proposed as important factors which constrain and shape the properties of the human visual system (Geisler, 2008; Nascimento et al., 2002; Regan et al., 2001; Webster & Mollon, 1994,

1997). Constraints based on the reflectance spectra of natural surfaces have also been used to provide rigorous descriptions of colour phenomena such as the asymmetry in the unique hues and frequencies of linguistic colour categories (Philipona & Regan, 2006).

Montag & Berns (2000) compared luminance thresholds for textures and uniform patches and reported a two-fold increase in the luminance thresholds for textures. More recently, Hansen, Giesel, & Gegenfurtner (2008) and Giesel, Hansen, & Gegenfurtner (2009) also estimated chromatic thresholds in isoluminant planes for natural objects and uniform patches. This chapter presents estimations of discrimination thresholds in a luminance-chromaticity 3-D colour space for another salient natural texture, namely, human skin. Besides skin images, thresholds were also estimated for textures derived from the relative colour distribution of skin, and furthermore, for each polychromatic patch that was tested, thresholds for a corresponding uniform patch with the same mean colour were also measured.

In the first experiment, the discrimination thresholds for ecologically valid simulations of skin patches from two ethnicities were estimated under dark, cool white fluorescent, and simulated daylight illumination conditions. Since the human visual system has evolved under natural daylight (Mollon, 2006), an interesting question that arises is whether the performance of the human visual system would significantly change if the skin texture were to be viewed under a non-natural illuminant. The choice of simulated daylight – a naturalistic illuminant, and white fluorescent light – an artificial illuminant, were driven by this question. Furthermore, studies (Xiao et al., 2012) have shown that the gamut of skin colours within a given ethnicity is highly constrained. For this reason, representative skin patches from two ethnicities – Caucasian and Oriental, were used prepare the stimuli used in the experiment.

In the subsequent two experiments the effects of the illumination condition and the location of the stimulus in the colour space were investigated. Here, ecologically invalid stimuli were tested to investigate how the discrimination thresholds change

when the stimulus has the same relative distribution of colours as skin, but with a displaced mean. The stimuli were ecologically invalid because after the displacement of the mean they no longer represented valid simulations of skin.

4.2 Methodology

This section gives methodological details common to all four experiments described in this chapter. Experiment-specific details are described in the corresponding sections to avoid confusion.

4.2.1 Lighting conditions

The experiments were carried out in a light-proof anechoic chamber fitted with a Graphic Technology Inc. GLE-M5/32 overhead luminaire. Two illumination modes from the overhead luminaire were used in the experiments – metameric daylight and cool-white-fluorescent light. In the first two experiments, an additional 'dark' condition was also used, wherein the luminaire was switched off.

The light reaching the screen in each luminaire mode was measured using a Photo Research PR-650 spectroradiometer and a standard white reflective tile. The white tile was placed at the same position on the screen as the centre of the stimulus. The measured SPDs (Spectral Power Distributions) of the two illuminants are shown in Figure 4.1.

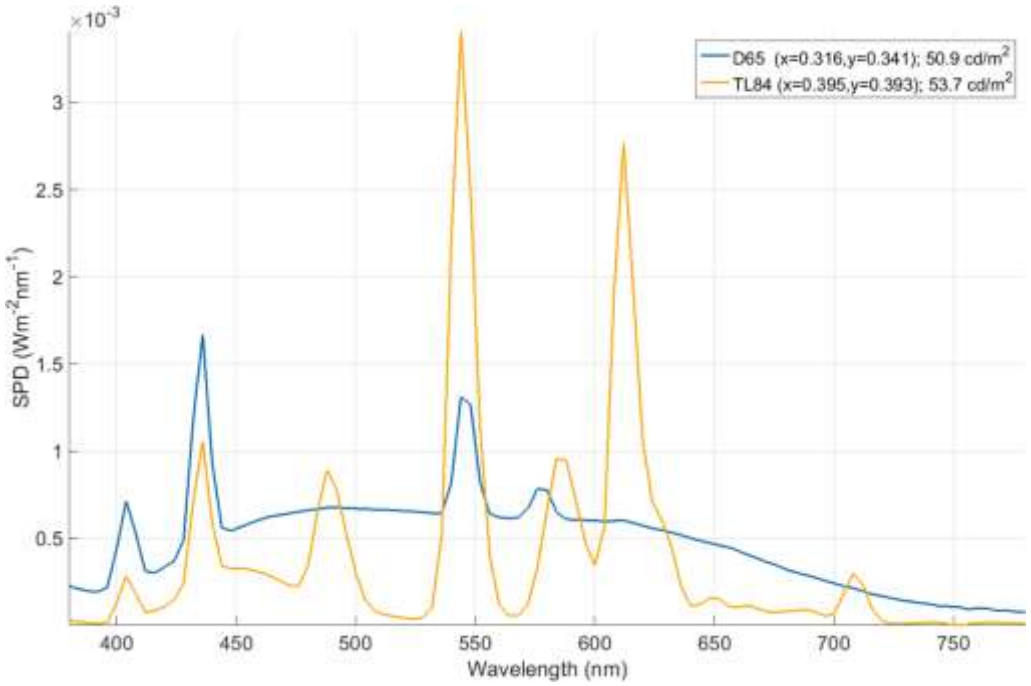


Figure 4.1: Spectral Power Distributions of illuminants. They are different from Figure 3.1 due to the change in luminaire lamps between the two studies.

The two sources are approximately metameric with standard D65 and F11 (Philips TL84) illuminants. The measured illuminant chromaticities and those of the closest CIE standard illuminants are shown in Table 4.1 below.

	D65		TL84	
	Luminaire	Standard	Luminaire	Standard
Correlated Colour Temperature (K)	~6100	~6500	~3900	~4000
x-y chromaticity	(0.32, 0.34)	(0.31, 0.33)	(0.39, 0.39)	(0.38, 0.38)

Table 4.1: Illuminant chromaticities.

4.2.2 The task and stimulus generation

Figure 4.2: The stimulus (exaggerated colour difference).

In all experiments, thresholds were estimated using a 4-AFC (4-Alternative Forced Choice) task. Four skin patches were simultaneously displayed on the screen, of which three were copies of the *reference image*, while one – the *test patch* – differed in colour (Figure 4.2). The observer's task was to indicate the odd-one-out by pressing the corresponding button on a response box. The *test patch* was generated by adding a *test vector* in 3-D $CIELAB$ colour space to each pixel of the original patch. An example of this is shown in Figure 4.3.

The $CIELAB$ space was chosen because of its wide acceptance as a uniform colour space. Although numerous other uniform colour spaces exist, e.g. DKL (Derrington et al., 1984), $CIE\ 1976\ UCS$ (CIE, 2004) , Boynton-McLeod (MacLeod & Boynton, 1979); for the purposes of this experiment an effective sampling of the space was more important than its structural properties, so long as it described an approximately uniform colour space. The white point used for the normalisation of this $CIELAB$ space was fixed as the white-point of the display used for the experiment, with $CIE\ 1931\ xyY$ coordinates of $[0.28\ 0.30\ 106.1445]^T$

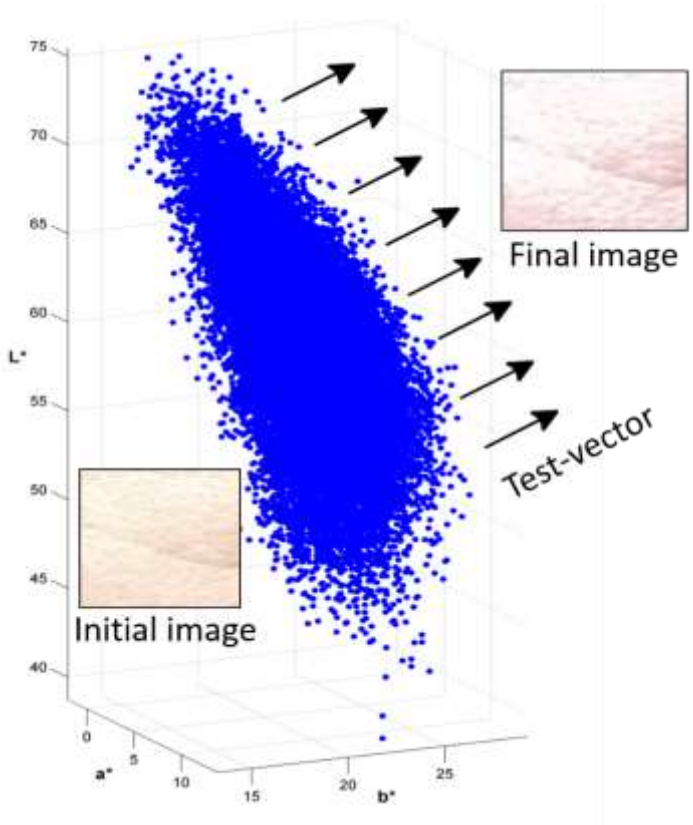


Figure 4.3: Generation of *test patches*. The blue points represent the *reference image*. The black arrows show the direction of the *test vector*. The result is a displacement of the colour of each pixel in the direction of the *test vector*.

The thresholds were estimated along 14 directions such that the space was sampled evenly. Six of these coincided with the cardinal $\pm L^*$, $\pm a^*$ and $\pm b^*$ directions while the other eight directions were along the centres of the eight octants. During the experiment, the length of the *test vector* in each direction was controlled by the QUEST adaptive algorithm (A. Watson & Pelli, 1983), leading to 14 interleaved staircases. The parameters of the QUEST algorithm were adjusted for the 4-AFC paradigm to optimise the sweat-factor (Robbins & Monro, 1951; Taylor, 1971), theoretically giving the measured threshold an 86% score on the psychometric function. In the best case scenario, staircases lasted approximately 40 trials; although some staircases lasted for as many as 90 trials if observers made frequent errors throughout the session. The optimisation of the QUEST parameters is described in the section that follows.

4.2.3 Optimisation of the QUEST procedure

As stated in the previous section, the thresholds were estimated using a variation of the QUEST adaptive procedure (A. Watson & Pelli, 1983). The response thresholds were modelled using the log-Weibull function (Weibull, 1951), the classical form of which is shown below.

$$W = 1 - (1 - \gamma)\exp\left(-10^{(\beta/20)\cdot(x-T+\epsilon)}\right) \qquad \text{Eq. 4.1}$$

Here W is the response modelled using the log-Weibull function, x is the stimulus intensity, γ is the probability of success at a stimulus intensity of zero, β is the slope of the psychometric function, T is the defined threshold point, and ϵ is a parameter introduced so that the efficiency of the testing procedure can be optimised at the threshold T. These parameters are illustrated in Figure 4.4. Since the experimental paradigm involved estimation of a very large number of thresholds in each session (14 interleaved staircases), an optimised ϵ was crucial for keeping the number of trials required per staircase to a controlled minimum.

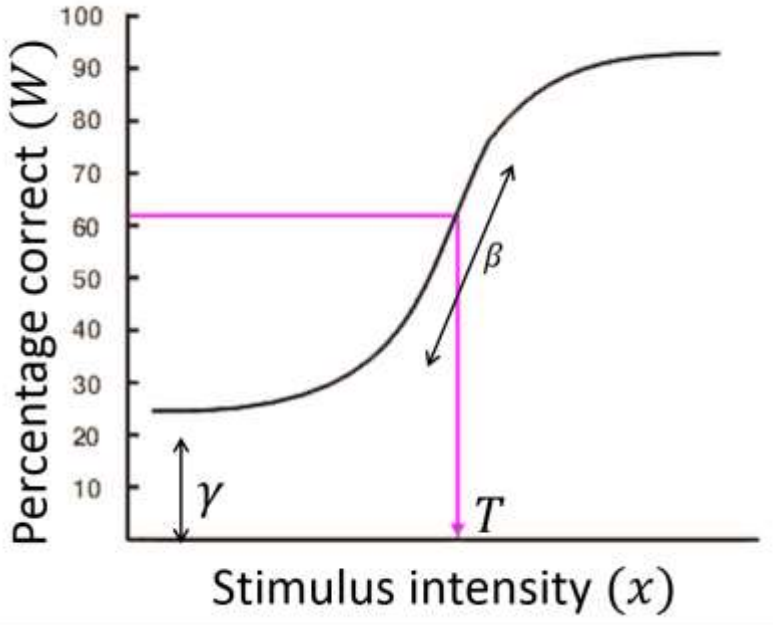


Figure 4.4: Parameters of a typical psychometric curve modelled using the log-Weibull function described in Eq. 4.1.

To characterise the efficiency of psychometric estimation procedures, Taylor & Creelman (1967) proposed a measure they called the S*weat-factor*. Given a parameter θ and a sequence (or procedure) which best-estimates the parameter as $\hat{\theta}$ after n trials, the S*weat-factor* S is defined as

$$S = \sigma_{\hat{\theta}}^2 \qquad \text{Eq. 4.2}$$

where $\sigma_{\hat{\theta}}^2$ simply denotes the variance in the estimated parameter $\hat{\theta}$. Typically, for an online adaptive procedure, it is difficult to have an a priori estimate of this parameter variance. To this end, often an ideal Sweat Factor is used to estimate the efficiency of the adaptive process (Taylor, 1971; Treutwein, 1995). Taylor (1971) defined this ideal S*weat-factor* in terms of Wetherill's (1966) derivation of the asymptotic variance of the Robbins-Monro process (Robbins & Monro, 1951). The Robbins-Monro process is a theoretical sequence of estimations which asymptotically converges to a given threshold, with each term of the sequence depending on the preceding terms, i.e., it is an adaptive estimation algorithm with proven convergence (Robbins & Monro, 1951) at a certain limit or threshold. Using the definition of *Sweat-factor* under the assumption of a Robbins-Monro process converging to a threshold T, the S*weat-factor* S_W for the log-Weibull psychometric function W defined above (Eq. 4.1) can be formulated as

$$S_W = \frac{W(1-W)}{\left[\frac{\partial W}{\partial x}\right]^2} \qquad \text{Eq. 4.3}$$

It must be noted here that the more efficient a procedure, the lower its *Sweat-factor*. This is because an efficient procedure will converge faster, leading to higher values of $\partial W/\partial x$, and in turn, lower values of the *Sweat-factor*. By definition, for a given process, the sweat-factor reaches a minimum at the threshold T, i.e.,

$$\left|\frac{\partial S_W}{\partial x}\right|_{x=T} = 0 \qquad \text{Eq. 4.4}$$

Substituting for S_W in Eq. 4.4 using Eq. 4.3, and solving for ϵ, one can show that the optimum value of the parameter ϵ_{opt} is given by

$$\epsilon_{opt} = \frac{20}{\beta \log(10)} \cdot \log\left(2 + \Lambda\left(0, \frac{2(\gamma-1)}{e^2}\right)\right) \qquad \text{Eq. 4.5}$$

Here, β and γ are parameters of the log-Weibull W (Eq. 4.1), and $\Lambda(N, x)$ denotes the N^{th} branch of Lambert's W-function, which is the set of solutions to the equation $w \cdot \exp(w) = x$. Of course, this solution for ϵ is only valid for $\gamma \in [0,1)$. It can also be numerically verified that for this range of γ, the second derivative of the *Sweat-factor* is always positive, i.e.,

$$\left|\frac{\partial^2 S_W}{\partial x^2}\right|_{x=T,\ \epsilon=\epsilon_{opt}} > 0, \qquad \forall \gamma \in [0,1) \qquad \text{Eq. 4.6}$$

Eq. 4.6 proves that the solution ϵ_{opt} given by Eq. 4.5 does indeed represent a minima of the *Sweat-factor* with respect to the stimulus.

One of the advantages of the Weibull formulation of the psychometric curve is that it allows for independent estimation of the slope β and the threshold T. Since the primary interest in these experiments was the magnitude of the thresholds rather than the psychometric slope, a slope of $\beta = 3.5$ was used throughout the study. To make the convergence more rapid, a 4-AFC task was used, corresponding to $\gamma = 0.25$. Using these values of β and γ in Eq. 4.5 one obtains an optimised value of $\epsilon_{opt} = 1.37$ which, theoretically, maximises the *Sweat-factor* of the estimation process. It is to be noted that the optimisation procedure proposed here does not allow for an explicit choice of the threshold. Instead, the threshold is adjusted for the most efficient performance of the adaptive estimation process, given the design of the study. For the current study, the optimised threshold roughly corresponded to a value of $W = 0.87$,

i.e., an 87% score on the psychometric curve.

4.2.4 Stimulus presentation

The stimuli were presented on a colour calibrated EIZO ColorEdge CG243W monitor using the CRS ViSaGe system. The participants were seated 175 $cm.$ away from the screen. At this distance, the opposing edges of the individual 5 $cm.\times$ 5 $cm.$ patches subtended an angle of $\approx 1.65°$ at the observer's retina (while the diagonals subtended and angle $\approx 2.3°$).

When the stimuli were displayed under the illumination from the luminaire, the screen was covered by a grey cardboard sheet with cut-outs such that only the four patches remained visible. The reasons for doing this were twofold. First, doing so occluded the self-luminous background, forcing the observer to further adapt to the ambient illumination. Second, it made the patches appear less like images presented on a self-luminous screen, akin to what is often described as an *object-mode* of stimulus presentation (Tangkijviwat, Rattanakasamsuk, & Shinoda, 2010). This is a more ecologically valid method of presenting stimuli such as natural or known textures and surfaces on a computer screen. In the dark condition, the stimuli were displayed against a grey background of the same chromaticity as the simulated daylight from the luminaire $(x = 0.32, y = 0.34)$ at 20 cd/m^2. This was done in order to avoid adaptation to an arbitrary unpredictable source (the self-luminous stimuli) in the Dark condition. It also ensured that the condition would be somewhat comparable to one of the luminaire-illuminated conditions (D65).

4.2.5 Response collection and analysis software

The observer responses were collected using a Cedrus RB-350 mechanical-contact response box. The experiment was completely programmed in MATLAB using the CRS (Cambridge Research Systems) Toolbox. The ellipsoid fitting and data analysis

were performed in MATLAB and R.

4.2.6 Experimental protocol

The study began with an initial briefing explaining the general aims of the experiment to the observers. After this the participants were tested for colour normal vision using the Cambridge Colour Test (Regan et al., 1994). Due to the nature of the study, only colour normal participants were allowed to continue.

In order to avoid observer fatigue, the testing was done in blocks held on separate days. The illumination condition within each block was kept constant. This was done because the stabilisation-period of the luminaire was prohibitively high to allow switching between illumination conditions within the same block. Before each block, the corresponding light source was allowed to stabilise for about half an hour.

Each block was divided into several sessions, each session corresponding to the measurement of a separate discrimination boundary. It must be noted that a discrimination boundary is described (in this study) by measurements along 14 directions. Thus, in other words, each session consisted of 14 randomly interleaved staircases operating along different directions. The sessions began with a test-run of 30-50 presentations of stimuli along random directions. In this test-run, the stimuli were deliberately chosen such that the observers found the task easy, with $\Delta E_{LAB} \geq 5$. The objective was to adapt the observers to the illumination condition and make sure they understood the task. After the test-run, the observers remained in the light-proof chamber for another minute before a long-beep signalled the start of the main experiment. At this point, the observer pressed a button to start the presentation of the trials. Each trial timed out after 5 seconds, irrespective of whether the observer responded or not. A button press or time-out was signalled by a beep, after which the experiment moved on to the next trial. If the observer failed to respond in a certain trial the experiment simply advanced to the next randomly chosen staircase while the state of the original staircase (for which the observer did not register a response) was

not changed, i.e., a failure to respond was not penalised.

The thresholds for each stimulus were estimated thrice. There were breaks of 5-10 minutes between the sessions, during which the observers were allowed to exit the light-proof chamber.

4.2.7 Ellipsoid fitting

Despite its intuitive interpretation for the purposes of colour differences and local colour directions, $CIELAB$ is, in effect, a relative colour space. More precisely, $CIELAB$ uses a white-point adaptation to normalise colour coordinates to a given reference. In display calibration applications, this reference point is often taken as the white-point of the screen. In most other applications where accurate transfer of absolute tristimulus information is critical (such colour management chains), a standard illuminant such as $D50$ is commonly employed. In this study, a natural choice for the reference point would be the illuminant colour for the luminaire-illuminated conditions where the stimulus is presented through a cardboard cut-out in a booth with ambient illumination, and the background chromaticity for the Dark condition where the stimulus is displayed with a luminous background and the testing-booth is dark.

Thus, any comparison of thresholds in $CIELAB$ coordinates defined using different reference points would implicitly be a test of the adaptation model implemented in $CIELAB$. This makes comparison of raw thresholds across illumination conditions in this study quite difficult. To address this issue, the analysis was performed in a 3-D space defined by the $CIE\ 1976\ UCS$ chromaticities and a scaled luminance. The $CIE\ 1976\ UCS$ chromaticities are projective mappings of the $CIE\ 1931\ XYZ$ tristimulus coordinates without any assumptions about adaptation, and thus allow a direct comparison of absolute thresholds between illumination conditions. The exact formulation of the projection is described in Appendix 1.2. The luminance axis in our analysis was scaled by 1/100 so that it followed the order of magnitude of the chromaticity values, akin to methods previously employed by Melgosa et al. (1997) in

reporting supra-threshold ellipsoids for surface colours. This colour space will be referred to as the $u'v'Y'$ space from here on.

First, average threshold (over the 3 repetitions) along each of the 14 directions was calculated for each observer in each condition. Colour discrimination thresholds can be modelled by various methods and surfaces (Poirson et al., 1990), the most common of which are ellipses and ellipsoids determined by differential or line-element methods. These models have long been used in vision, and in particular, colour perception. Poirson & Wandell (1990b) showed the reliability of an ellipsoidal fit in interpolation of thresholds from measurements. Similar line element theory inspired ellipsoids were fitted to the estimated thresholds, both for colours and for textures. An ellipsoid was fitted to each set of 14 data-points by minimising the least-squared distance of the points from the ellipsoid surface.

During the calculation of each ellipsoid, the mean colour of the corresponding stimulus (calculated in $CIELAB$ space and then converted to the illuminant-independent $u'v'Y'$ space) was taken as the centre of the ellipsoid. An ellipsoid with a known centre is completely defined by six independent parameters - the lengths of the three axes, and the three Brian-Tait angles. The Brian-Tait angles of an ellipsoid represent the sequence of orthogonal rotations about the cardinal axes required to achieve its axes-orientations. Although the Brian-Tait angles encode the axes-orientations of the ellipsoid, they are not readily interpretable in terms of physical parameters of the ellipsoid. Therefore, after optimisation, the six optimised parameters were converted to more intuitive quantities for analysis, such as the volume, the axes lengths and orientation vectors, and the area and orientation of the projected ellipse on the chromaticity plane. These parameters are further illustrated in Figures 4.5 and 4.6, and a description of the ellipsoid fitting process is given below.

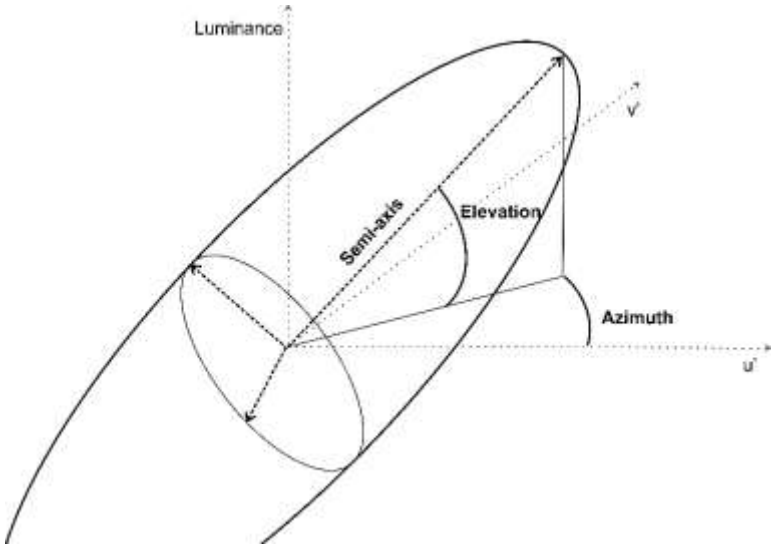


Figure 4.5: Ellipsoid parameters.

Let a discrimination ellipsoid in a real 3-dimensional $\mathbb{R}^3$ space ($u'v'Y'$ space in our case) be denoted by E. The threshold boundary of this ellipsoid E can be written as

$$E \equiv \{\boldsymbol{x} \mid \|\boldsymbol{\Sigma U^T}(\boldsymbol{x}-\boldsymbol{c})\| = 1\} \qquad \text{Eq. 4.7}$$

where $\boldsymbol{x} \in \mathbb{R}^3$ is a colour on the threshold boundary, $\boldsymbol{\Sigma}$ is a diagonal matrix $[\sigma_{ii}]_{i=1}^{3}$ with its entries representing the squared inverse lengths of the semi-axes, $\boldsymbol{U}$ is an orthogonal matrix with each column representing a unit vector along one of the ellipsoid axes, and $\boldsymbol{c}$ is the centre of the ellipsoid (defined as the average colour of the tested stimulus in this study). Of course, the above equation is superfluous in terms of the number of variables as $\boldsymbol{U}$ is an orthogonal matrix, i.e., the three columns of $\boldsymbol{U}$ are orthogonal. This fact is used to reduce the number of unknown variables in the equation, thereby making the optimisation process much faster. One beings by recognising the fact that $\boldsymbol{U^T}$ is essentially a rotation matrix in 3 dimensions, and thus can be decomposed into a set of Brian-Tait angles or rotation angles. The Brian-Tait angles of an ellipsoid essentially describe the sequence of rotations one must perform on the cardinal axes of the colour space such that they are aligned with the ellipsoid axes. Although the mapping from $\boldsymbol{U^T}$ to Brian-Tait angles is not bijective (it is many-one), a branch of the solution suffices to cover all real 3-D cases computationally. If

$\boldsymbol{R} = \boldsymbol{U^T} = [r_{ij}]_{i,j=1}^{3}$ is the rotation matrix, $\boldsymbol{\Theta} = \left[\theta_x\ \theta_y\ \theta_z\right]^T$ are Brian-Tait angles for the three axes, and $\mathrm{atan}(x, y)$ represents the 4-quandrant arctangent (which takes two arguments, and has a range different from that of the standard arctangent which is denoted here as $\tan^{-1}x$), a sufficient mapping for $\boldsymbol{\Theta}$ is given by

$$\theta_y = -\,\mathrm{asin}(r_{31})$$

$$\theta_x = \begin{cases} \mathrm{atan}(r_{12}, r_{13})\ , & r_{31} \in \{-1,1\} \\ \mathrm{atan}\left(\dfrac{r_{32}}{\cos\theta_y}, \dfrac{r_{33}}{\cos\theta_y}\right), & \text{otherwise} \end{cases} \qquad \text{Eq. 4.8}$$

$$\theta_z = \begin{cases} 0\ , & r_{31} \in \{-1,1\} \\ \mathrm{atan}\left(\dfrac{r_{21}}{\cos\theta_y}, \dfrac{r_{11}}{\cos\theta_y}\right), & \text{otherwise} \end{cases}$$

In the context of the ellipsoids in the given experiment, since c is fixed as the mean $CIELAB$ colour of the stimulus, the optimisation was performed for only for $\boldsymbol{\Sigma}$ and $\boldsymbol{\Theta}$, i.e., six parameters. Suppose the estimated thresholds for a given condition and a given observer are given by the set $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_{14}}\}$, such that each of $\boldsymbol{x_1}$ through $\boldsymbol{x_{14}}$ represents the threshold along one of the 14 measured directions. Furthermore, let $E(\boldsymbol{\Sigma}, \boldsymbol{\Theta})$ be an ellipsoid defined by the parameters $\{\boldsymbol{\Sigma}, \boldsymbol{\Theta}\}$. Let us also define a metric $d(\boldsymbol{x}, E)$ which denotes the Euclidean distance between a point $\boldsymbol{x}$, and an ellipsoid E. The optimisation problem to be solved for fitting an ellipsoid with optimised parameters $\{\boldsymbol{\Sigma_{opt}}, \boldsymbol{\Theta_{opt}}\}$ can now be written down as

$$\{\boldsymbol{\Sigma_{opt}}, \boldsymbol{\Theta_{opt}}\} = \min_{\boldsymbol{\Sigma},\boldsymbol{\Theta}} \sum_{i=1}^{14} d(\boldsymbol{x_i}, E(\boldsymbol{\Sigma}, \boldsymbol{\Theta})) \qquad \text{Eq. 4.9}$$

As can be seen from Eq. 4.8, $\boldsymbol{\Theta}$ is not a very intuitive parameter for a discrimination ellipsoid. For this reason, after optimisation, it was converted back to the more easily interpretable orthogonal matrix $\boldsymbol{U}$ of unit vectors along the axes of the ellipsoid. From this, the elevations (θ_i) and azimuths (φ_i) for each column $\boldsymbol{u_i} = [u_{1i}\ u_{2i}\ u_{3i}]^T$ of $\boldsymbol{U}$ were derived simply by converting them to polar coordinates using

$$\theta_i = \tan^{-1}\frac{u_{i3}}{\sqrt{u_{i1}^2 + u_{i2}^2}}, \qquad \theta_i \in \{-\frac{\pi}{2}, \frac{\pi}{2}\}$$ Eq. 4.10

and

$$\varphi_i = \operatorname{atan}(u_{i1}, u_{i2}), \qquad \varphi_i \in \{-\pi, \pi\}$$ Eq. 4.11

The volume V of a given ellipsoid with scaling matrix $\mathbf{\Sigma}$ was calculated by using

$$V = \frac{4}{3}\pi|\mathbf{\Sigma}|^{-1/2}$$ Eq. 4.12

Furthermore, assume that there is a plane L defined by

$$L \equiv \{\boldsymbol{x} \mid \boldsymbol{x} = \boldsymbol{d} + \boldsymbol{T}\boldsymbol{t}\}$$ Eq. 4.13

such that $\boldsymbol{d}$ is a point on the plane, $\boldsymbol{T}$ is the set of basis vectors in $\mathbb{R}^3$ defining the plane, and $\boldsymbol{t}$ is the vector of weights for each of the basis vectors, i.e. the local coordinates of a given point in the plane L. With $\boldsymbol{d} = \mathbf{0}$ and $\boldsymbol{T} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, one can model this plane L to represent the zero luminance chromaticity plane in the $u'v'Y$ colour space. The parallel-projection P_E of an ellipsoid E (as defined in Eq. 4.7 above) on this plane L can be calculated to be

$$P_E \equiv \{\tilde{\boldsymbol{x}} \mid \left\|\tilde{\mathbf{\Sigma}}\tilde{\boldsymbol{U}}(\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{c}})\right\| = 1\}$$ Eq. 4.14

where $\tilde{\boldsymbol{x}}$ is a point on the projected ellipse P_E, $\tilde{\boldsymbol{c}} = \boldsymbol{T}^{\boldsymbol{T}}(\boldsymbol{c} - \boldsymbol{d})$ is the centre of the ellipse, and $\tilde{\boldsymbol{U}}$ and $\tilde{\mathbf{\Sigma}}$ are the left singular-vector matrix and the inverse of the truncated singular-value matrix respectively, found by the singular value decomposition of $\boldsymbol{T}^{\boldsymbol{T}}\boldsymbol{U}\mathbf{\Sigma}^{-1}$. The matrices $\tilde{\boldsymbol{U}}$ and $\tilde{\mathbf{\Sigma}}$ have interpretations in a 2-D chromaticity plane

analogous to those of $\boldsymbol{U}$ and $\boldsymbol{\Sigma}$ in a 3-D space, i.e., $\widetilde{\boldsymbol{U}}$ is a matrix such that each of its columns is one of the axes of the projected ellipse P_E, and $\widetilde{\boldsymbol{\Sigma}}$ is a diagonal matrix with each entry denoting the squared inverse length of the corresponding semi-axis. The area of this projected ellipse is given by

$$A = \pi\left|\widetilde{\boldsymbol{\Sigma}}\right|^{-1/2} \qquad \text{Eq. 4.15}$$

The azimuth $\tilde{\varphi}_1$ of the major axis of the projected ellipse (often referred to as the orientation of the ellipse) is calculated simply by using the first column of the $\widetilde{\boldsymbol{U}}$ matrix $\widetilde{\boldsymbol{u}}_{\boldsymbol{1}} = [\tilde{u}_{11}\ \tilde{u}_{12}]^T$ which denotes a unit vector along the major axis of P_E.

$$\tilde{\varphi}_1 = \tan^{-1}(\tilde{u}_{11}, \tilde{u}_{12}), \qquad \tilde{\varphi}_1 \in \{-\frac{\pi}{2}, \frac{\pi}{2}\} \qquad \text{Eq. 4.16}$$

The volume (Eq. 4.12) of the ellipsoids can be taken as a measure of the actual number of non-discriminable stimuli. Further analysis was divided into analysis of the projections of these ellipsoids on the chromatic plane (theoretically, the envelope of chromatic discrimination ellipses across luminances), and analysis of the luminance projections of the ellipsoids (illustrated in Figure 4.6). This analysis of the discrimination ellipsoid in terms of luminance and chromaticity projections was driven, in part by the difficulty of demonstrating 3-D plots and their parameters, and in part, by the independence in the chromaticity and luminance projections of discrimination ellipsoids found by other researchers (Melgosa, Pérez, El Moraghi, & Hita, 1999), and later verified by results from the present study (Figures 4.11 and 4.12). It also breaks down the 3-D discrimination boundaries into two components which are relatively easier to interpret.

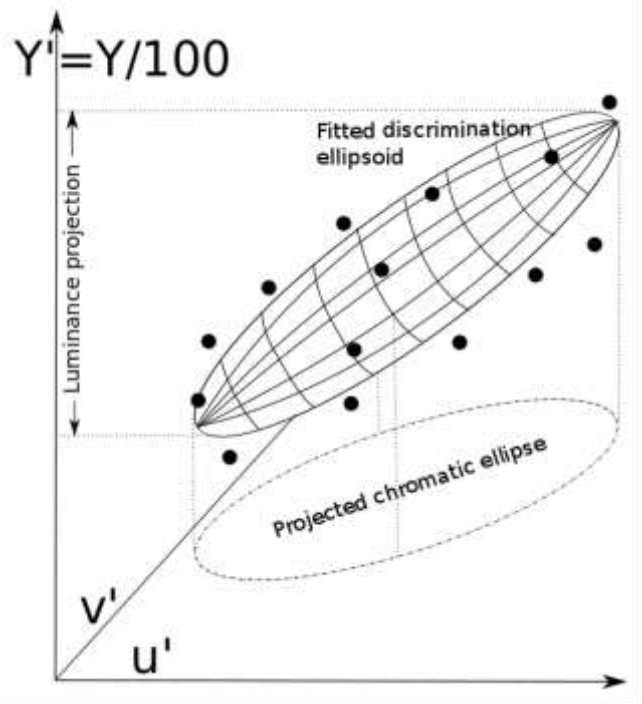


Figure 4.6: Ellipsoids were fitted to 14 thresholds in each condition for each observer. These ellipsoids were projected on the chromaticity plane and the luminance axis for further analysis. Please note that this is only an illustration to demonstrate the projections; the actual ellipsoids showed a much closer alignment with the vertical luminance axis (see Figures 4.9 and 4.17).

Both the area of the chromatic projections (Eq. 4.15), as well as the azimuth of their major-axes (Eq. 4.16) were analysed. This gives us an insight as to which colour directions generate the maximum confusion for a given stimulus. This analysis of the discrimination ellipsoid in terms of luminance and chromaticity projections was driven, in part by the difficulty of demonstrating 3-D plots and their parameters, and in part, by the independence in the chromaticity and luminance projections of discrimination ellipsoids found by other researchers (Melgosa et al., 1999), and later verified by results from the present study (see Figures 4.11 and 4.12).

4.2.8 Analysis

The analysis was carried out in R and MATLAB. Circular variables were analysed using directional statistics (Fisher, 1953). This was implemented using the 'circular' package in R and the 'CircStat' toolbox (Berens, 2009) in MATLAB. Both these

packages use routines based on Jammalamadaka & Sengupta (2001). Repeated-measures ANOVAs were used to conduct inferential statistics on all parameters except directional quantities – for these, the Harrison-Kanji test (Harrison & Kanji, 1988) was used. Post-hoc investigation of interactions was done by testing simple effects at the relevant levels using Bonferroni corrected pairwise t-tests. The circular analogue of this was the Watson-Williams test (Stephens, 1969; G. Watson & Williams, 1956).

It must be noted that satisfactory circular analogues of repeated-measures statistical tests (such as t-tests and ANOVAs) have so far not been proposed in the relevant literature (McMillan, Hanson, Saunders, & Gallun, 2013). Hence, the directional tests conducted here do not take repeated measures into account.

4.3 Experiment 1a: Skin images

The first experiment was aimed at investigating the effects of illumination condition and skin-type on the discrimination thresholds for skin images. To this end, the participants were tested in three different lighting conditions (dark, daylight D65 and cool-white-fluorescent TL84) using calibrated images of two skin-types (Caucasian and Oriental).

4.3.1 Image acquisition and spectral reconstruction

Images of Caucasian and Asian female faces were captured under controlled D65 lighting in a Verivide® DigiEye light-booth using a calibrated Nikon D7000 camera. The images were controlled for size by including a marker of known dimensions in the frame. This was important for reproducing the actual size of the skin patches during the experiment. Approximately $5\ cm. \times 5\ cm.$ patches were cropped from the forehead regions of two selected images. The cropping was done such that the patches looked uniformly lit, planar and textured. Care was taken to minimise cues besides colour and texture – such as obvious illumination gradients, shadows, furrows, wrinkles,

blemishes, and facial and stray hair. Since these patches were obtained under controlled lighting with a known illuminant and a booth with grey non-reflective walls, they were taken as stock images for the reconstruction of the spectra at each pixel. Since the illuminants in the experiment-booth were different from the Verivide® DigiEye light-booth used to register the images, access to the reconstructed spectral image (i.e., the reflectance spectrum of each pixel) instead of a tristimulus image was preferred as it allowed for calculation of the colour of the skin patches under any given illuminant.

The workings of the reflectance reconstruction algorithm are inspired by PCA based methods such as those proposed by (Agahian, Amirshahi, & Amirshahi, 2008; Babaei, Amirshahi, & Agahian, 2011; Shen, Cai, Shao, & Xin, 2007). An essence of the spectral reconstruction model is provided in what follows.

The model was fit using measurements of a Silicon skin-colour chart manufactured by Spectromatch® as the training data. The reflectance spectra of this chart closely resemble those of measured human skin. This skin specific calibration was performed in order to gain better accuracy as compared to broader gamut calibration charts such as the MacBeth® chart calibration. Let the reflectance and 3-channel RGB measures of the skin-samples in the chart be $\boldsymbol{R_T}$ and $\boldsymbol{X_T}$ respectively, with the subscript $\boldsymbol{T}$ denoting the training dataset. Both $\boldsymbol{R_T}$ and $\boldsymbol{X_T}$ are matrices, with measurements for individual skin-chart samples placed column-wise. The optimisation can be made more robust by considering a Singular Value Decomposition (SVD) of $\boldsymbol{R_T}$ and only using the first few bases which explain most of the variance, thereby reducing noise. Thus, we write $\boldsymbol{R_T}$ as

$$\boldsymbol{R_T} - \overline{\boldsymbol{R}}_{\boldsymbol{T}} = \widetilde{\boldsymbol{B}}_{\boldsymbol{T}}^{\boldsymbol{k}} \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{T}}^{\boldsymbol{k}} \qquad \text{Eq. 4.17}$$

where $\overline{\boldsymbol{R}}_{\boldsymbol{T}}$ stands for the mean reflectance of the training set and $\widetilde{\boldsymbol{B}}_{\boldsymbol{T}}^{\boldsymbol{k}}$ stands for the first k bases of the training set, with $\tilde{\boldsymbol{L}}_{\boldsymbol{T}}^{\boldsymbol{k}}$ denoting the corresponding loadings.

The main idea here is to optimise a regression matrix $\boldsymbol{\Lambda}$ which would help transform a function f of any 3-channel skin colour measurement vector $\boldsymbol{x}$ to a reflectance spectrum, say, $\boldsymbol{r}$. Let us denote the output of f as $\boldsymbol{y}$, i.e., $\boldsymbol{y} = f(\boldsymbol{x})$. We now propose a regression model of the form

$$\boldsymbol{R_T} - \bar{\boldsymbol{R}}_{\boldsymbol{T}} = \boldsymbol{\Lambda} \cdot f(\boldsymbol{X_T}) = \boldsymbol{\Lambda} \cdot \boldsymbol{Y_T} \qquad \text{Eq. 4.18}$$

This model allows the determination of $\boldsymbol{\Lambda}$ by a simple pseudo-inverse calculation.

$$\boldsymbol{\Lambda} = (\boldsymbol{R_T} - \bar{\boldsymbol{R}}_{\boldsymbol{T}}) \cdot \boldsymbol{Y_T^T} \cdot \left(\boldsymbol{Y_T} \cdot \boldsymbol{Y_T^T}\right)^{-1} \qquad \text{Eq. 4.19}$$

Combining Eq. 4.17 and Eq. 4.19 we now get a noise-reduced regression matrix

$$\boldsymbol{\Lambda} = \tilde{\boldsymbol{B}}_{\boldsymbol{T}}^{\boldsymbol{k}} \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{T}}^{\boldsymbol{k}} \cdot \boldsymbol{Y_T^T} \cdot \left(\boldsymbol{Y_T} \cdot \boldsymbol{Y_T^T}\right)^{-1} \qquad \text{Eq. 4.20}$$

Let the $\boldsymbol{I_{RGB}}$ represent the camera three channel image of a skin patch recorded in the VeriVide booth such that the RGB values at a given location (x, y) are given by

$$\boldsymbol{I_{RGB}}(x, y) = \boldsymbol{x_{pixel}} = \begin{bmatrix} R_{pixel} \\ G_{pixel} \\ B_{pixel} \end{bmatrix} \qquad \text{Eq. 4.21}$$

Furthermore, let the estimated reflectance at (x, y) be denoted by $\hat{\boldsymbol{r}}_{\boldsymbol{pixel}}$, with the hat over the $\boldsymbol{r}$ denoting that it is an estimate. Using Eq. 4.18 and Eq. 4.20, the spectral reconstruction now reduces to the regression

$$\hat{\boldsymbol{r}}_{\boldsymbol{pixel}} = \bar{\boldsymbol{R}}_{\boldsymbol{T}} + \tilde{\boldsymbol{B}}_{\boldsymbol{T}}^{\boldsymbol{k}} \cdot \tilde{\boldsymbol{L}}_{\boldsymbol{T}}^{\boldsymbol{k}} \cdot \boldsymbol{Y_T^T} \cdot \left(\boldsymbol{Y_T} \cdot \boldsymbol{Y_T^T}\right)^{-1} \cdot f(\boldsymbol{x_{pixel}}) \qquad \text{Eq. 4.22}$$

Two critical aspects of the algorithm are the number of bases k used in the SVD, and the transformation function f. In other collaborative experiments outside the scope of the current thesis (Xiao, Qin, Chauhan, Li, & Wuerger, 2014), the optimal value for k

was found to be 5, while the function f providing the best regression results was found to be the 4th order polynomial. In other words, the first 5 bases were used to extract the reflectance information from the training set, while the three-channel RGB values were converted to 4th order polynomial forms for the regression.

The spectral reconstruction described above was applied to each pixel of the VeriVide booth images, resulting in corresponding spectral images where the reflectance at each pixel was known. These spectral images allowed for the simulation of the colour of each pixel under any given illuminant using the simple illuminant-reflectance-sensor equation (Eq. 1.1) for each pixel.

$$\boldsymbol{X}_i = \int_{visible} \boldsymbol{L}(\lambda) \cdot \hat{\boldsymbol{r}}_{\boldsymbol{pixel}}(\lambda) \cdot \bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda) \cdot d\lambda\,, \qquad i \in \{1,2,3\} \qquad \text{Eq. 4.23}$$

Here, λ is the wavelength in the visible spectrum, $\boldsymbol{L}(\lambda)$ is the spectrum of the illuminant, $\hat{\boldsymbol{r}}_{\boldsymbol{pixel}}(\lambda)$ is the estimated reflectance spectrum calculated for each pixel (Eq. 4.22), $\bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda)$ is the i^{th} CIE 1931 XYZ colour matching function and $\boldsymbol{X}_i$ is the i^{th} tristimulus coordinate corresponding to $\bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda)$.

In this experiment, values of $\boldsymbol{L}(\lambda)$ corresponding to the curves shown in Figure 4.1 were used for simulating the patches. The colour gamuts of the patches simulated using both overhead luminaires are shown below (Figure 4.7).

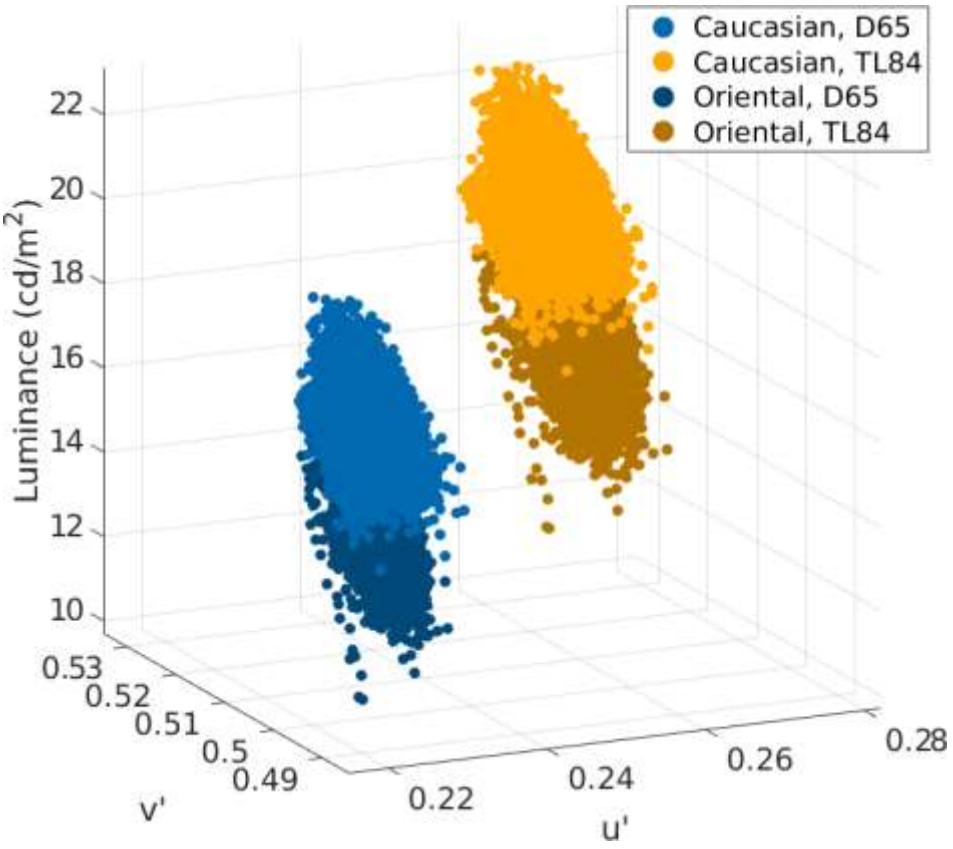


Figure 4.7: Colour distributions of the stimuli. Caucasian and Oriental skin patches were simulated (Eq. 4.23) using D65 and TL84 illuminant SPDs (Figure 4.1), and pixel-wise reconstructed reflectance spectra (Eq. 4.22). Displayed stimuli were always consistent with the ambient illumination in the test-booth.

The luminance and chromatic projections of these distributions are shown in Figure 4.8. The first row shows plots of luminance (ordinate) against the u' coordinate (abscissa), while the second row shows $u'v'$ chromaticity plots (v' ordinate, u' abscissa). The Caucasian patch gamuts are plotted in the left column while the Oriental patch is plotted in the right column.

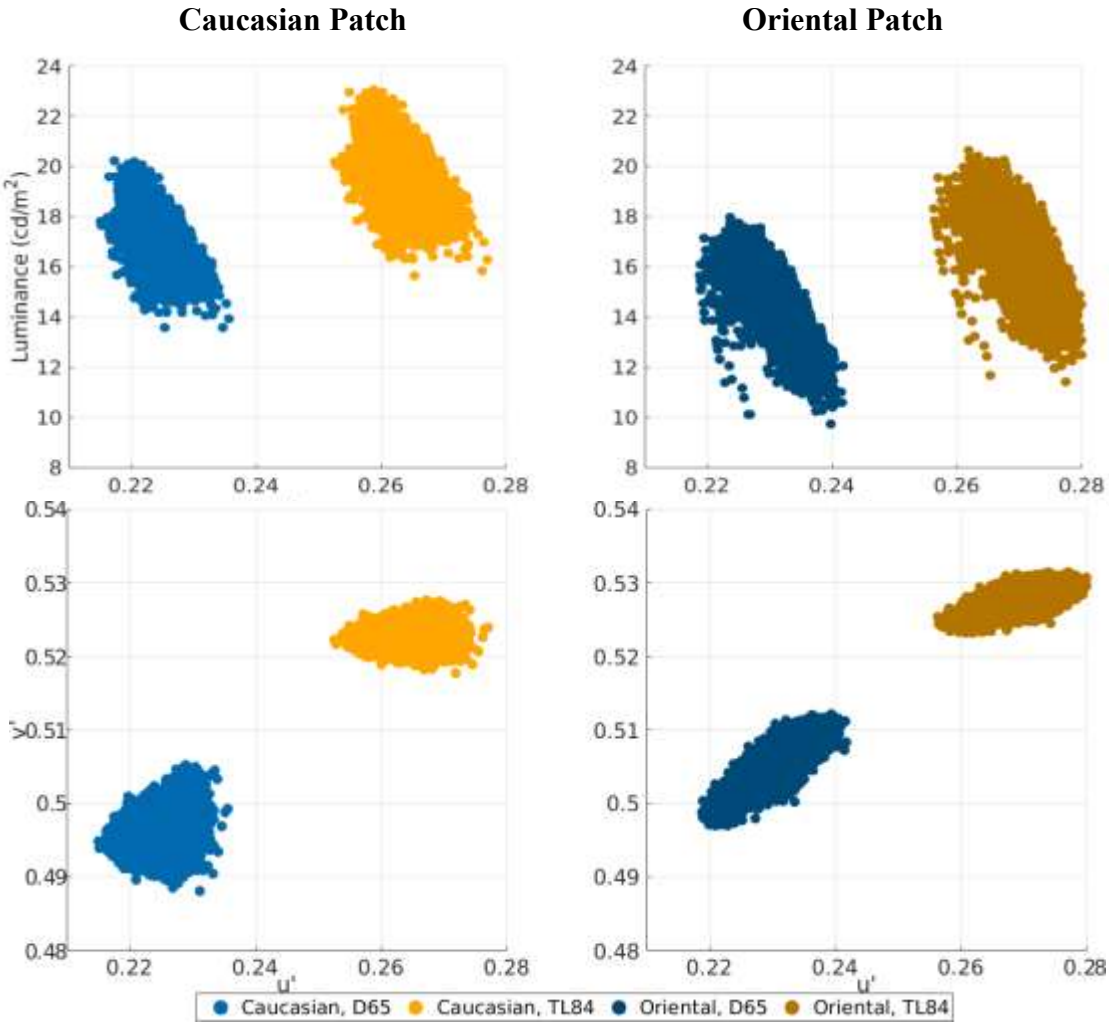


Figure 4.8: Luminance and chromatic spreads of the stimuli simulated using the luminaire illuminants: D65 and TL84. Patches are shown column-wise (left column, Caucasian; right column, Oriental). First row shows the luminance spreads (luminance in cd/m^2 along the ordinate, u' along abscissa) while the second row shows chromatic spreads (v' ordinate, u' abscissa).

The Caucasian patch shows a luminance spread of about $6.5 - 7.5\ cd/m^2$ while the Oriental patch has an $8 - 9\ cd/m^2$ spread. Interestingly, for a given patch, the luminance spreads are very similar (within $1\ cd/m^2$) under the two illuminants. Chromatically, while the two patches show an overlap under each illumination, the Oriental patch has a more elongated distribution of chromaticities. Furthermore, for each patch, the chromatic gamut in D65 seems to be larger than in TL84. A more detailed description of the gamuts for the two patches is shown in Table 4.2 below.

	Volume ($\times 10^{-6}$)	Luminance range (cd/m^2)	Area ($\times 10^{-4}$)	Orientation of first PC (°)
Caucasian Patch				
D65	7.44	6.63	2.42	5.1 (var. exp. = 61%)
TL84	5.82	7.42	1.69	-2.7 (var. exp. = 87%)
Oriental Patch				
D65	7.78	8.24	1.97	34.3 (var. exp. = 89%)
TL84	6.46	9.18	1.40	13.5 (var. exp. = 92%)

Table 4.2: The colour distributions of the two skin patches (Caucasian and Oriental) described using four parameters. *Volume*: Calculated by fitting a convex hull to the distributions in $u'v'Y'$ space. *Luminance range*: Calculated by using maximum and minimum luminance values in the distributions. *Area*: Calculated by fitting a convex hull to the chromatic projections of the data on the $u'v'$ plane. *Orientation of the first principal component*: Calculated by performing a principal component analysis on the chromatic projections and computing the angle made by the first principal component with the positive u' axis.

As noted earlier, the oriental patch shows a higher luminance range in each illumination condition. One also sees that both patches show higher volumes and areas in $D65$ compared to $TL84$. The variance explained by the first principal component was high in all cases ($87\% - 92\ \%$) except the Caucasian patch in the $D65$ condition (61%), and the orientation of the first principal component shows that while the colour distribution of the Caucasian patch varied along the u' axis in both illumination conditions, the Oriental patch showed variation along an inclined axis with the inclination changing markedly with the illuminant.

4.3.2 Experimental protocol

The protocol as described in Section 4.2.6 was used. Within a given illumination block, the two skin patches (Caucasian and Asian) were tested alternately, three times each,

leading to a total of six sub-blocks. On average, the observers responded to approximately 40 trials per staircase; and since there were 14 interleaved staircases, each sub-block consisted of at least 550 trials lasting from 20 to 25 minutes. In total, 252 thresholds (3 illuminants × 2 patch-ethnicities × 3 repetitions × 14 measurement directions) were measured per participant, amounting to about 7.5 hours of testing. The participants were compensated for their time with a fee. In total, 18 participants were tested in this experiment.

4.3.3 Results

The mean ellipsoids derived from the data are shown in Figure 4.9 below. The average measured thresholds are plotted as dots, with the standard error for each threshold represented by a black line along the direction of measurement. The fluorescent illuminant (TL84) condition is plotted in yellow, simulated daylight (D65) in blue, and the Dark condition in grey.

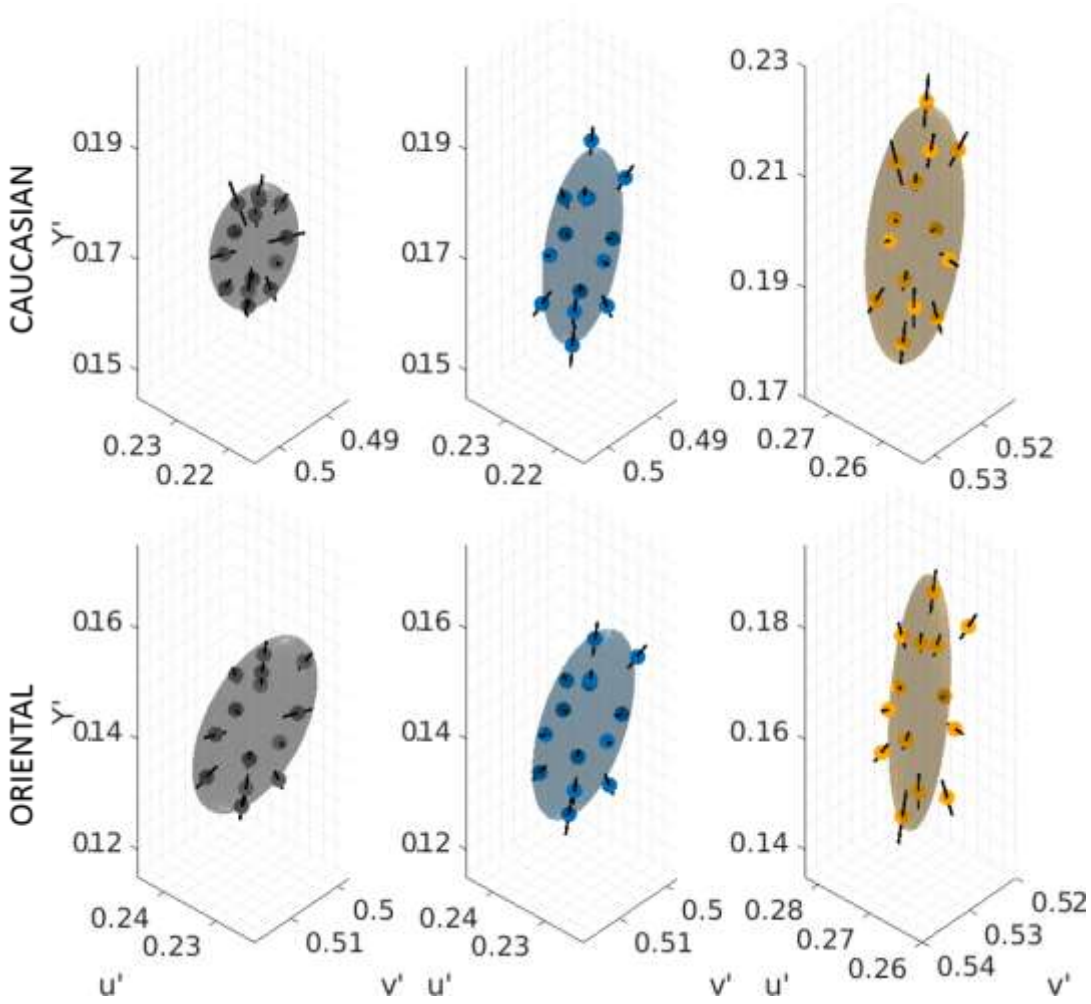


Figure 4.9: Mean ellipsoids for skin stimuli. The ellipsoids correspond to average values of the parameters. The average thresholds across the observers in each direction are marked with a small sphere of the corresponding colour, with the standard error being marked as a black line through the sphere.

The volumes of the ellipsoids across the observers is shown in Figure 4.10 below, with the error bars marking the 95% confidence-intervals (Cousineau, 2005; Morey, 2008). This form of presentation will be used throughout the chapter for other parameters as well.

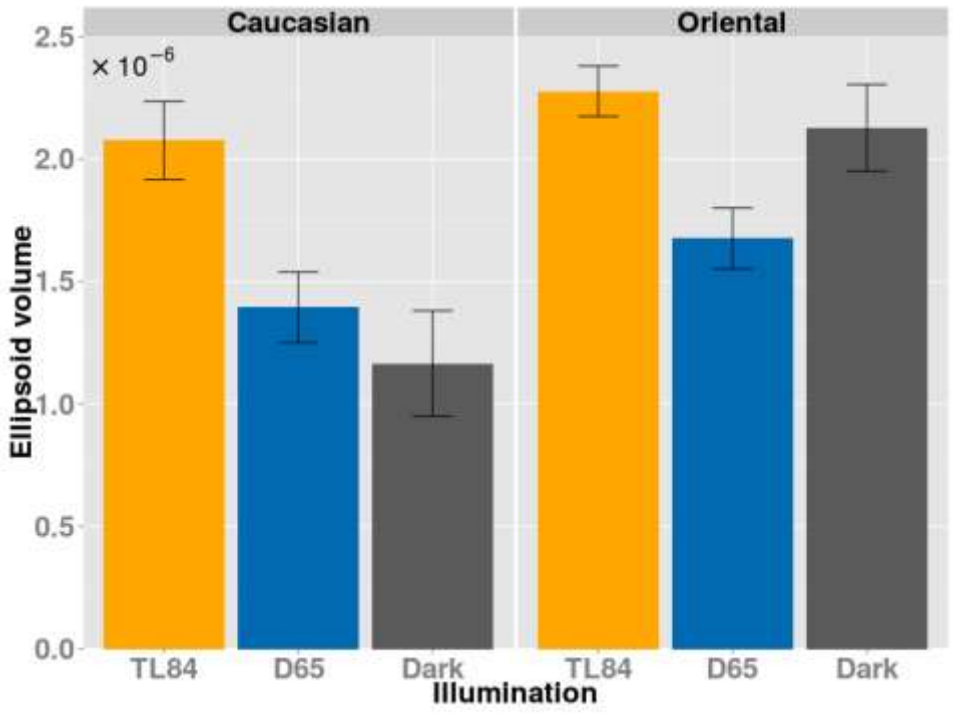


Figure 4.10: Mean ellipsoid volumes for skin images.

Repeated-measures two-way ANOVA reveals main effects of both illumination condition ($F_{2,34} = 4.655, p = 0.0163$) and the patch-ethnicity ($F_{1,17} = 38.78, p < 0.001$), along with an interaction between the two ($F_{2,34} = 11.36, p < 0.001$). Post-hoc tests reveal larger discrimination volumes for the TL84 condition compared to the D65 condition ($p < 0.001$ for both Caucasian and Oriental patches). They also suggest that the main-effect of the patch-ethnicity is found in the Dark ($p < 0.001$) and D65 ($p = 0.013$) conditions, but not in TL84 ($p = 0.21$).

The next two figures (4.11 and 4.12) show plots of the average semi-lengths of the major axes and their projection on the luminance axis respectively.

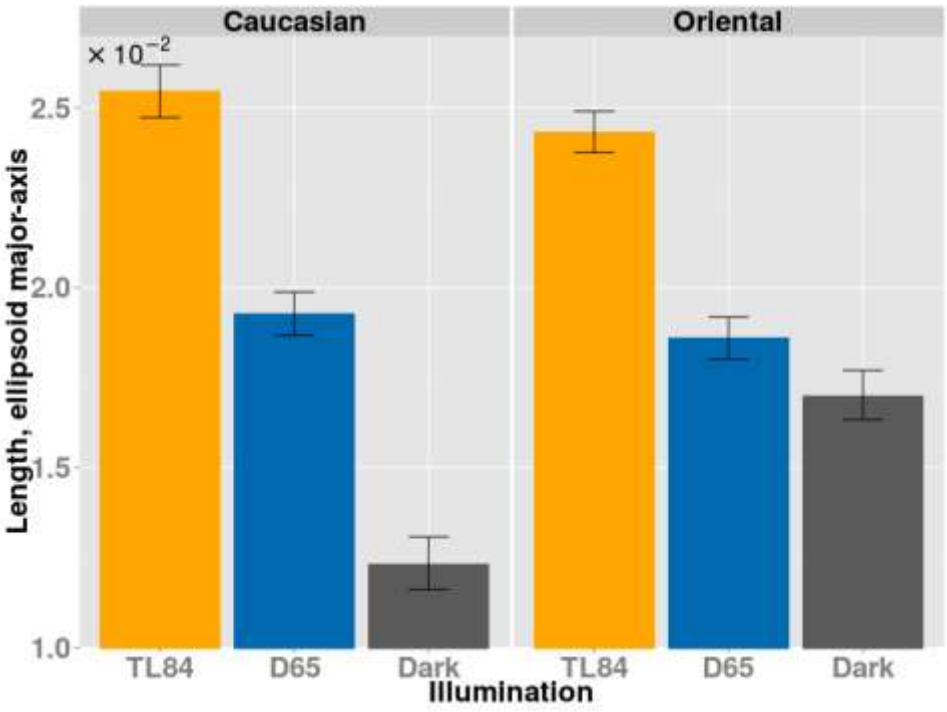


Figure 4.11: Mean major axes for discrimination ellipsoids for skin images.

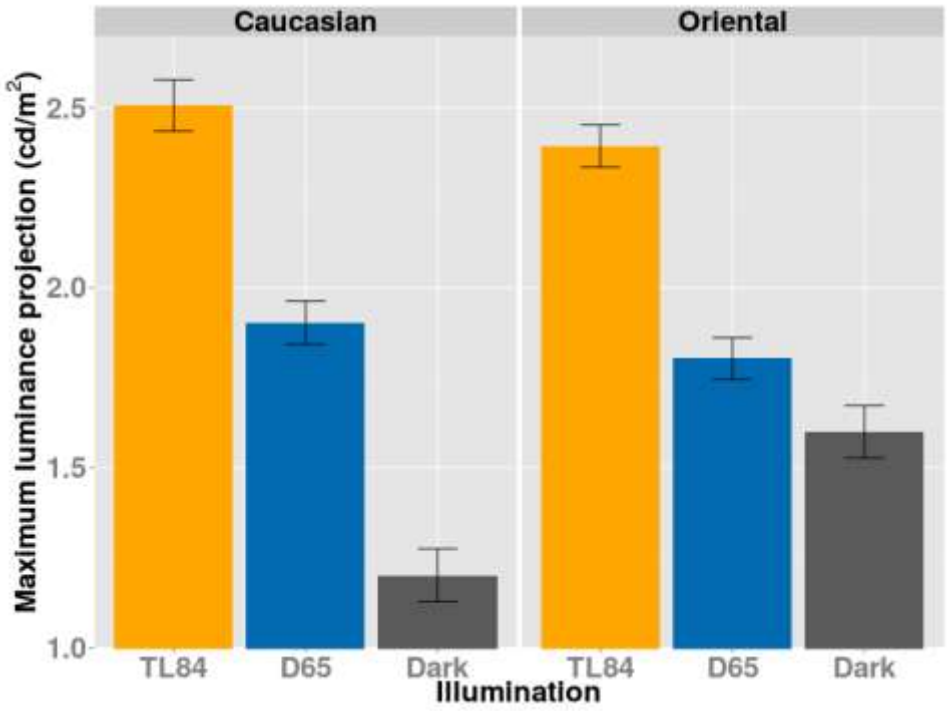


Figure 4.12: Luminance axis projections of the discrimination ellipsoids for skin images.

Both these parameters show almost identical trends, which also come out in the statistical analysis. This is true for all the analyses in all subsequent experiments reported in the paper. Since the luminance axis projection is easier to interpret, all result sections in this chapter will use this parameter instead of the length of the major axis.

Repeated-measures two-way ANOVA show a main effect of illumination condition ($F_{2,34} = 106.7, p < 0.001$), and a strong interaction between the illumination condition and the patch-ethnicity ($F_{2,34} = 11.33, p < 0.001$). No main effect of the patch-ethnicity is observed. Post-hoc tests reveal higher luminance thresholds in TL84 compared to D65 ($p < 0.001$) for both patch types and a higher threshold for TL84 than the Dark condition ($p < 0.001$) for the Caucasian patch.

Figures 4.13 and 4.14 show the mean projections of the ellipsoids on the $u'v'$ chromaticity plane for the two patch-ethnicities.

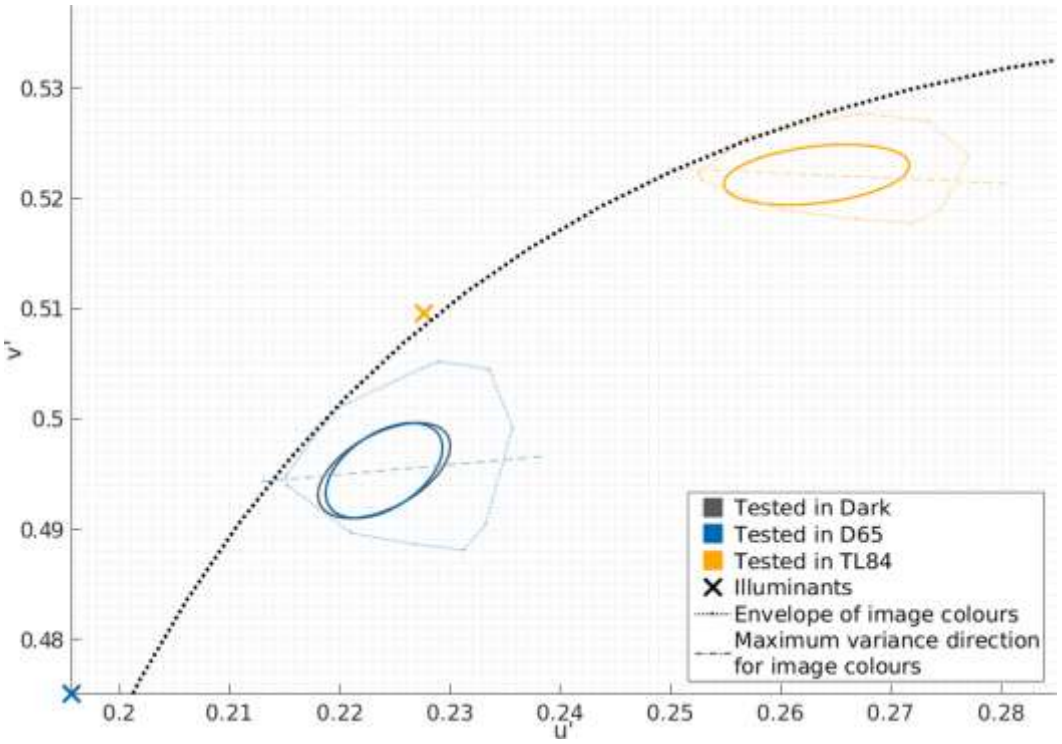


Figure 4.13: Mean projected ellipses for the Caucasian patch.

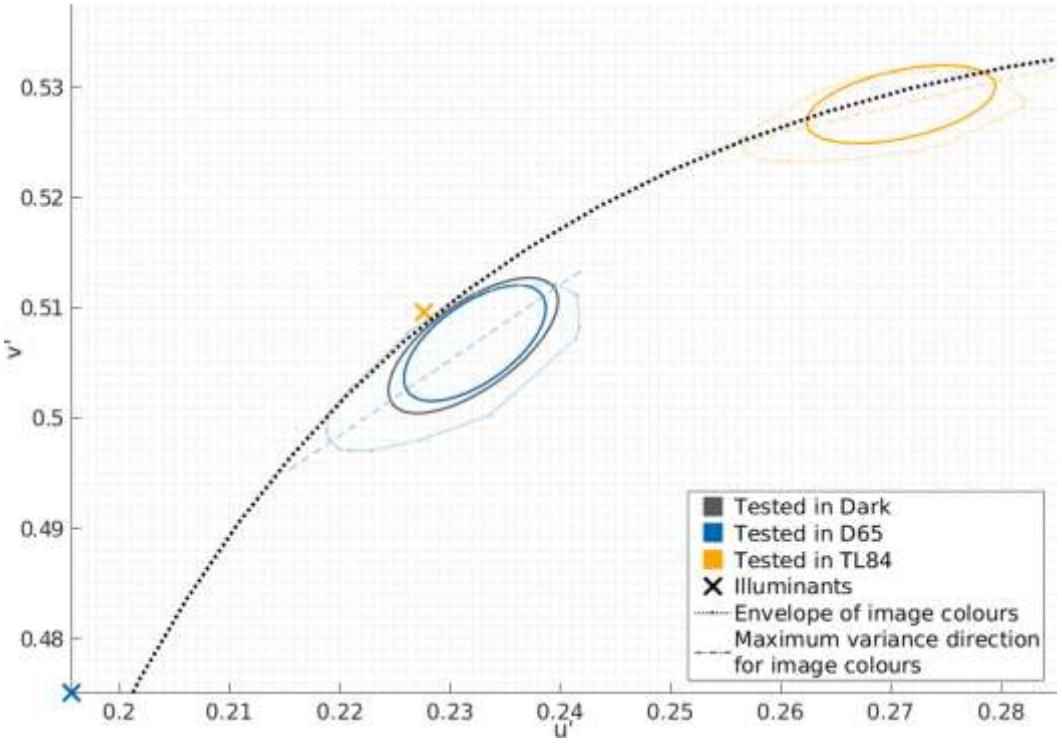


Figure 4.14: Mean projected ellipses for the Oriental patch.

To explore these projections, let us first consider the area of the ellipses (Figure 4.15).

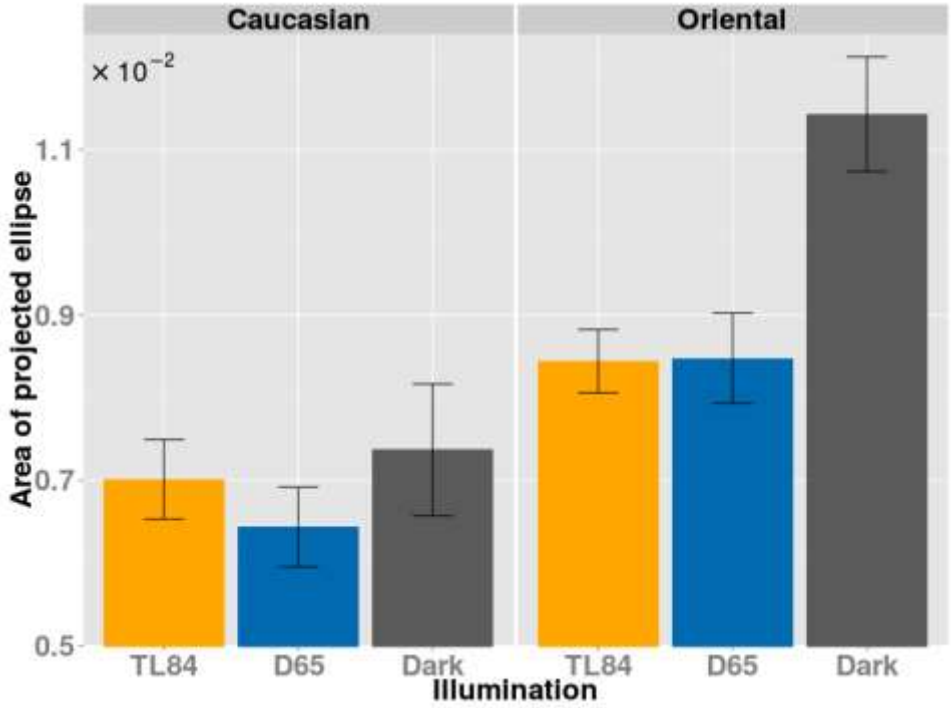


Figure 4.15: Area of the projected chromaticity ellipses for skin images.

A repeated-measures two-way ANOVA reveals a marginal main effect of the illumination condition interaction ($F_{2,34} = 3.211, p = 0.0528$), a significant main effect of the patch-ethnicity ($F_{2,17} = 99.25, p < 0.001$) and a significant interaction ($F_{2,34} = 8.988, p < 0.001$). Bonferroni corrected post-hoc comparisons show that the area of the chromatic projections is higher for the Oriental patch than the Caucasian patch ($p < 0.001$ for Dark and D65, and $p = 0.005$ for TL84). For the Oriental patch, one also observes that the area of the Dark condition ellipse is higher than that of the D65 ($p = 0.034$) and the TL84 ($p = 0.012$) conditions.

The orientation of these projections are further explored by plotting the azimuth of the major axis in Figure 4.16. The azimuth denotes the angle made by the major axis of the ellipse with the positive u' axis of the chromaticity plane.

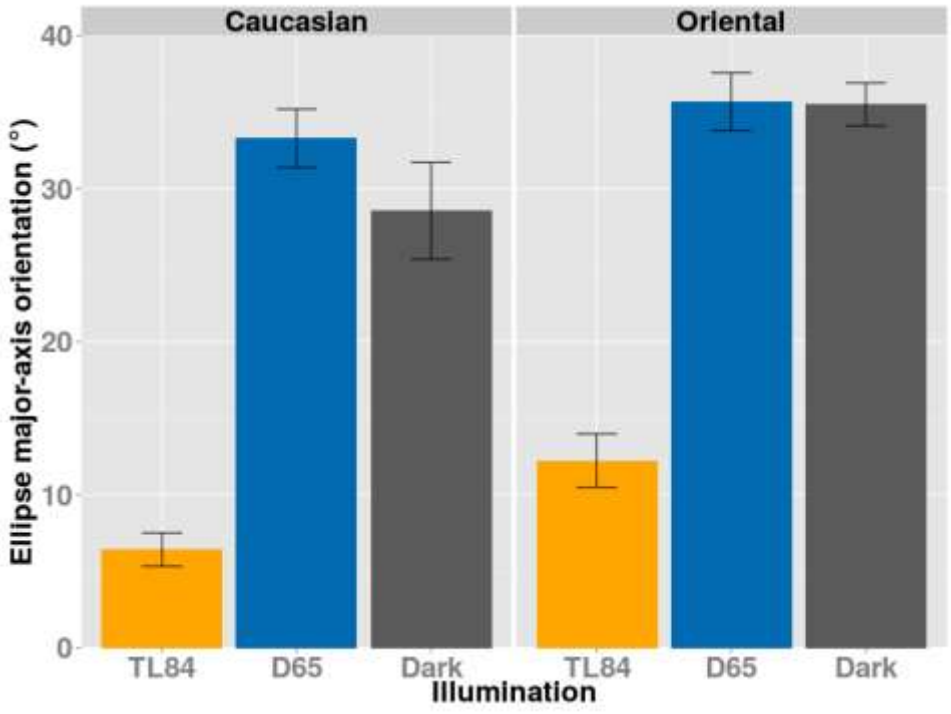


Figure 4.16: Azimuth of the projected ellipse major axis for skin images.

The Harrison-Kanji test shows main effects of both the illumination condition ($F_{2,102} = 71.94, p < 0.001$) and the patch-ethnicity ($F_{1,102} = 7.358, p = 0.0078$), with no interaction between the two factors. Post-hoc analysis using the Watson-Williams test shows that for both patch-types the azimuth is about 25° smaller in the TL84 condition than in the D65 and Dark conditions ($p < 0.001$ in all cases). The D65 and Dark conditions do not differ statistically for both the patches. It is also found that the effect of patch-ethnicity is driven by mainly by the TL84 condition ($p = 0.0077$) with the azimuth being about 5° higher for the Oriental patch.

4.3.4 Discussion

Figure 4.10 shows that the discrimination volume is larger in TL84 compared to D65 for both patch-ethnicities. In other words, irrespective of the ethnicity of a given patch, human observers are better at discriminating small differences in daylight than under

artificial fluorescent lighting. This variation in discrimination ability is not reflected equally in the luminance and chromatic dimensions. While the area of the projected chromaticity ellipses remains roughly the same (Figure 4.15), the luminance projections of the ellipsoids seem to follow the overall changes in the discrimination volume (compare Figure 4.12 and Figure 4.10). One possible reason for this could be the spikes in the TL84 illuminant spectral power distribution (see Figure 4.1) in the 525-650 nm region. This range of wavelengths covers the yellowish part of the visible spectrum, and thus spikes in this region of the illuminant spectrum could imply higher luminance of the simulated skin patch under TL84, thereby leading to an increase in thresholds.

Another effect that merits closer examination is the systematic difference in the azimuths of the projected chromatic ellipses (Figure 4.16). In the light of Figures 4.13 and 4.14, this could be explained in two ways. First, one observes that the azimuths for both the patches seem to be aligned with the daylight locus. This would support the theory that discrimination thresholds are minimal orthogonal to the 'caerulean line' – the line representing natural illuminants (Danilova & Mollon, 2010); in other words, observers tend to confuse colours that lie along the daylight locus more than the colours that lie orthogonal to it. Second, the ellipses also seem to be aligned with the first principal component direction of the colours in the respective skin patches. This is similar to the results obtained by Hansen et al. (2008) who found that isoluminant discrimination ellipses roughly follow the direction of maximum chromatic variation in natural stimuli (banana, orange and lettuce). These two explanations are by no means exclusive, and could be reconciled by the interesting possibility that colour distributions of natural surfaces and textures under varied lighting conditions fall along the daylight locus.

This discussion has thus far mainly focussed on the daylight and fluorescent lighting conditions. This is because the stimuli in these two conditions represent ecologically valid simulations of the skin patches and produce consistent trends in the results. The dark condition, on the other hand, does not represent an ambient illumination, but rather a forced surround illumination. The patches in this condition could easily be

made out by the observers to be self-luminous images displayed on a screen.

In this case, an interesting comparison to analyse is that between the simulated daylight and the dark condition. Although the two conditions use stimuli simulated using the same illuminant (luminaire D65), they have different viewing parameters in terms of the display mode (object mode in D65 vs. self-luminous surface patch in Dark), surround (grey cardboard, reflecting ambient lighting in D65 vs. self-luminous grey screen in the Dark condition), and the ambient lighting ($\approx 51\ cd/m^2$ simulated daylight from an overhead luminaire in D65 vs. $\approx 20\ cd/m^2$ simulated daylight from the surround in the Dark condition). Bearing this in mind, one observes that the ellipsoids under these two conditions display very similar orientations and chromatic ellipse areas (for both patch types). One also observes that in the Dark condition, the two ethnicities display a difference in the volume of the ellipsoids (with the ellipsoids being smaller for the Caucasian patch as compared to the Oriental patch) while no such difference is seen in the D65 condition. This difference is driven entirely by the luminance projections of the discrimination volumes. This could suggest that while the bases mechanisms (which determine the orientations of the ellipsoids) involved in the discrimination are primarily determined by the spectrum of light emitted by the stimulus in the foveal field of view (as the stimuli in the two conditions were identical), the aspect ratio of the discrimination ellipsoid, i.e., the relative weights of the mechanisms involved, are affected by the adaptation and viewing conditions.

4.4 Experiment 1b: Uniform skin-colour patches

Discrimination thresholds for uniform colour stimuli have been measured time and time again – from the initial experiments by MacAdam (1942) and Brown (1951) to the more recent studies such as those by Poirson & Wandell (1990a), Krauskopf & Gegenfurtner (1992), Melgosa et al. (1997), and Yebra et al. (2001). Attempts have also been made to extend these measurements to polychromatic natural stimuli by Hansen et al. (2008).

Experiment 1a looked at discrimination thresholds for polychromatic skin images in detail. But how do these thresholds compare to those for uniform colour patches in the same location in colour space? Furthermore, does the relation change with changes in the ambient illumination condition? The aim of this experiment (Experiment 1b) was to measure thresholds for uniform colour patches under the same three ambient illumination conditions as Experiment 1a.

4.4.1 Stimuli

The stimuli used in this experiment were uniform patches, i.e., patches wherein each pixel was the same colour. These uniform patches had the same colour as the mean $CIELAB$ colour of the skin images used in Experiment 1a. The idea was to investigate how colour discrimination abilities change as one removes texture and second-order colour information from a skin image.

4.4.2 Experimental protocol

The experiment followed the exact same protocol as Experiment 1a, except that the polychromatic skin images were replaced by the uniform patches. A subset ($N = 8$) of the observers tested in Experiment one was recalled. Once again, each observer was tested for about 7.5 hours and 252 thresholds per observer were measured.

4.4.3 Results

In this section, the results of this experiment are presented in a manner analogous to Experiment 1a. Figure 4.17 shows the average ellipsoids for each condition, with symbols carrying the exact same meaning as in Figure 4.9.

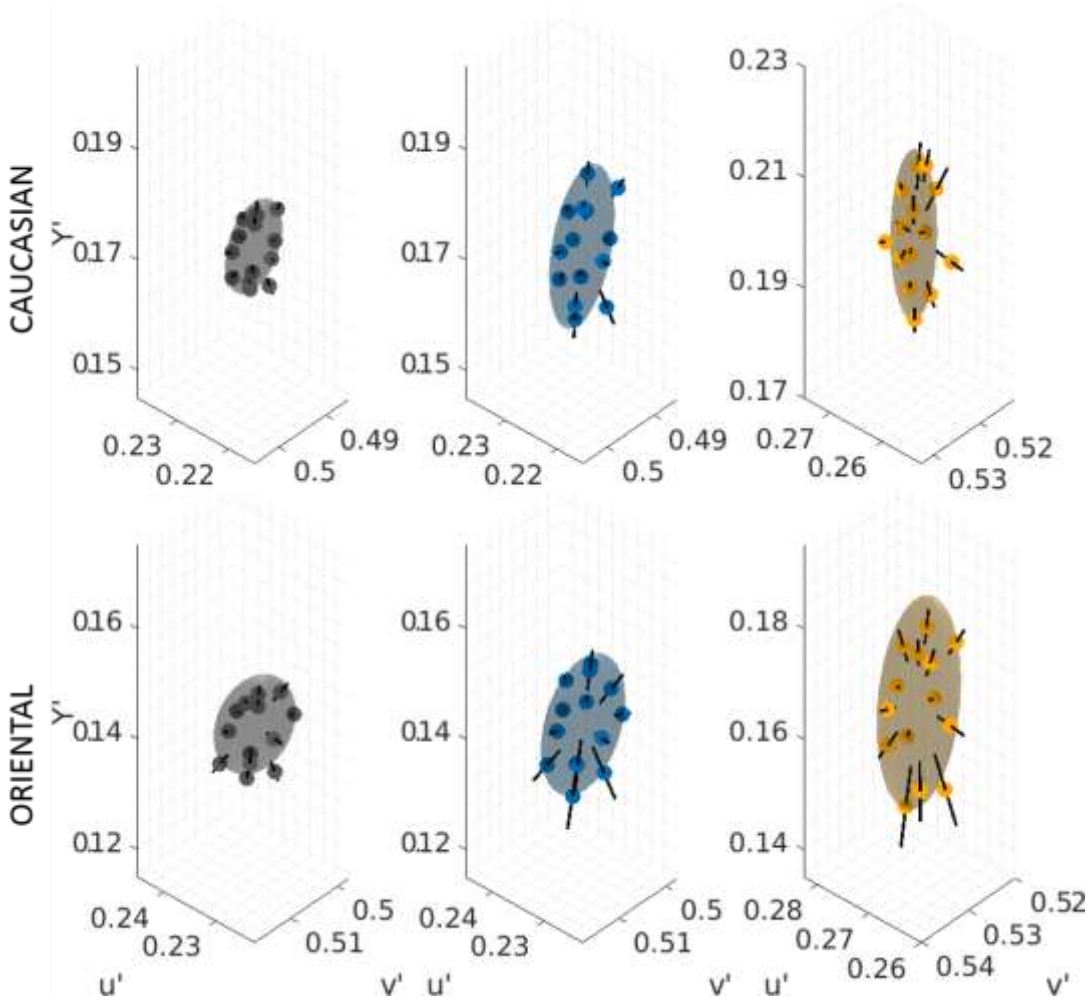


Figure 4.17: Mean ellipsoids for uniform stimuli. The plots use the same symbols as Figure 4.9.

Since the rationale of this experiment relies on the measurements in Experiment 1a, all the subsequent results in this section are presented with the results for Experiment 1a (only for the participants who participated in both experiments). This avoids plotting the same data twice as these comparisons will be frequently referred to in the Discussion section to follow. The statistical analysis presented in this section, though, will only consider data from Experiment 1b. The two patches will still be labelled as the 'Caucasian' and the 'Oriental' patch, the ethnicity names simply referring to the original skin image they were derived from.

Figure 4.18 shows the average volumes of the discrimination ellipsoids. It also shows average volumes for Experiment 1a and 2 for the participants who participated in all the experiments

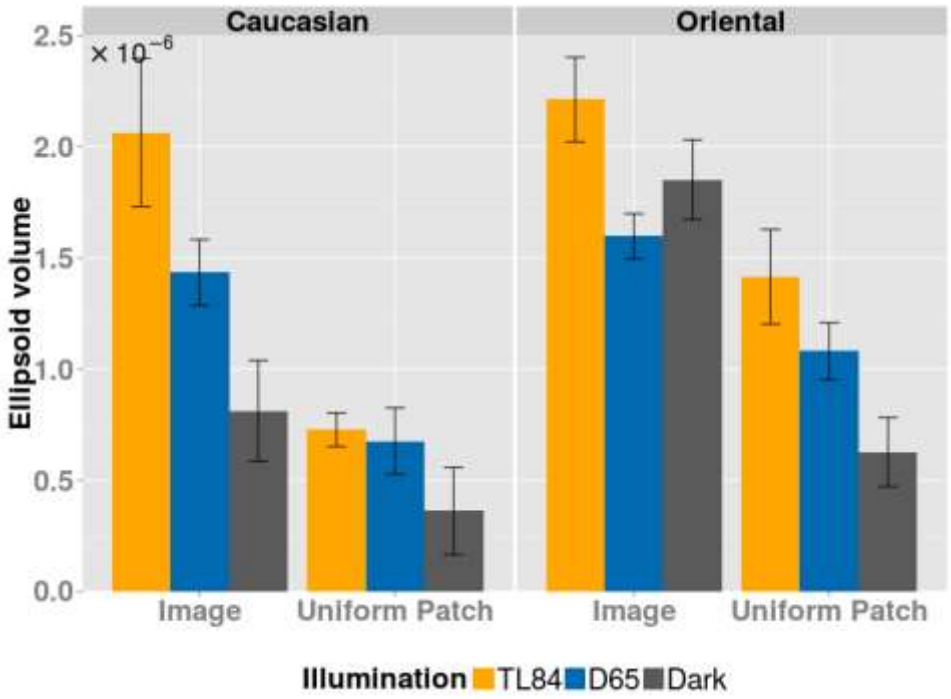


Figure 4.18: Ellipsoid volumes for uniform patches

Repeated-measures two-way ANOVA reveals main effects of both illumination condition ($F_{2,14} = 5.93, p = 0.0136$) and the patch-type ($F_{1,7} = 5.821, p = 0.047$), with no interaction between the two ($p = 0.0991$). Post-hoc tests reveal lower volumes for the Dark condition compared to the D65 ($p = 0.031$) and TL84 ($p = 0.021$), and a higher volume for the Oriental patch ($p = 0.0014$).

Analogous to the analysis for Experiment 1a, these volumes are first projected on the luminance axis. The length of these projections is shown in Figure 4.19.

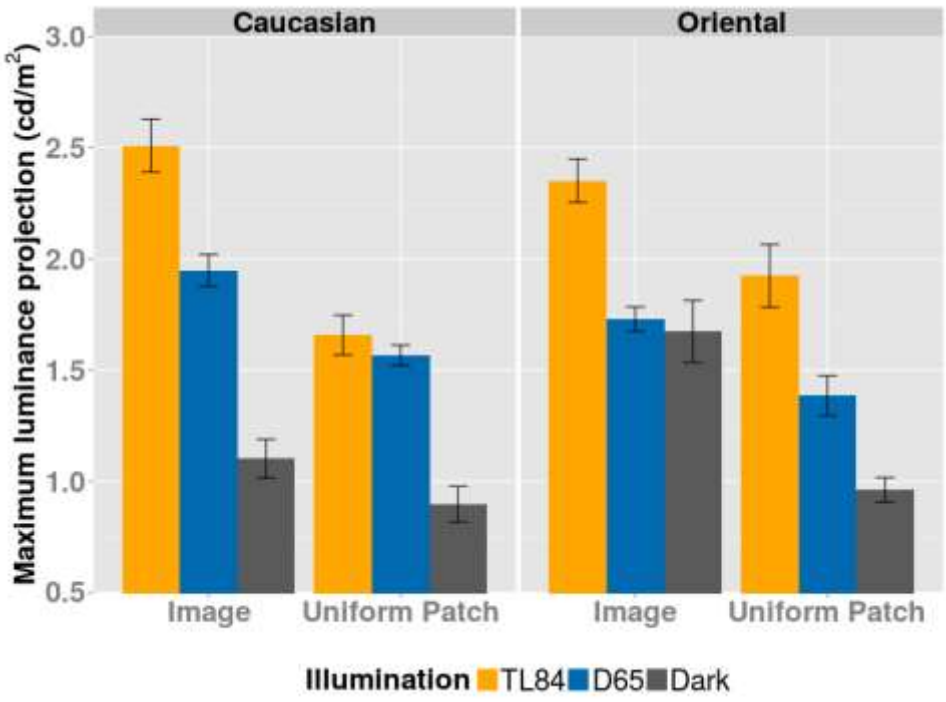


Figure 4.19: Luminance projections for uniform patches

A repeated measures ANOVA shows a main effect of illumination ($F_{2,14} = 48.48, p < 0.001$), no main effect of patch-type ($p = 0.641$), and a significant interaction of the two ($F_{2,14} = 4.84, p = 0.0252$). Post-hoc tests show the thresholds for the Dark condition to be lower than the D65 and TL84 conditions for both patch types (Caucasian: $p < 0.001$ for both TL84 and D65; Oriental $p = 0.002$ for TL84 and $p = 0.044$ for D65).

Next, chromaticity projections of the ellipsoids are examined in the figures below (Figures 4.20 and 4.21). The average ellipses for the uniform patches (this experiment) as well as skin images (Experiment 1a) are plotted, with only participants common to both experiments being included.

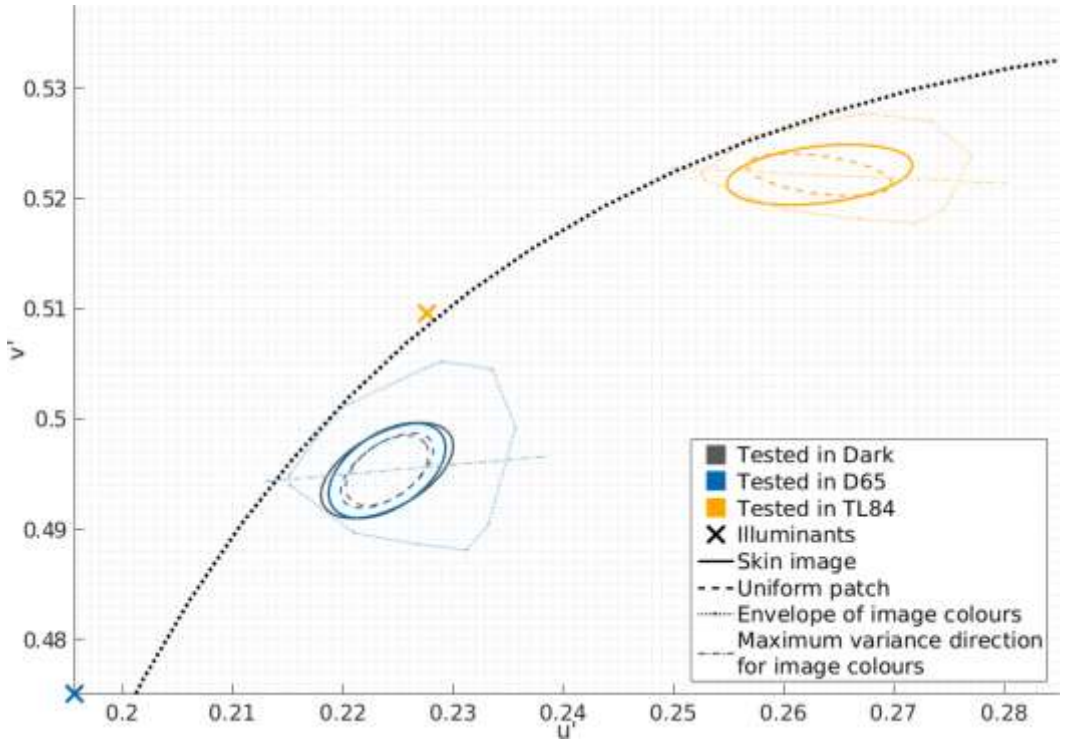


Figure 4.20: Caucasian patch-type: Average ellipses for uniform and skin patches

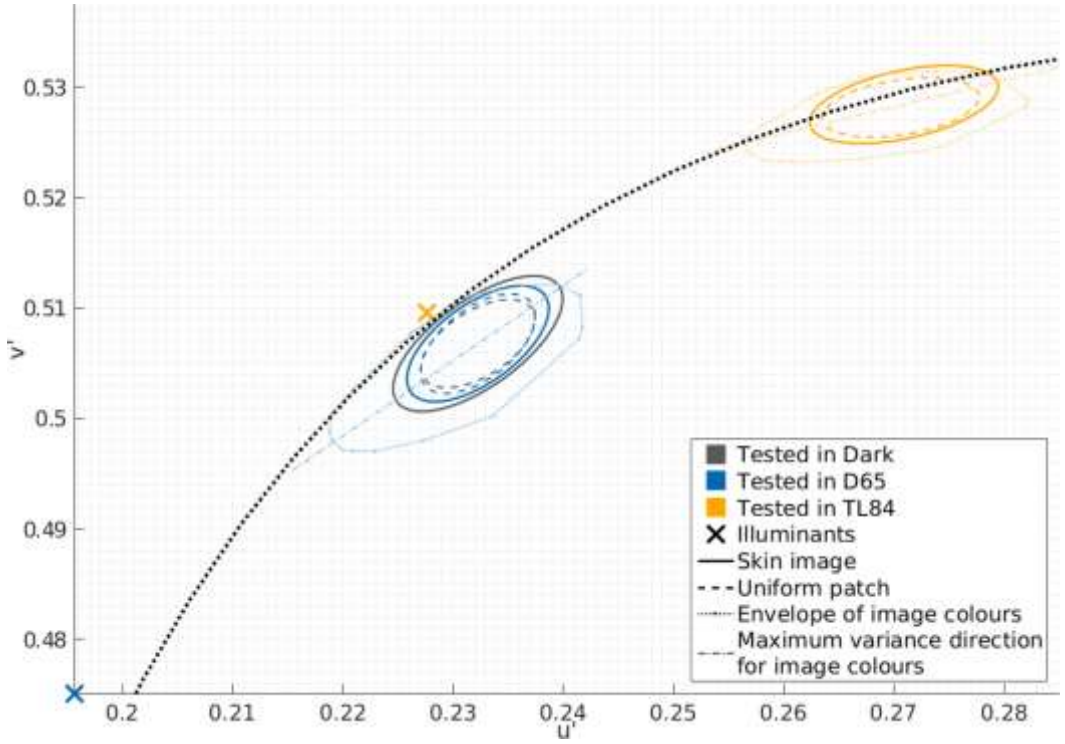


Figure 4.21: Oriental patch-type: Average ellipses for uniform and skin patches

The area of these ellipses is shown in Figure 4.22.

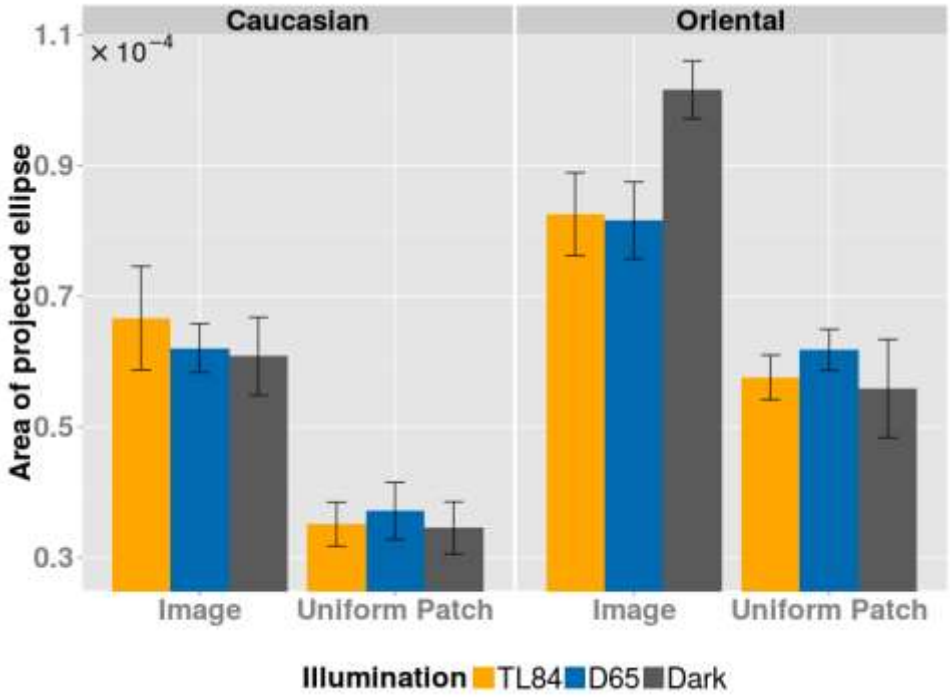


Figure 4.22: Area of chromatic projections for uniform patches

A repeated measures ANOVA shows a main effect of patch-type ($F_{1,7} = 28.78, p = 0.001$) with no effect of the illumination condition. The areas of the projections for the Oriental patch are much larger than those for the Caucasian patch.

Further, Figure 4.23 shows the azimuths for the projected chromatic ellipses. The Harrison-Kanji test shows significant main effects of the illumination condition ($F_{2,42} = 24.57, p < 0.001$), whilst showing no effect of the patch-type or an interaction effect. Post-hoc analysis using the Watson-Williams test shows that for both patch-types the azimuth is smaller in the TL84 compared to the D65 and Dark conditions ($p < 0.001$ in all cases). The Dark and D65 conditions do not show any difference. Although the effect is not powerful enough to be statistically significant as a main effect, from the figure one can observe a difference in the azimuths between the two patches in the TL84 condition. This is indeed verified by a post-hoc paired t-

test ($F_{1,14} = 10.5445, p = 0.0058$).

The results shown in this section are representative of the sample, verified by the individual projected ellipses for each observer which are included in Appendices 2.1 and 2.2.

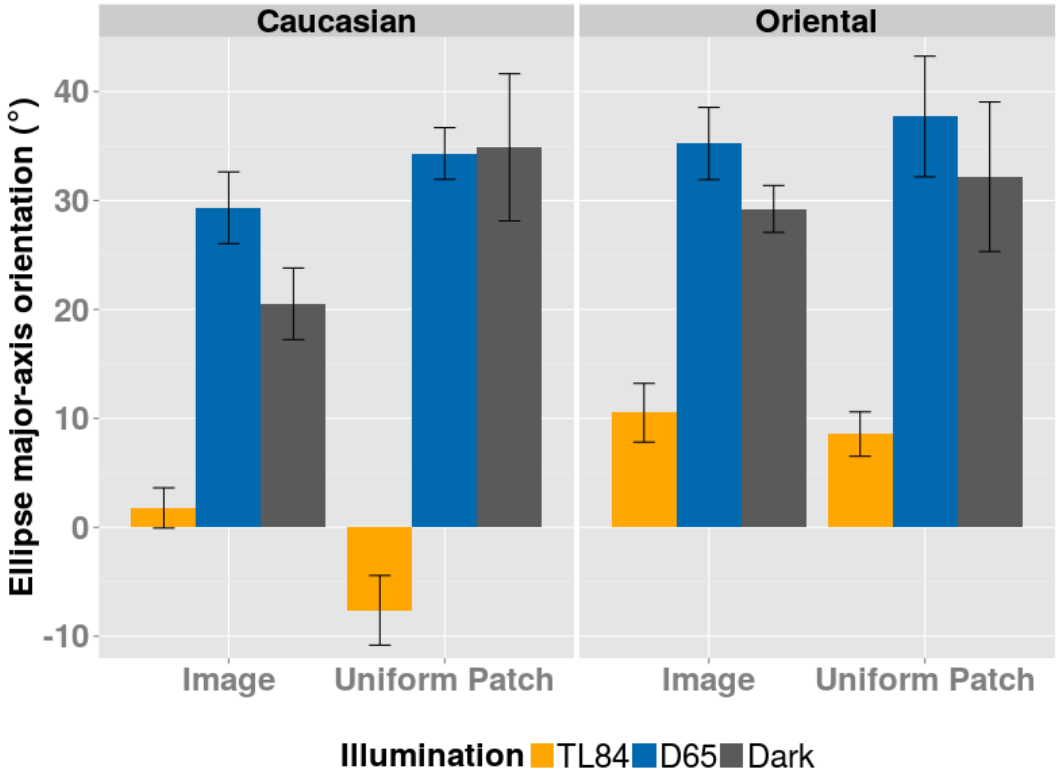


Figure 4.23: Azimuth of chromatic projections for uniform patches

4.4.4 Discussion

Experiment 1a had estimated the discrimination thresholds for skin-patch images under varying illumination conditions. An important question to emerge from these measurements was how different these thresholds were from those for uniform colours under the same experimental conditions. Clearly, the first trend one notices is that the

ellipsoids for skin images are about 2-3 times larger than those for the corresponding uniform patches (mean $CIELAB$ colour). This is better shown in Figure 4.24.

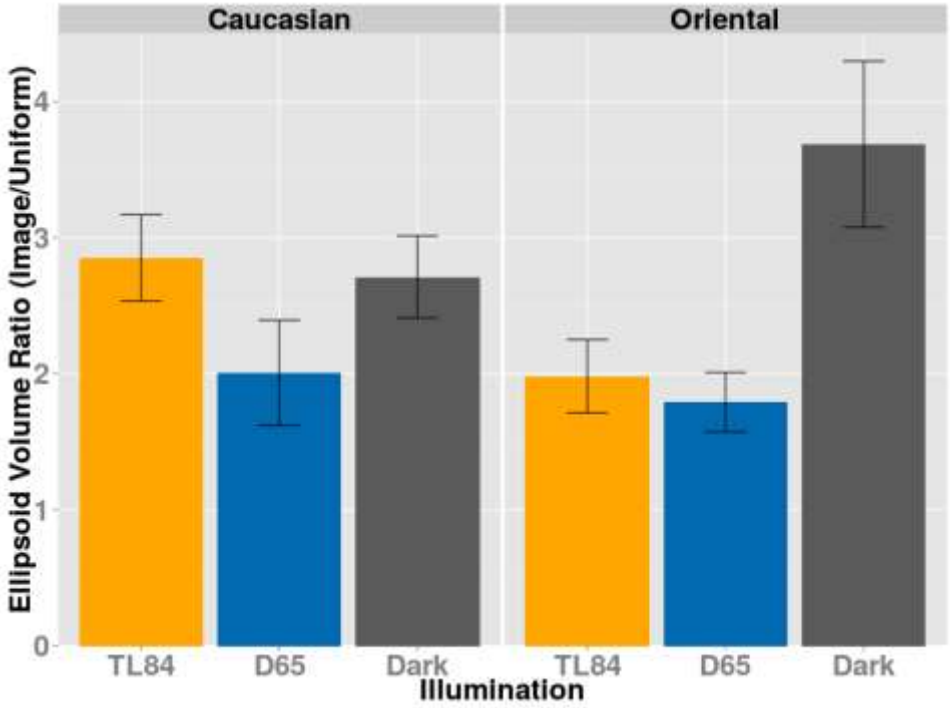


Figure 4.24: Ratio of volumes for the skin image and the uniform patch

This shows that skin textural properties lower the discriminability of the patch compared to uniform patches, leading to an increase in the thresholds. This difference is found in both luminance and chromatic thresholds, though the ratios are higher for chromatic projections (Figure 4.25) than luminance projections (Figure 4.26).

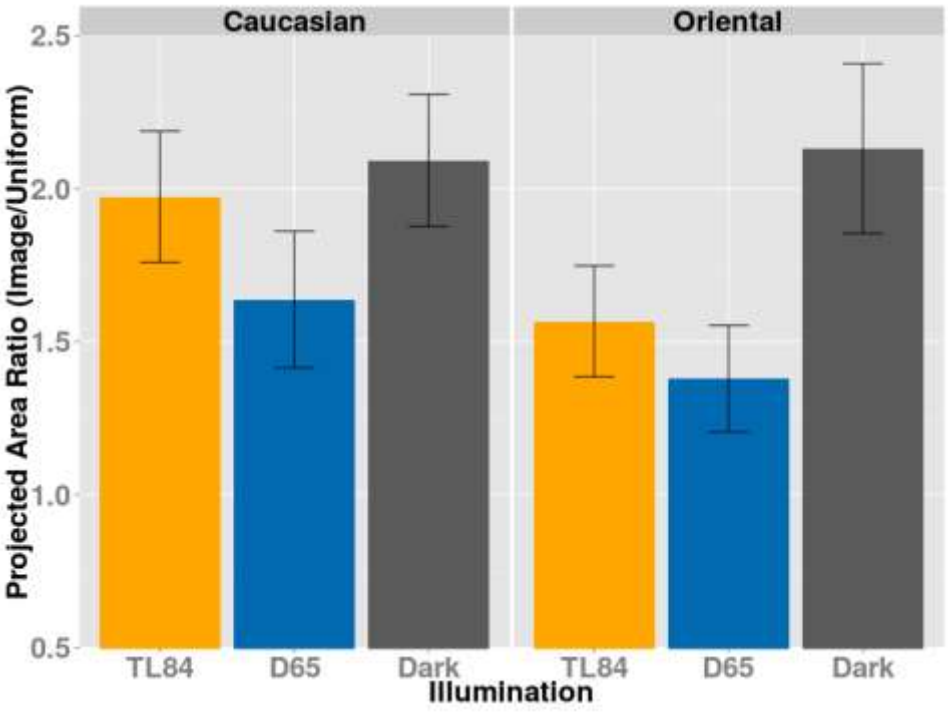


Figure 4.25: Ratio of area of chromatic ellipses for the skin image and the uniform patch

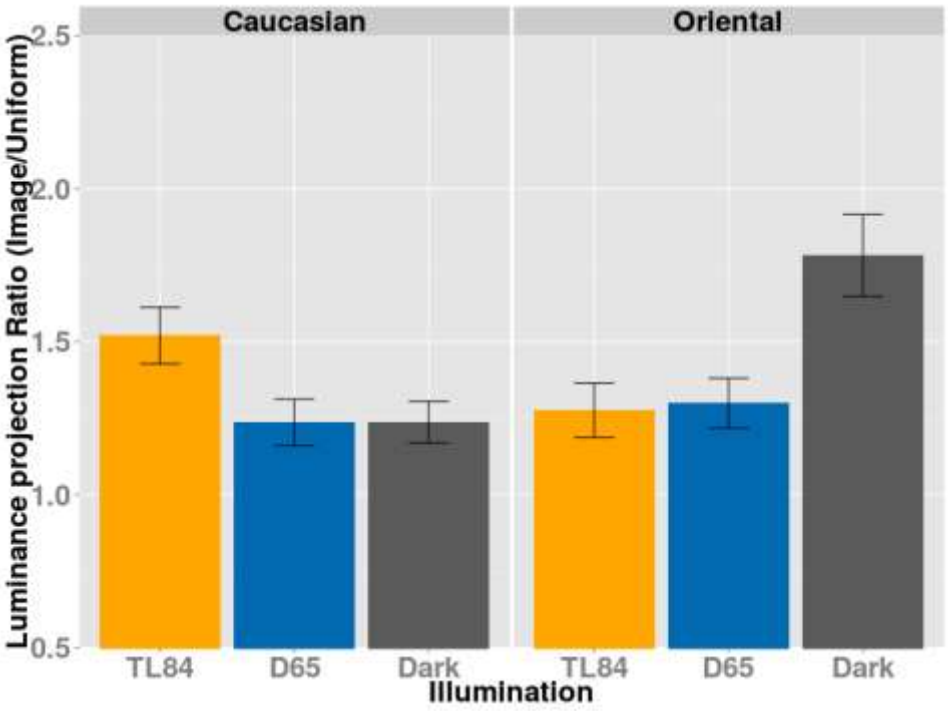


Figure 4.26: Ratio of luminance projections for the skin image and the uniform patch

A similar increase in chromatic thresholds was also reported by Hansen et al. (2008) for natural textures, and synthetic textures with colour distributions similar to natural textures. Montag & Berns (2000) also reported similar effects in luminance thresholds. A possible explanation of these results could lie in the theory proposed by Webster & Mollon (1997) that polychromatic natural stimuli entail adaptation not only to the mean luminance of the scene, but also a contrast adaptation to the colour distribution within the scene. They reason that although light adaptation could adjust for changes in mean colour, it cannot compensate for changes in the statistics of the colour distributions. They further propose that contrast adaptation mechanisms might operate by whitening the stimulus colour distribution based on changes in post-receptoral channel tunings, with new tunings emerging due to inhibition between channels which produce the most correlated responses (Atick, Li, & Redlich, 1993; Barlow & Földiák, 1989; Webster & Mollon, 1997). Considering that in the current study the observer could view the whole scene, even if they fixated on the stimulus, one cannot completely ignore contrast adaptation regardless of whether the tested stimuli were uniform or textured. Even so, it is likely that the amount of possible contrast adaptation in case of the uniform patch stimuli was lower than that for the simulated skin patches (since there is no contrast within a uniform stimulus). This would mean that in the case of uniform patches, the observers were comparatively less adapted, and hence less capable of constancy or discounting the illuminant, resulting in the ability to notice smaller changes in the stimulus colour, i.e., better discrimination performance (or lower thresholds) compared to the simulated skin patches.

It was noted in Experiment 1a that the thresholds for both patch-ethnicities were higher in TL84 than in D65 (Figure 4.10). Figure 4.18 shows that uniform patches produce statistically similar discrimination volumes for the TL84 and D65 conditions (although an analysis of simple effects suggests a marginally significant difference for the Oriental patch with $p = 0.042$). This would suggest that second order properties of the colour distribution seem to modulate the discrimination thresholds such that it not only increases the number of discriminable configurations, but also makes observers markedly better at discrimination under D65 as compared to TL84.

Perhaps the most interesting effect when comparing discrimination of uniform patches and skin images is the orientation of the chromatic ellipses. In Experiment 1a, the ellipses for skin images were shown to lie along the daylight locus (Figures 4.13 and 4.14). Figures 4.20 and 4.21 show that although this is still the case for the corresponding 'Oriental' uniform patch, the 'Caucasian' uniform patch shows a clear deviation. This effect is robust and can be observed for most observers (refer to figures in Appendix 2). The fact that this effect was found under a specific illumination (TL84) once again points towards an effect of second or higher order textural properties on the discrimination boundaries, hinting at possible contrast adaptation mechanisms (Webster & Mollon, 1997).

4.5 Experiment 2: Swapped colour distributions

The previous two experiments measured discrimination thresholds for skin patches, and compared them to those for uniform patches of the same mean colour. It was found that the presence of skin texture changes the thresholds one would otherwise obtain for uniform patches. Going back to Eq. 4.23 which was used to generate these stimuli, one sees that the set of tristimulus values X_i is the form of the stimulus that reaches the photoreceptors in the eye, while the reflectance spectrum $\hat{\boldsymbol{r}}_{\boldsymbol{pixel}}(\lambda)$ is the physical property of skin which generates the tristimulus values under a given illumination. Thus, the stimuli used in these experiments can be interpreted in two, mutually non-exclusive ways. First, as representations of a physically plausible or known surface – skin being viewed under the given illumination, and second, as textures with a colour distribution which resembles skin in appearance (under the given illumination) but may or may not necessarily represent an actual skin surface. The first interpretation (known surface) is driven by an implicit knowledge of the physical reflectance of the stimulus while the second interpretation can be sufficiently represented simply by considering the distribution of the colour in the stimulus.

If one were to disregard the fact that the origin of the textures used in the previous two

experiments lies in the interaction between the illumination and the reflectance spectrum of human skin, one arrives at the interesting question as to whether higher order properties of the texture in colour space would, by themselves, allow for predictable changes in discrimination thresholds. In other words, if one were to change the mean colour of a textured patch while maintaining its higher order colour characteristics intact, would the thresholds change as a function of the new mean colour; or would they remain constant because the illuminant and the relative distribution of colours has not changed? If one were to abstract the idea further, are the spectral properties of the textured stimulus which identify it as skin reducible simply to relative colour distributions of the texture? The next set of experiments were aimed at answering these questions. The colour distributions of the simulated skin patches in the two illumination conditions (D65 and TL84) were translated such that their means were swapped, while the relative distributions remained intact in $CIELAB$ space. Three initial hypotheses were considered. H1: If the thresholds are affected only by the illumination condition and the relative $CIELAB$ distribution of colours, swapping the means should not affect the thresholds and they should be identical to Experiment 1a. H2: On the other hand, if the thresholds are completely driven by the location of the textures in colour space, their discrimination thresholds should be swapped as well. H3: The most likely possibility is that the thresholds neither remain unchanged, nor are completely swapped, but show an intermediate change.

4.5.1 Reference stimuli and illumination conditions

In Experiment 1a, the *reference stimuli* were colour-accurate skin-patch images under simulated daylight and fluorescent illuminants such that their colour values were consistent with the ambient illumination. For a given patch, the reference images used in the current experiment were obtained by swapping the average $CIELAB$ colour of the D65 and CWF simulations, while maintaining the same relative colour-distributions (also in the $CIELAB$ space). In other words, in each illumination condition, the second order properties of the colour distribution of the image were the same as in Experiment 1a, but the first order properties between the two illumination conditions were swapped (Figure 4.27).

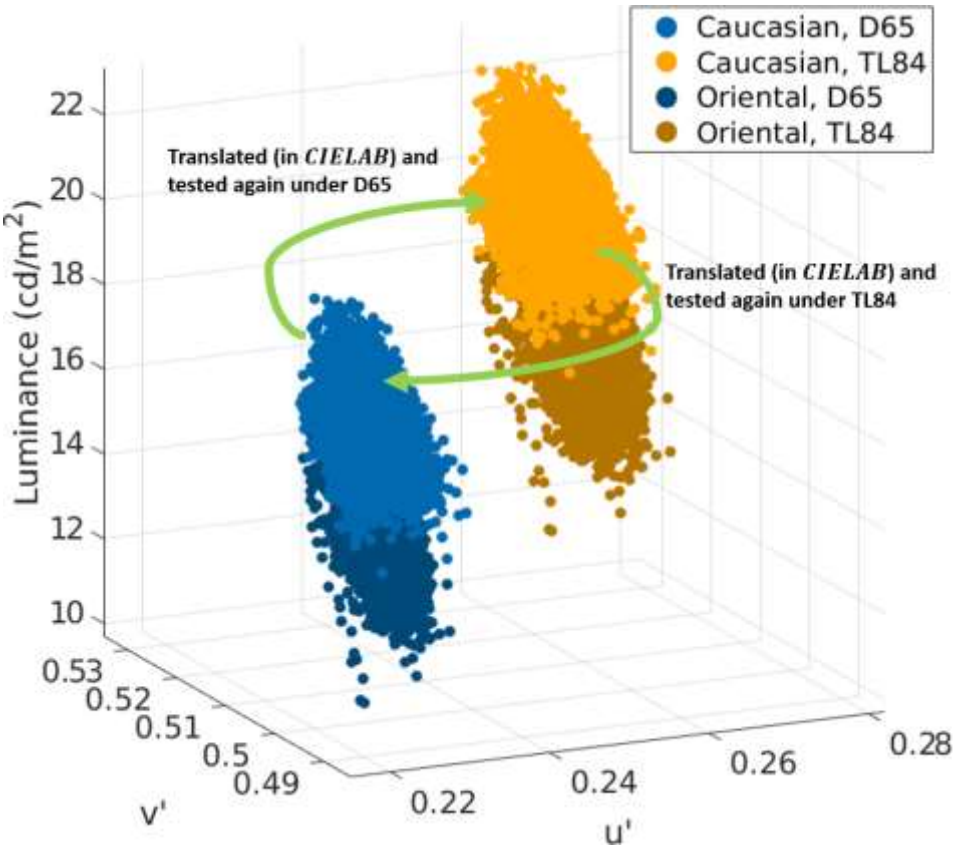


Figure 4.27: Polychromatic stimuli for Experiment 2.

Only stimuli based on the original Caucasian patch were tested to reduce the testing time per participant. This choice was also based on the fact that the difference in orientation of the chromatic ellipses for the polychromatic and uniform stimuli was observed only in the Caucasian case. The observers were also tested with uniform patches derived from the mean $CIELAB$ colours of the stimuli. The experiments were carried out in simulated daylight and fluorescent lighting conditions. The dark condition was dropped.

Given a certain (say, Caucasian) patch with a pixel-wise reflectance $\hat{\boldsymbol{r}}_{\boldsymbol{pixel}}(\lambda)$ we have two sets of absolute tristimulus values corresponding to each illumination condition. These are given by applying Eq. 4.23. If the two illuminant SPDs are denoted by $\boldsymbol{L}_{\boldsymbol{D65}}(\lambda)$ and $\boldsymbol{L}_{\boldsymbol{TL84}}(\lambda)$ respectively, the two sets of tristimulus values are given by

$$\boldsymbol{X}_{i,j} = \int_{visible} \boldsymbol{L_j}(\lambda) \cdot \hat{\boldsymbol{r}}_{\boldsymbol{pixel}}(\lambda) \cdot \bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda) \cdot d\lambda\,, \qquad \text{Eq. 4.24}$$

$$i \in \{1,2,3\}\,, j \in \{D65, TL84\}$$

where λ and $\bar{\boldsymbol{x}}_{\boldsymbol{i}}(\lambda)$ represent the wavelength and the CIE colour matching functions while the variable j represents the illumination. A question that naturally arises out of the required manipulation is: In which colour-space should the swapping of the means be defined? If we were to define them in an absolute space such as the $CIE\ 1931\ XYZ$ space or the $u'v'Y'$ space (used for plotting all the results in the study), we swap the absolute tristimulus value distributions, with no guarantee that the manipulated patches will still retain the appearance of skin to any credible degree. In fact, such transforms were tested by displaying textures manipulated in $u'v'Y'$ space (in the light-proof booth under the intended illumination conditions), and the patches were found to have an appearance quite different from skin. None of the consulted observers (CN, CT and PR) found the patches to be identifiable as skin.

To maintain the *skin-like* appearance of the stimuli, it was decided to transform the colour distributions of the patches to a common $CIELAB$ space (using the same white-point for normalisation as used for defining sampling directions in colour space, see Section 4.2.2) and swapping the means in this space. The same set of observers (CN, CT and PR) now reported the patches obtained by the $CIELAB$-swap to resemble skin much more than the patches obtained by a $u'v'Y'$-swap. It must be noted that when one uses an absolute space (such as the $u'v'Y'$ used here) to plot the $CIELAB$-swapped distributions, the shapes of the distributions appear to change, and distributions centred at the same location tend to resemble each other (see Figure 4.28). This also explains why observers found these $CIELAB$-swapped distributions to resemble skin much more than the $u'v'Y'$-swapped ones.

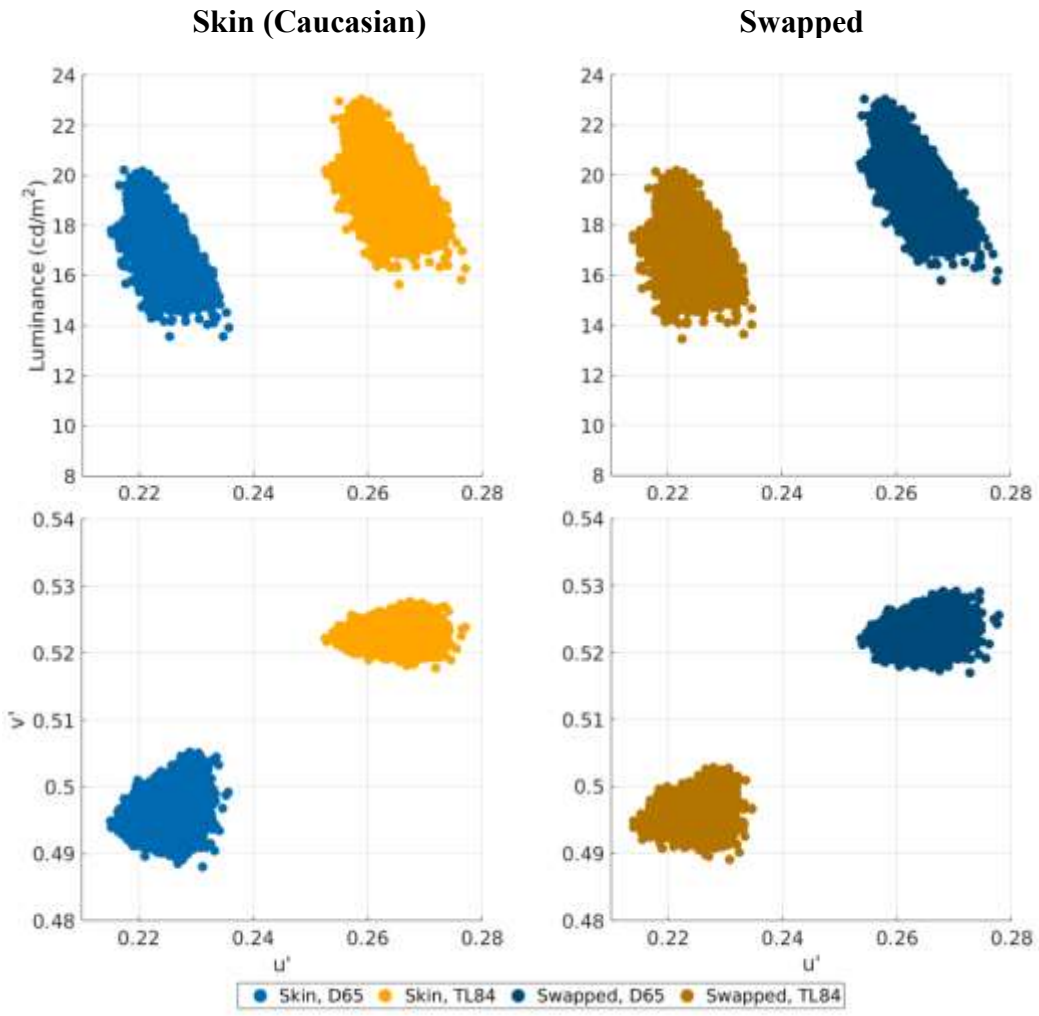


Figure 4.28: Luminance and chromatic spreads of the simulated skin patches (left column) and the stimuli generated by swapping their relative distributions in $CIELAB$ space (right column). In each case, the colour of the points represents the illumination used for testing the corresponding patch (light and dark blue: D65; yellow and brown: TL84). The first row shows the luminance spreads (luminance in cd/m^2 along the ordinate, u' along abscissa) while the second row shows chromatic spreads (v' ordinate, u' abscissa). Notice that simply swapping the distribution means in $CIELAB$ space results in changes in the shapes of the entire distributions in $u'v'Y'$ space.

To further quantify the change in the shapes of the graphs, the following table (Table 4.3) provides the key parameters of the Caucasian and the swapped-Caucasian patches. An interesting observation that emerges from Table 4.3 is that swapping the mean colour of the patches in $CIELAB$ space clearly changes the relative distribution of the colours in the $u'v'Y'$ space as well. The effects are a swapping, not only of the mean colour, but also of the approximate luminance range and the projected chromatic area of the two gamuts. Thus, by maintaining the relative distributions of the colours in a

common appearance space (namely, the $CIELAB$ space normalised using a common white-point), we have managed to preserve the skin-like appearance of the stimuli.

	Volume ($\times 10^{-6}$)	Luminance range (cd/m^2)	Area ($\times 10^{-4}$)	Orientation of first PC (°)
Caucasian patch				
D65	7.44	6.63	2.42	5.1 (var. exp. = 61%)
TL84	5.82	7.42	1.69	-2.7 (var. exp. = 87%)
Swapped patch				
D65	5.93	7.28	2.18	2.7 (var. exp. = 79%)
TL84	7.28	6.75	2.04	-4.8 (var. exp. = 74%)

Table 4.3: The colour distributions of Caucasian and the swapped-Caucasian patches described using the same four parameters as Table 4.2 – volume of the 3-D hull of the gamut in $u'v'Y'$ space, luminance range, area of the hull covering the projections of the gamut on the $u'v'$ space chromaticity plane, and the orientation of the first principal component found by a Principal component analysis of the chromatic projection of the data.

4.5.2 Experimental protocol

The experiment followed the exact same protocol as Experiments 1 and 2. A subset of the observers ($N = 6$) who had also participated in both Experiment 1a and 1b was recruited. In total, 168 thresholds (2 illuminants × 2 stimuli types (uniform and textures) × 3 repetitions × 14 measurement directions) were measured per participant, amounting to about 6 hours of testing.

4.5.3 Results

This section presents results for observers ($N = 6$) common to both experiments. The statistical analysis here only pertains to the results for this experiment (Experiment 2), while comparisons with previous experiments (Experiments 1a & 1b) are made in the subsequent discussion (Section 4.5.4). To avoid multiple plots of the same data, the plots in this section present the results from both Experiment 1 (left) and Experiment 2 (right).

The ellipsoid volumes are plotted in Figure 4.29. The results from Experiment 2 are shown in the panel to the right. The locations of the two skin simulations used in Experiment 1a (shown in the small window on the top-right) are labelled as **T** (TL84) and **D** (D65). Since the positions of the stimuli are swapped in Experiment 2, these symbols also denote the location of the swapped stimuli under TL84 (**D**) and D65 (**T**) respectively. These labels will be used throughout this section to code the location of the stimulus in $u'v'Y'$colour space. Thus, while the labels **T** and **D** encode the spatial location of patches, the colour of the bars represents the illumination condition. In addition, the figure also marks the predictions as per the first two hypotheses proposed in Section 4.5. H1 denotes the hypothesis that the thresholds are driven solely by the illumination condition, while H2 represents the hypothesis that the thresholds are driven exclusively by the location of the stimulus in colour space.

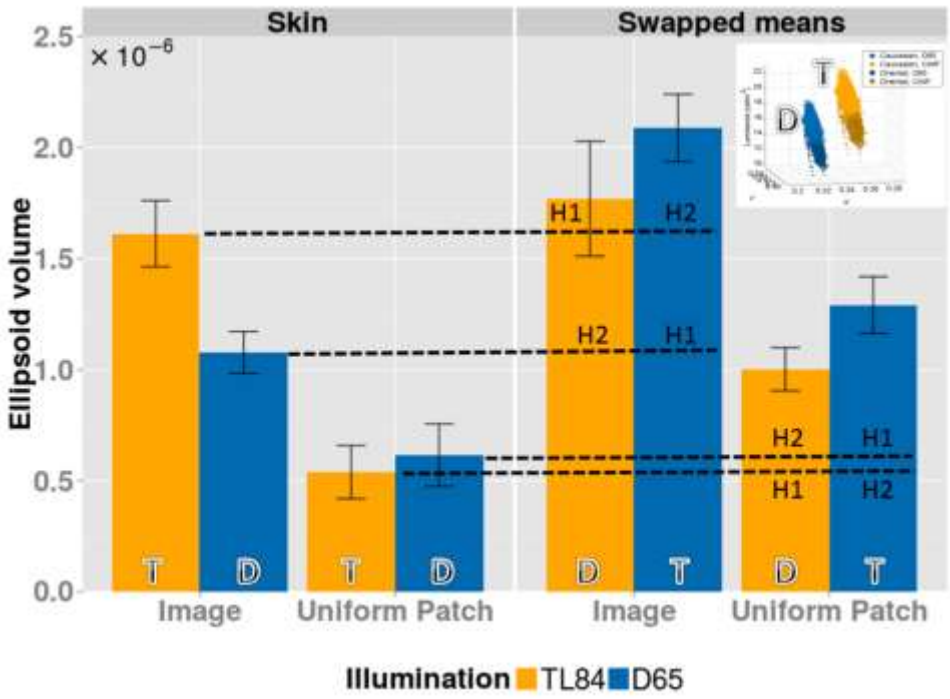


Figure 4.29: Ellipsoid volumes for participants common to all conditions. Left panel shows data for stimuli representing simulations of skin colour and texture under TL84 and D65 illuminants. Right panel shows the thresholds for stimuli derived by swapping mean colour of skin simulations for the two illumination conditions while maintaining the relative colour distribution of the stimuli. The locations of the patches in $u'v'Y'$ colour space are labelled as **T** and **D**. The label **T** corresponds to the location of the simulated skin patch under TL84 (and the swapped patch under D65), while the label **D** corresponds to the location of the simulated skin patch under D65 (and the swapped patch under TL84). The original distributions are shown in a small window on the top-right. Furthermore, the graph also shows predictions as per the first two hypotheses proposed in Section 4.5. H1: The thresholds are driven by the illumination condition only, H2: The thresholds are driven by the location in colour space only. The results show neither H1 nor H2 holds completely.

A two-way repeated measures ANOVA shows that there is no statistical effect of illumination condition ($p = 0.0688$), but a clear difference in the thresholds between the textures and the uniform patch ($F_{1,5} = 15.05, p = 0.0116$), with the uniform patches having lower thresholds. Here, the degrees of freedom derive from the fact that it is a within-subjects 2×2 design with $N = 6$ participants. There are two factors, illumination condition (say, A) and the stimulus-type (say, B), each with two levels (Illumination condition: D65 and TL84, Stimulus type: Textured and Uniform patches), say n_A and n_B. Here, $n_A = n_B = 2$. In a two-way repeated measures ANOVA, the main effect (say, of factor B) has the degree of freedom $n_B - 1$, with an

associated $(n_B - 1) \times (N - 1)$ degrees of freedom for the error. In the present case, since $n_A = n_B = 2$ and $N = 6$, the F-statistic for the main effect is reported with the degrees of freedom $(2 - 1, (6 - 1) \times (2 - 1)) = (1, 5)$.

The next set of figures examines these threshold ellipsoids in terms of their luminance (Figure 4.30) and chromatic (Figure 4.31) projections. The symbols used in these plots are identical to Figure 4.29. It can be seen that the differences between the volumes of the uniform and texture stimuli emerges from the chromatic component of the ellipsoids ($F_{1,5} = 31.3, p = 0.00252$), with the luminance projections remaining very similar ($p = 0.156$). It is also found that although the volumes of the ellipsoids are not different under the two illuminants, both the luminance and chromatic projections show significant differences. The luminance projections are higher under TL84 ($F_{1,5} = 95.24, p < 0.001$) whereas the area is higher under simulated daylight ($F_{1,5} = 45.41, p = 0.001$), the two effects cancelling each other out to give similar volumes.

The chromatic thresholds in the in $u'v'Y'$ colour space are further illustrated in Figure 4.32 (polychromatic stimuli) and Figure 4.33 (uniform patches). In addition to the chromatic projections of the thresholds, these plots also show the chromatic envelopes of the colour gamuts of both the actual and the swapped skin patches, along with the directions of the first principal components of each gamut.

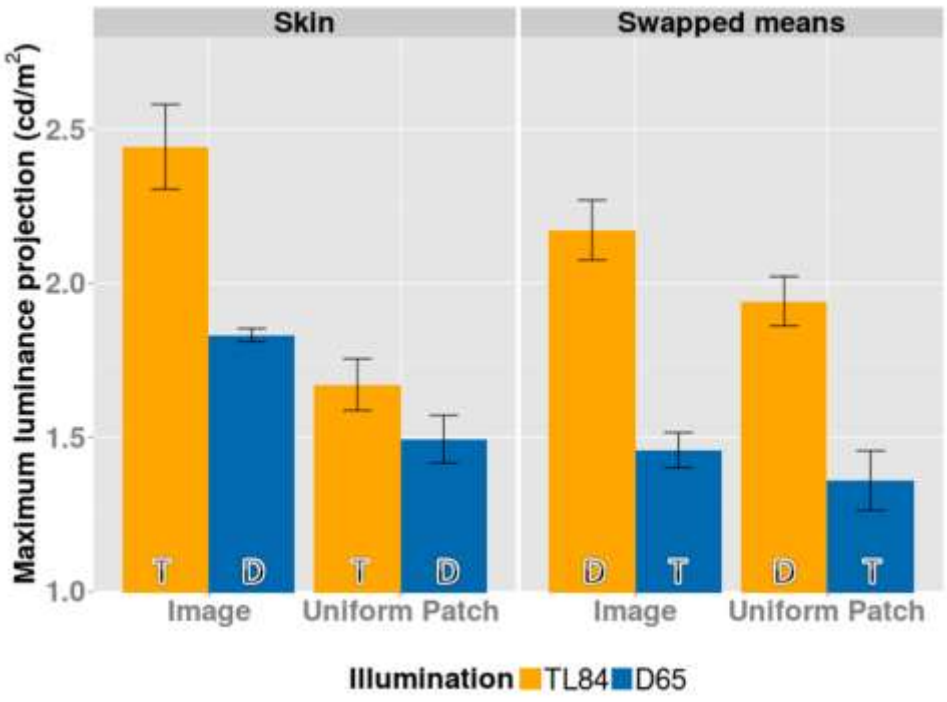


Figure 4.30: Luminance projections of ellipsoids for participants common to all conditions.

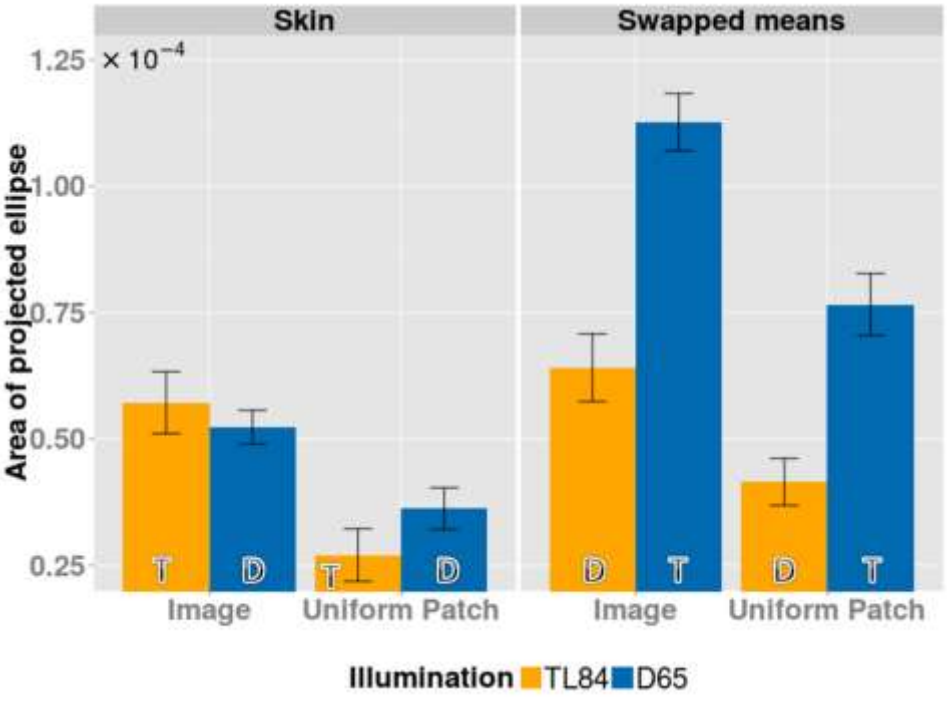


Figure 4.31: Area of chromatic projections of the ellipsoids for participants common to all conditions.

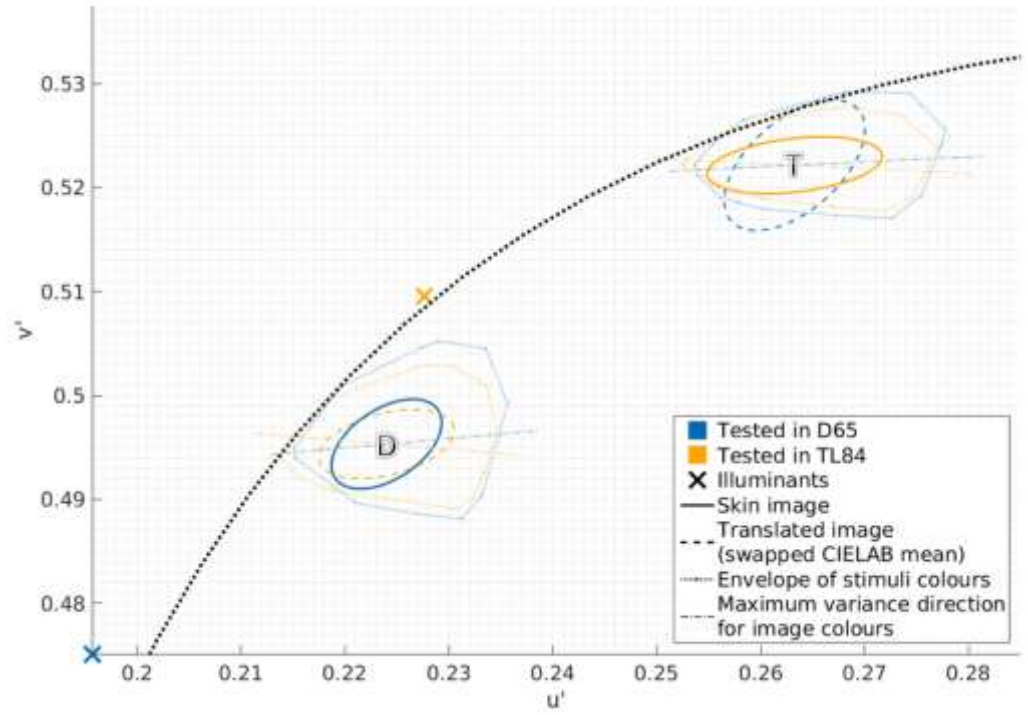


Figure 4.32: Average ellipses for Experiments 1a and 2a (polychromatic textured stimuli)

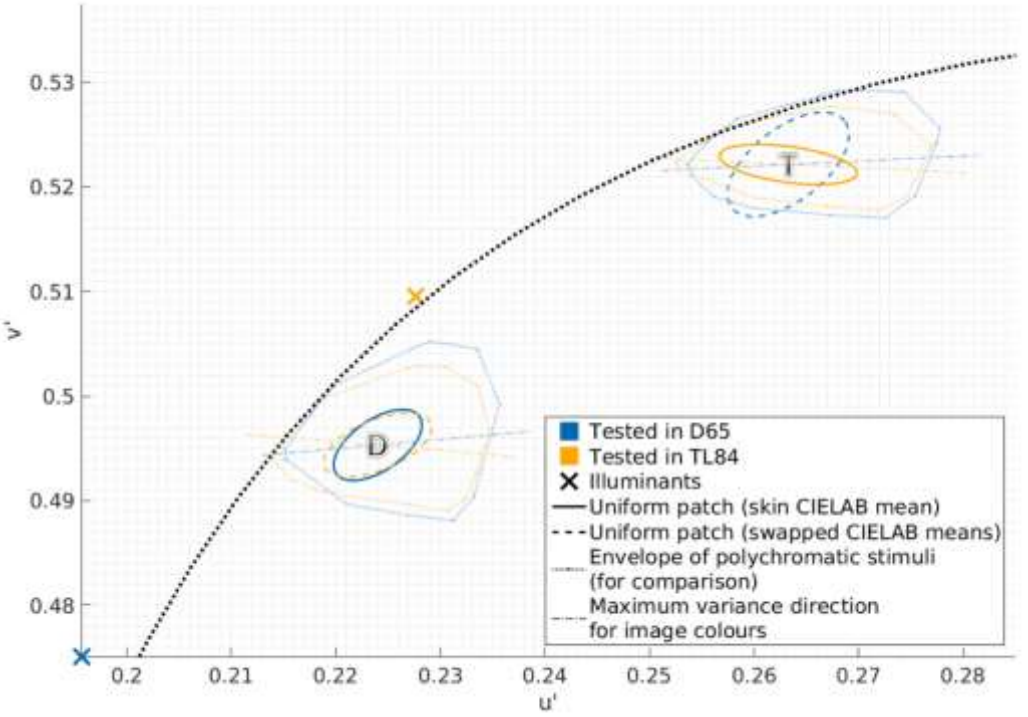


Figure 4.33: Average ellipses for Experiments 1b and 2b (uniform stimuli)

Figure 4.34 below shows the azimuth angles for the chromatic ellipses. One sees a clear effect of illumination condition on the azimuths (the Harrison-Kanji test yielding $F_{1,20} = 52.5276, p < 0.001$) with the ellipses for TL84 lighting being closer in orientation to the u' axis than the D65 ellipses. This is true for both the textured and the uniform conditions. No effect of the type of stimulus (textured/uniform) or an interaction thereof is observed.

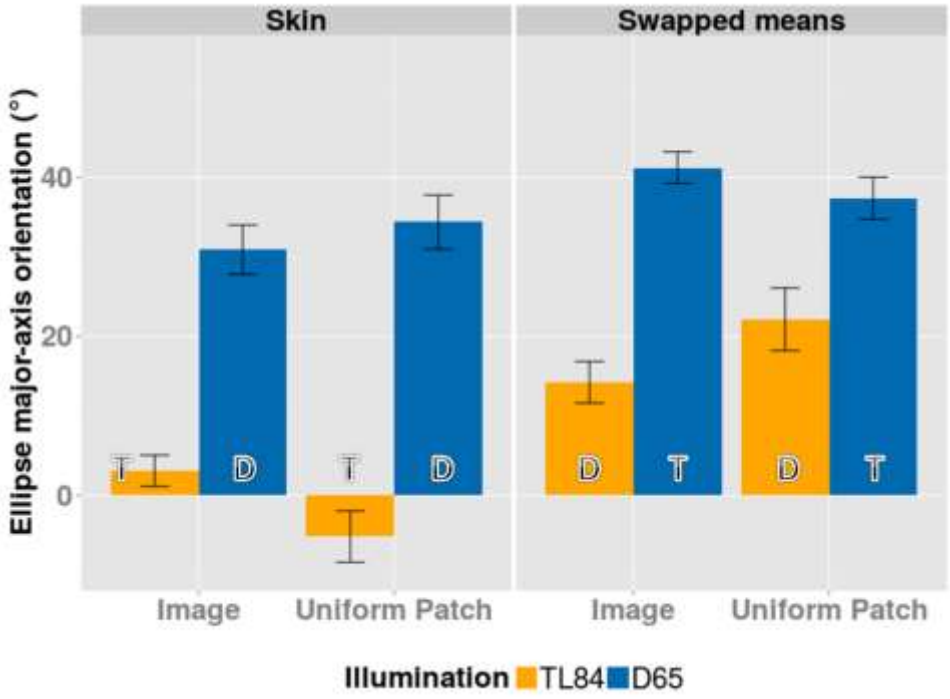


Figure 4.34: Azimuth of chromatic projections for participants common to all conditions.

4.5.4 Discussion

Experiments 1a and 1b measured discrimination thresholds for simulated skin and uniform patches in various illumination conditions. Experiment 2 further investigated the effect of changing the mean colour of the stimuli while maintaining the relative distribution of colours in the textures. This translation of the mean chromaticity was done such that the mean colours of the stimuli under the two illumination conditions

were swapped in $CIELAB$ space.

It must be noted that in doing so, between Experiment 1 and 2, one examines the same mean location in colour space under two different illumination conditions. The hypotheses are that any change in thresholds at the same location would prove an effect of the ambient illumination (H1), and any change between two locations would prove an effect of the mean colour of the stimulus (H2). From the results one does indeed note that the thresholds measured at the same location but under different illumination conditions show a difference in volume (Figure 4.29), although the direction of the effect is not readily apparent – being different for the two locations. In other words, neither H1 or H2 is completely fulfilled. This effect is consistent for both uniform and textured patches. Instead, as proposed by Giesel et al. (2009), an analysis of the chromatic projections (Figures 4.32 and 4.33) shows a consistent trend, with the areas of chromatic ellipses being proportional to distances from the ambient illumination. The luminance projections (Figure 4.30), on the other hand, show a higher threshold for the fluorescent TL84 illuminant, irrespective of the location in colour space. In combination, the luminance and chromatic effects do indeed predict the trend in the results obtained for the volumes.

Furthermore, a comparison of the change in the gamut of the stimuli (Table 4.3 and Figure 4.28) and the change in the thresholds (volume: Figure 4.29, luminance: Figure 4.30, chromatic area: Figure 4.31) yields no consistently predictable effects. For example, the luminance range of the stimulus decreases when it is translated from location **D** to **T**, predicting the direction of change of the discrimination thresholds (which also decrease); but this predictability breaks down when one examines the change from **T** to **D**. Here, the change in the luminance range (an increase) does not predict the direction of change in the thresholds (a decrease). Thus, these results cannot be completely explained on the basis of the change in the range of the gamuts of the stimuli.

Irrespective of the swap, one observes that the textured patches always have a higher

discrimination volume compared to the uniform patches. This is evident from in Figure 4.35 which shows the ratio of the volumes of discrimination ellipsoids obtained for polychromatic stimuli to those for the corresponding uniform patches (the $CIELAB$ mean colour). A ratio greater than one demonstrates that the volume was always higher for polychromatic stimuli. This increase is consistent, both in the chromatic (Figure 4.36) and the luminance (Figure 4.37) projections. This indicates that the decrease in discriminability for textures compared to their means is independent of the illumination condition and the stimuli's exact location in colour space. These results, once again reinforce those by Hansen et al. (2008) who found a similar increase in thresholds for textures.

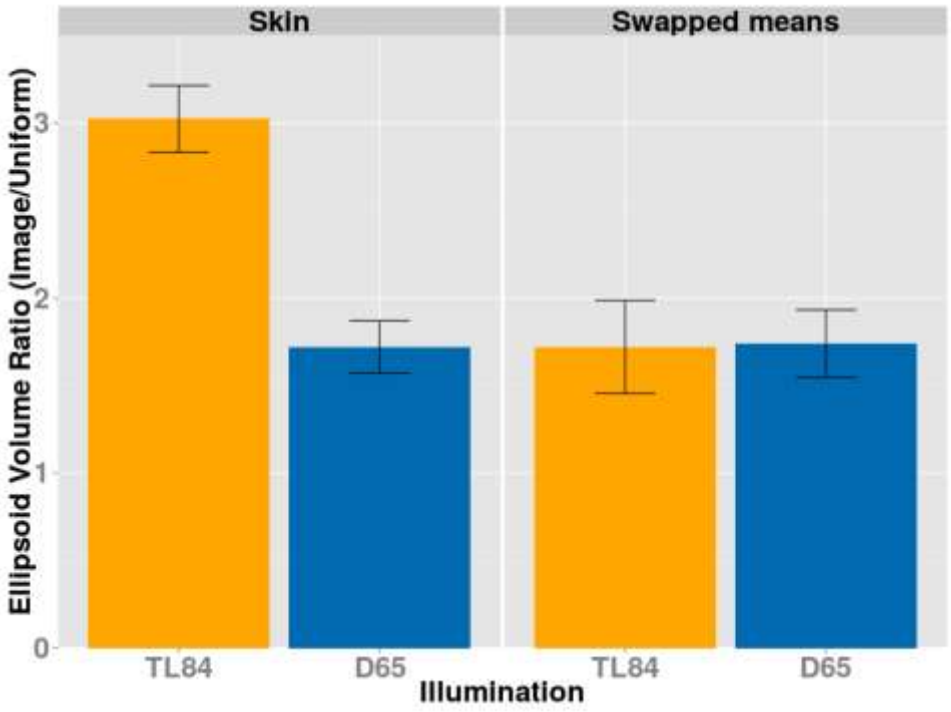


Figure 4.35: Ratio of discrimination volumes for the polychromatic and uniform stimuli

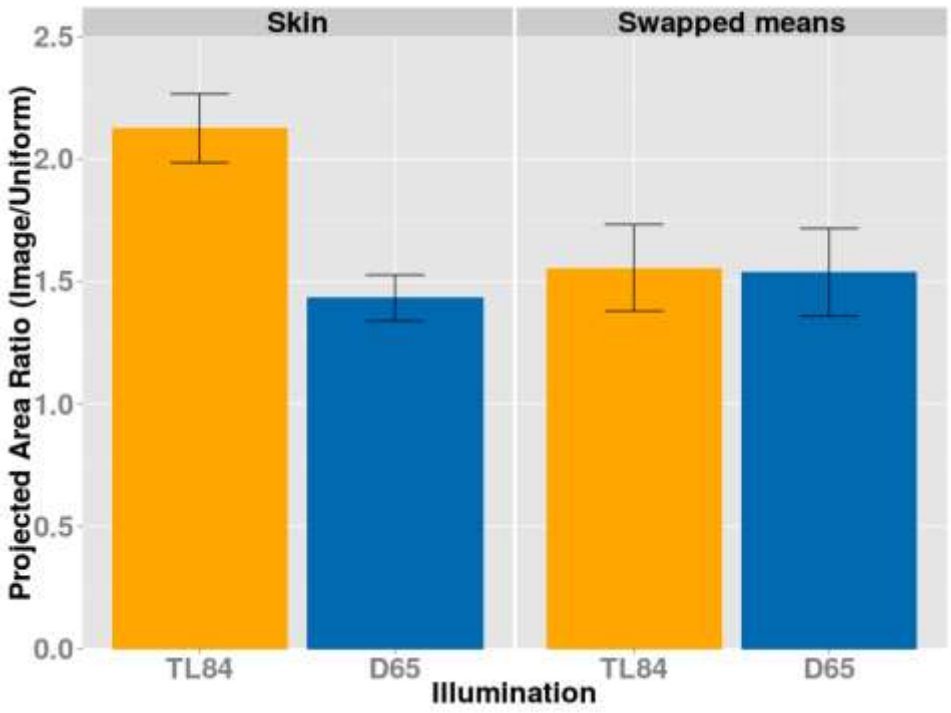


Figure 4.36: Ratio of area of chromatic ellipses for the polychromatic and uniform stimuli

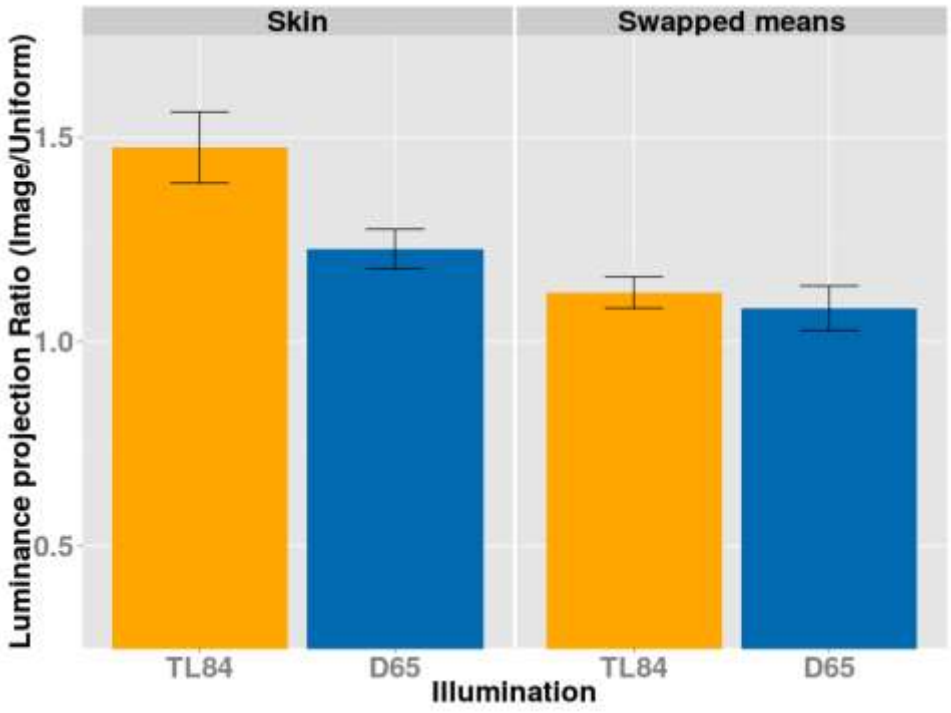


Figure 4.37: Ratio of luminance projections for the polychromatic and uniform stimuli

The fact that similar effects of ambient illumination and chromatic distance of the stimuli from the illuminant are seen for both textured and uniform stimuli, while maintaining larger thresholds for textures (Figure 4.38), suggests that the mechanisms governing these effects are perhaps dominated by the average colour signal. In the discussion to Experiment 1b (Section 4.4.4) it was proposed that this increase in threshold for textured stimuli compared to uniform stimuli could perhaps be a result of contrast adaptation mechanisms proposed by Webster & Mollon (1997). They suggest that these contrast adaptation mechanisms might operate by decorrelating the colour distribution of the stimuli (Atick et al., 1993; Webster & Mollon, 1991) using multiple chromatic channels tuned to different directions in colour space, and that the adaptation might emerge from inhibition between channels with the most correlated responses (Barlow & Földiák, 1989; Webster & Mollon, 1994). The results from this experiment offer further support for this possibility, proving that this increase of thresholds holds not only for ecologically valid simulations of skin (Experiment 1), but also for stimuli with similar *skin-like* relative colour distributions which are incongruent with the ambient illumination due to a shift in mean colour.

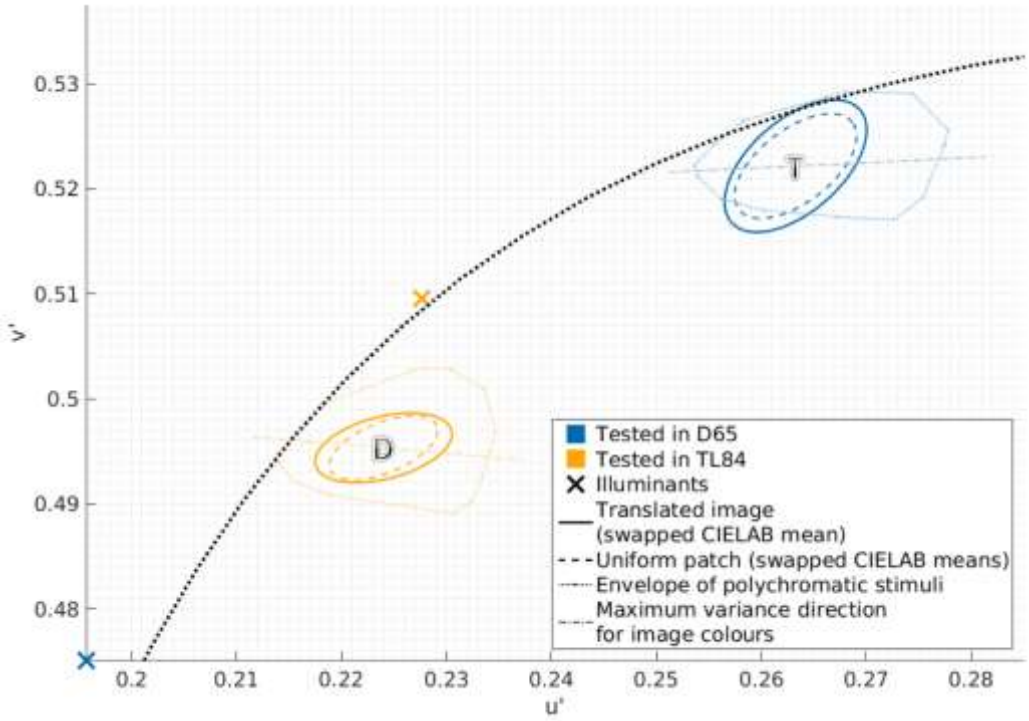


Figure 4.38: Average ellipses for Experiments 2a and 2b

4.6 Experiment 3: Measurements at ambient illuminant chromaticity

The experiments up to this point explore the effects of the illumination condition and the extent of adaptation on the discrimination volume in terms of luminance component and the area of the chromatic component. It has been reported by Hansen et al. (2008) and Giesel et al. (2009) that chromatic discrimination ellipses at the adaptation chromaticity are smaller in comparison to those further away. They also report that for polychromatic natural stimuli these ellipses are driven by the direction of the spread of the colour distribution. The next experiment was conducted in order to explore the discrimination thresholds at the ambient illuminant chromaticity in the light of these findings.

4.6.1 Reference stimuli and illumination conditions

The stimuli were generated in a manner analogous to Experiment 2 above, and the same two illumination conditions (D65 and TL84) were used. The *reference images* for the current experiment were generated by translating the mean chromaticity of the *reference images* from Experiment 1 (which, to reiterate, were ecologically accurate simulations derived from original skin reflectance spectra) to the chromaticity of the respective illuminant in $CIELAB$ space. The translation was only done in the chromaticity plane while holding the luminance ($CIELAB\ L^*$ coordinate) constant. Each patch was then tested under the illuminant used to simulate the original patch so that the relative distribution under each illuminant was the same in Experiment 1 and Experiment 3. The observers were also tested with uniform patches of the same mean $CIELAB$ colour as the *reference images*:

4.6.2 Experimental protocol

The experiment followed the exact same protocol as Experiments 1 and 2. All

participants ($N = 6$) who had also participated in both Experiment 1 and 2 were recruited. In total, 168 thresholds (2 illuminants × 2 stimuli types (uniform and textures) × 3 repetitions × 14 measurement directions) were measured per participant, amounting to about 6 hours of testing.

4.6.3 Results

The results in this section are presented along with results from Experiments 1 and. Only observers common to all the experiments ($N = 6$) are included. This section only presents the results for the current experiment, while comparisons with previous experiments are made in the discussion.

Figure 4.39 shows the volumes of the discrimination ellipsoids. There is a statistical effect of illumination condition ($F_{1,5} = 13.41, p = 0.0146$) with higher volumes for TL84 than D65. No statistical difference in the thresholds between textures and uniform patches ($p = 0.456$) is observed.

Figures 4.40 and 4.41 show that this effect is observable both in the luminance ($F_{1,5} = 8.757, p = 0.0315$) and the area of the chromatic ellipses ($F_{1,5} = 20.25, p = 0.0064$). Figure 4.42 shows the ellipses for both the uniform and the texture patches in the experiment. The envelopes and first principal component directions of the polychromatic patches are also provided.

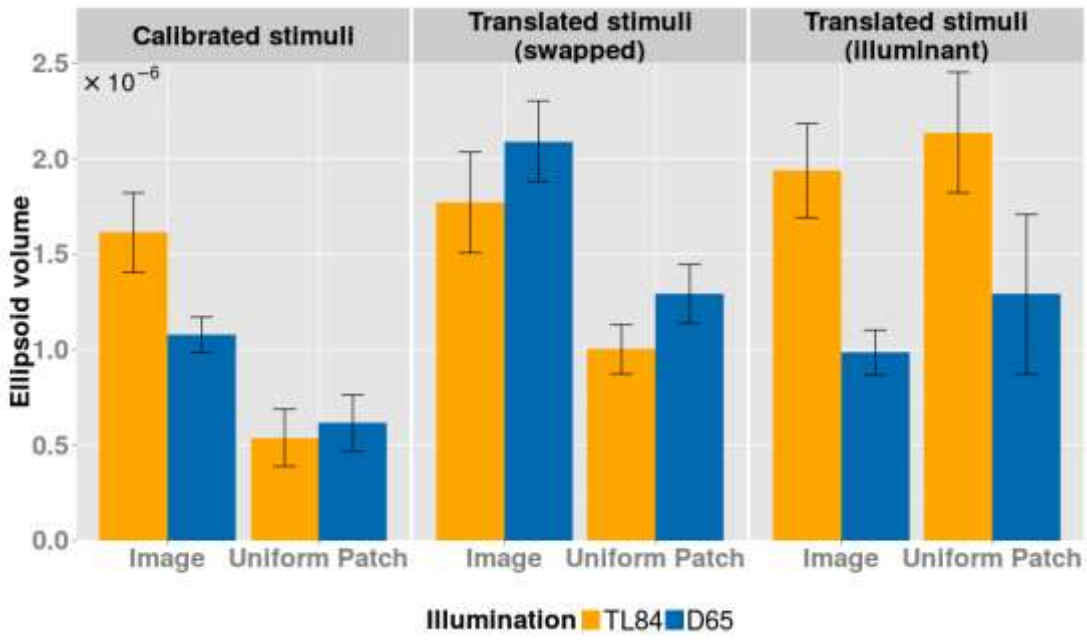


Figure 4.39: Ellipsoid volumes for participants common to all conditions. The data for the current experiment is shown in the rightmost panel. The left panel shows data for stimuli representing simulations of skin colour and texture under TL84 and D65 illuminants. Centre panel shows stimuli derived by swapping mean colour of skin simulations for the two illumination conditions while maintaining the relative colour distribution of the stimuli.

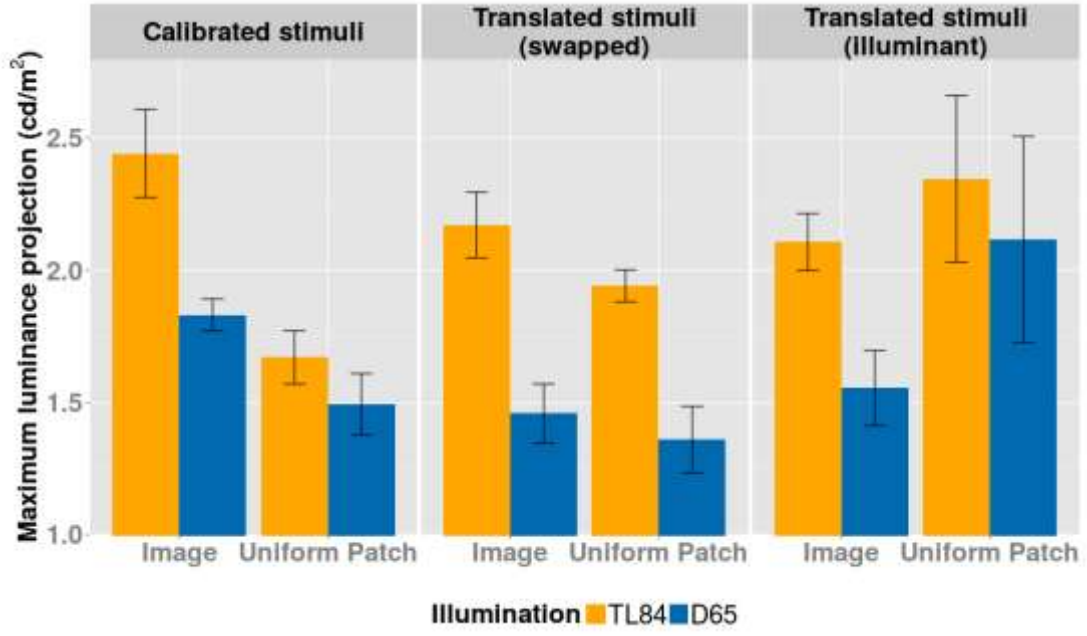


Figure 4.40: Luminance projections of ellipsoids for participants common to all conditions. The left, centre and right panels correspond to Experiment 1, 2 and 3 respectively.

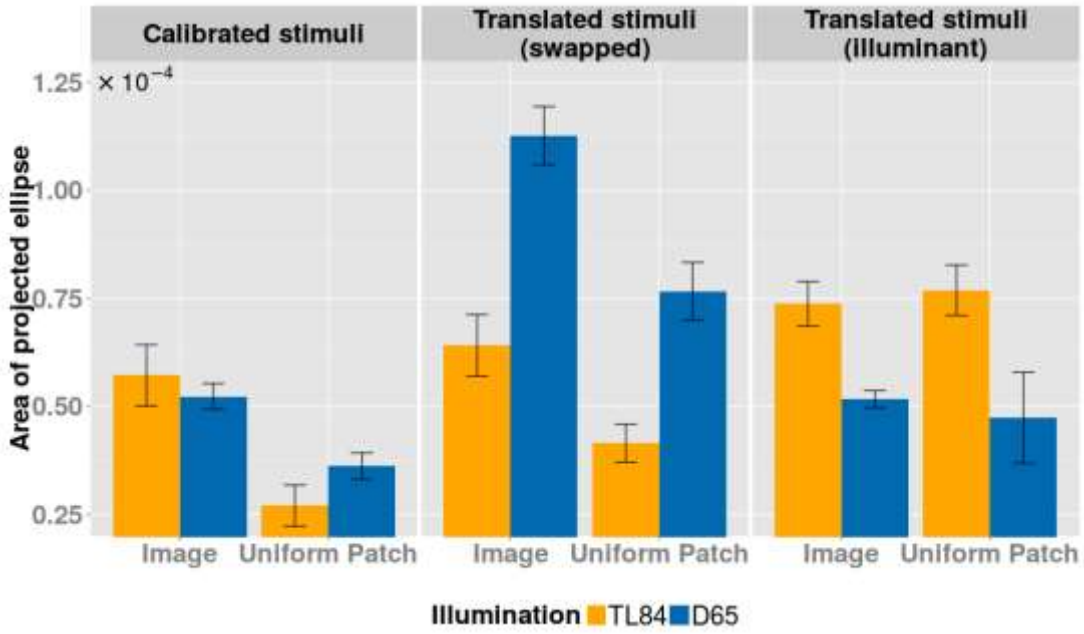


Figure 4.41: Area of chromatic projections of the ellipsoids for participants common to all conditions. The left, centre and right panels correspond to Experiment 1, 2 and 3 respectively.

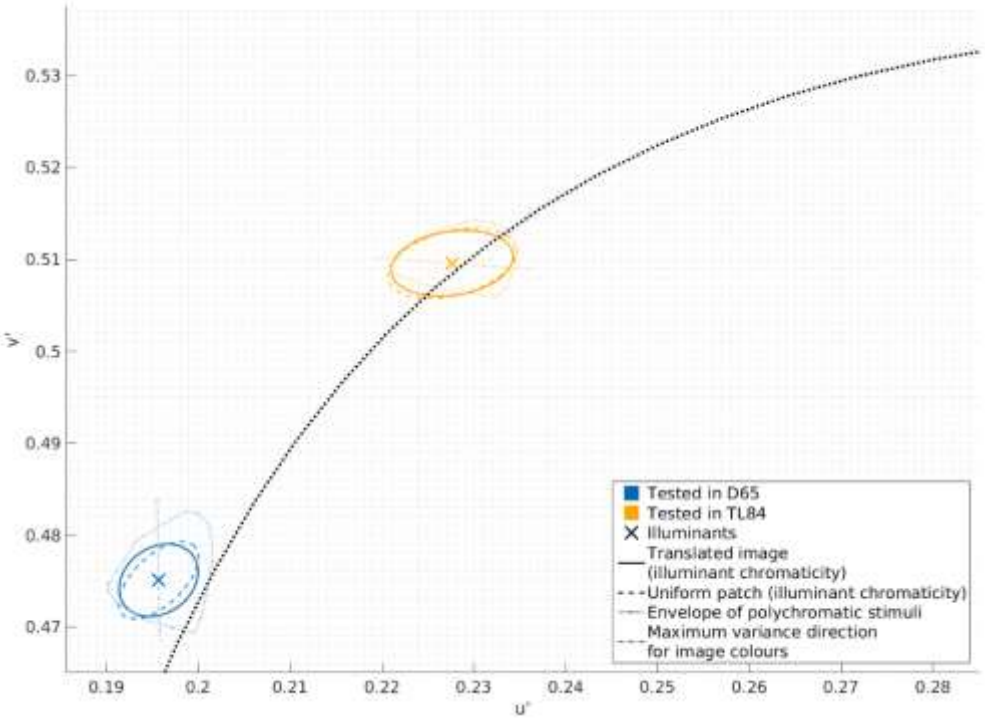


Figure 4.42: Average ellipses for Experiment 3. The ellipses for textures are drawn in solid lines while ellipses for uniform patches are in dashed lines.

The azimuths of these ellipses are shown below (Figure 4.43). Although one still sees the overall statistical effect ($F_{1,20} = 6.6340, p = 0.0181$ in the Harrison-Kanji test) of the TL84 ellipses being aligned to a greater extent with the u' axis than the D65 ellipses, the variability, especially for the textures is very high. A simple effects post-hoc analysis using the Watson-Williams test indeed shows that the effect only exists for uniform patches ($p = 0.0073$) and vanishes altogether for textures ($p = 0.2519$).

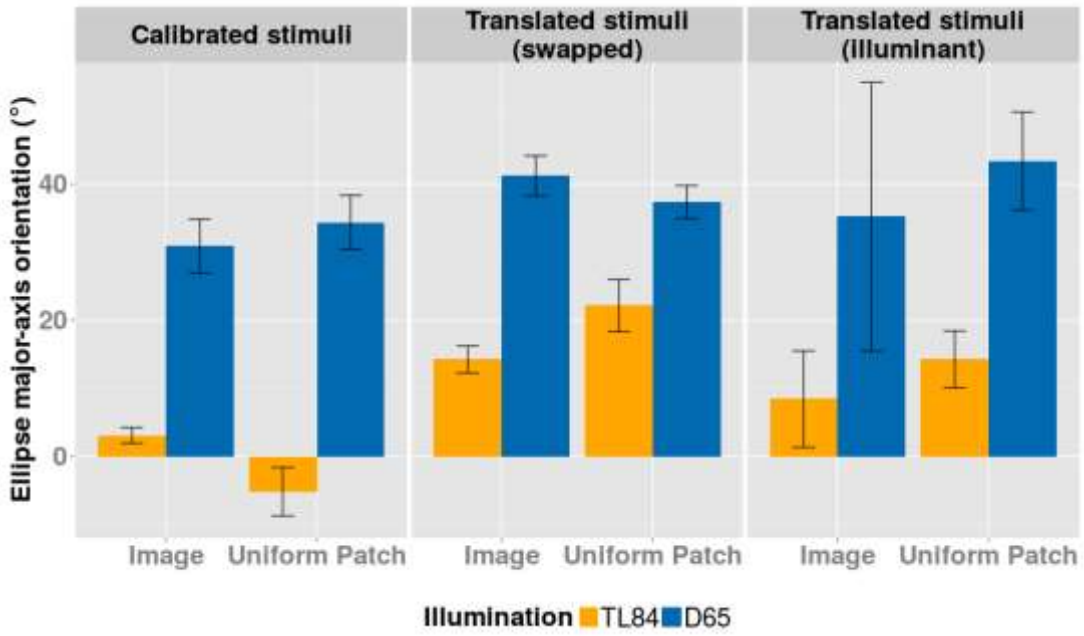


Figure 4.43: Azimuths of chromatic ellipse major axes for Experiment 3.

4.6.4 Discussion

In this set of experiments (textures and uniform patches), one observes that the ratios of the thresholds for images and uniform patches is much lower than those observed in Experiments 1 and 2, reaching values below 1 for some cases. Figure 4.44 shows the ratio of the discrimination volumes for polychromatic images to the volumes for the corresponding uniform patches. Similarly, Figure 4.45 shows the ratios (polychromatic stimuli vs. corresponding uniform patches) for the areas of the

projected discrimination ellipses, and Figure 4.46 shows the ratios for the luminance projections of the discrimination volumes.

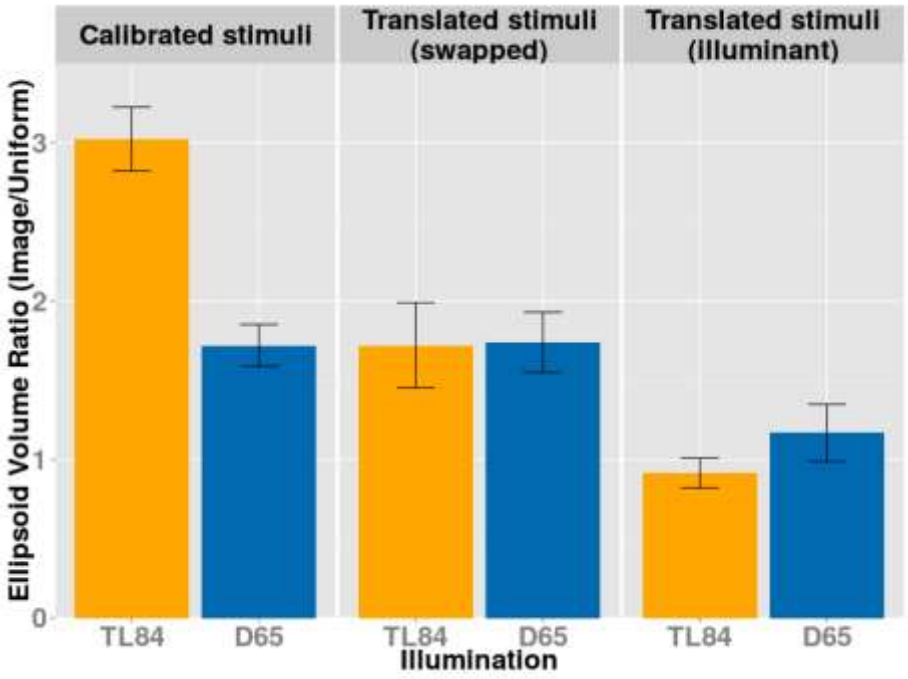


Figure 4.44: Ratio of discrimination volumes for the polychromatic and uniform stimuli

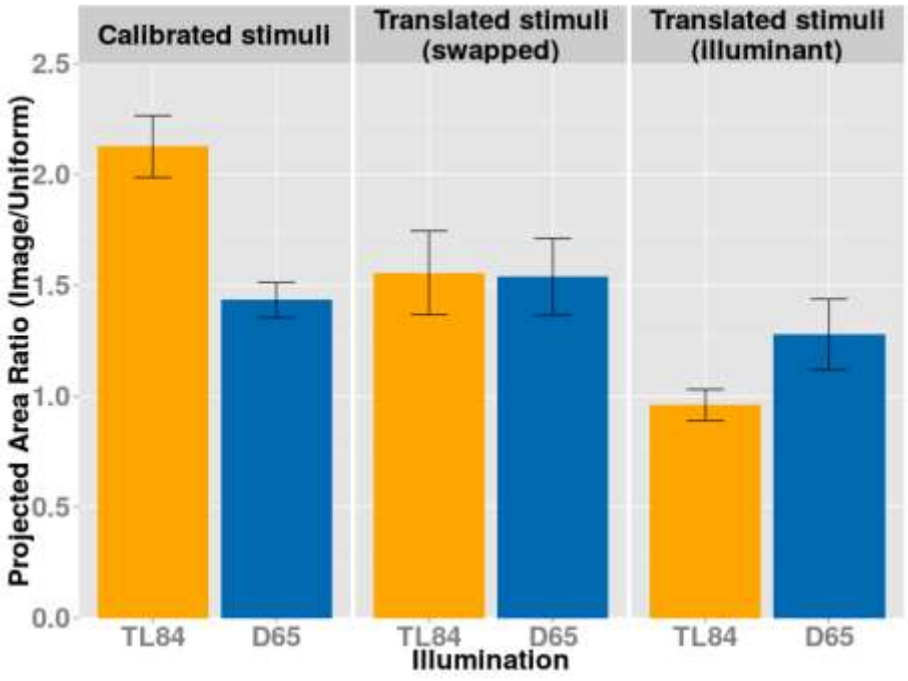


Figure 4.45: Ratio of area of chromatic ellipses for the polychromatic and uniform stimuli

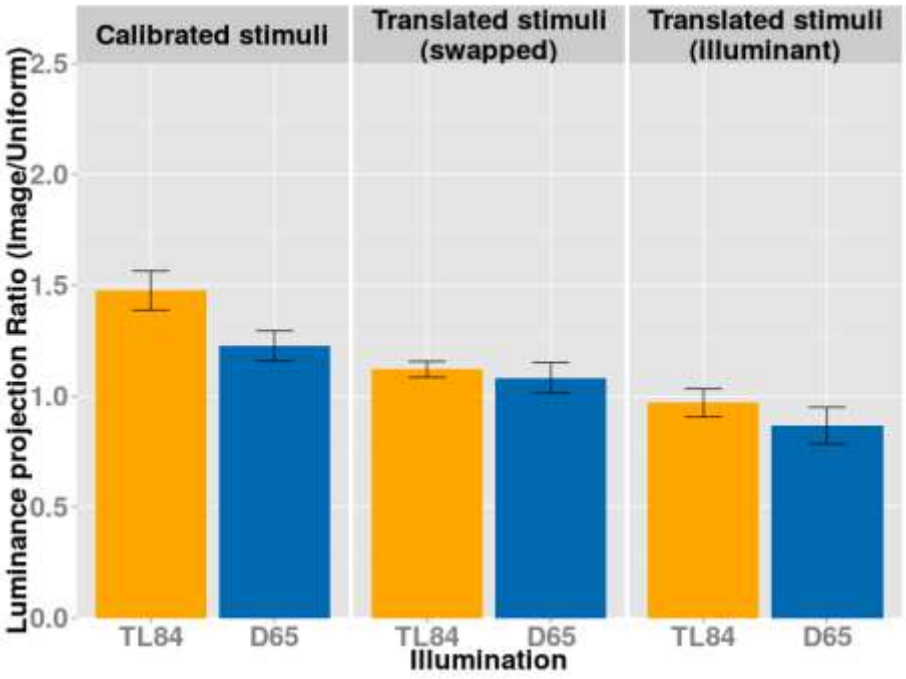


Figure 4.46: Ratio of luminance projections for the polychromatic and uniform stimuli

This lowering of the ratios is observed for ellipsoid volumes, luminance projections and chromatic projections, thereby suggesting that if the mean colour of the texture is approximately the same as the ambient illumination, the effect of the higher order statistical properties of the textural stimuli is diminished. It is also interesting to note that in Figure 4.40 the effect of larger thresholds being observed under fluorescent light than for simulated daylight does not hold for uniform patches, with very large variances in both lighting conditions.

The findings from Experiments 1 and 2 had strongly supported those reported by Hansen et al. (2008) and Giesel et al. (2009). In addition to the results for stimuli away from the adaptation point, they had also found that at the adaptation point the chromatic ellipses for uniform patches were smaller than those for textures. The results from Experiment 3 do not agree with this finding. To an extent, this can be explained by the difference in the experimental setup and procedure – first, subjects in the current study adapted to the ambient illumination while their subjects adapted to the surround-field illumination from the self-luminous screen background; second, the stimuli used in the

current study were displayed in object-mode through cardboard cut-outs with no self-luminous background while their stimuli were displayed on a screen with a self-luminous background (the adaptation point). These differences could suggest that the adaptation state in the current experiment was a result of a mixed adaptation to the ambient and the surround fields, with the observers adapting to an effective colour that is different from the ambient illumination (Henley & Fairchild, 2000; Katoh, Nakabayashi, Ito, & Ohno, 1998). It is unlikely that this is the case, as the grey cardboard which served as the surround had a relatively flat reflectance spectrum (Figure 4.47), with colour chromaticities in each ambient illumination condition falling very close to the source spectrum (Figure 4.48). The effect of this combination of ambient and surround illumination cannot produce an adaptation that reconciles the current results with those obtained by Hansen et al. (2008) and Giesel et al. (2009).

A close examination of Figures 4.42 and 4.43 reveals that the azimuths for the texture and uniform stimuli are very similar. Hansen et al. (2008) and Giesel et al. (2009) had concluded that the chromatic ellipses for textures measured at the adaptation chromaticity tend to align along the maximum variance direction of their colour distribution – a fact that is reflected in the average ellipses for the data from the current experiments, but with a large amount of variability. Data for individual observers suggest that this is not a robust effect (refer to individual observer plots in Appendix 2).

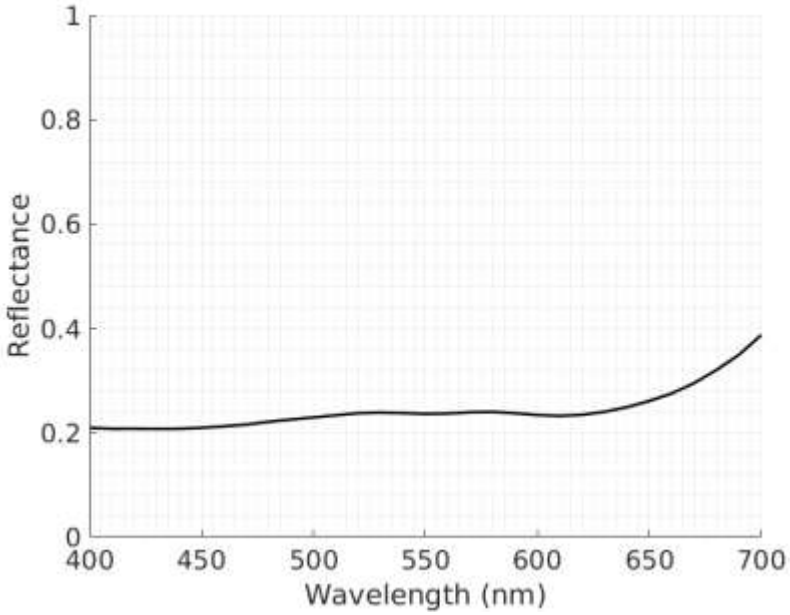


Figure 4.47: Reflectance spectrum of the grey cardboard

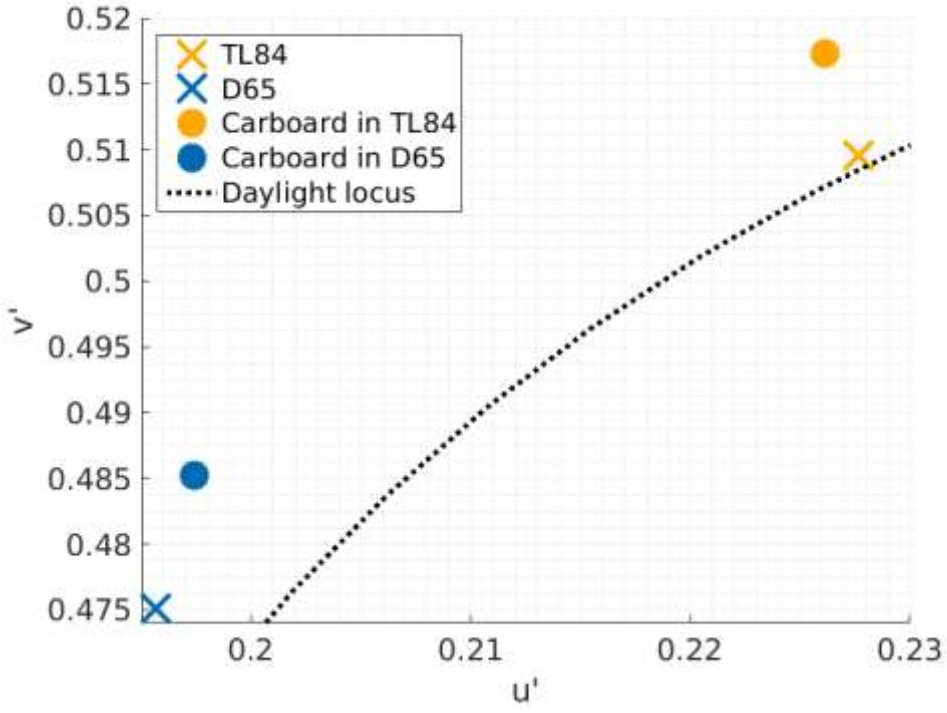


Figure 4.48: Chromaticities of the cardboard under D65 and TL84 illuminants

4.7 General discussion

This section reports results and analyses which span across the three experiments. Previous discussions have remarked on the higher thresholds of textures compared to uniform patches. In the discussion to Experiment 1b it was pointed out how this could be explained in terms of the idea of contrast adaptation proposed by Webster & Mollon (1994). However, it is interesting to note that this increase in thresholds for isolated skin patch images stands in stark contrast to findings by Tan & Stephen (2013) who found a decrease in the thresholds for facial images. They found no changes in luminance thresholds, and increased sensitivity along chromatic directions for images of complete faces as compared to uniform patches. This suggests, once again, that skin without shape cues, is not as discriminable as the complete face – even on flat displays. Furthermore, care must be taken not to interpret the estimates of skin discrimination thresholds as acceptability thresholds, which normally tend to be higher (Paravina, Majkic, Del Mar Perez, & Kiat-Amnuay, 2009).

One of the most interesting observations by Hansen et al. (2008) was that if the textured stimuli are sufficiently away from the adaptation point, the orientation of the isoluminant discrimination ellipses is dictated by the direction of adaptation, i.e., the direction of the vector connecting the illuminant and the stimulus chromaticities. While the current set of studies does indeed confirm this for the ecologically valid skin stimuli in Experiment 1a (Figures 4.13 and 4.14), the trend is absent for textures in Experiment 2 (Figure 4.32) which resemble actual skin simulations only in second and higher order statistic, but not in the mean colour. This might further indicate that the reflectance of natural stimuli and scenes (and consequently their appearance) introduces certain additional constraints which make estimation of the illuminant properties easier (Golz & MacLeod, 2002; Philipona & Regan, 2006). As noted earlier, the adaptation to natural textures could be explained in terms of contrast adaptations to colour distributions of the stimuli by the decorrelation of signals from multiple chromatic channels with a large array of colour tunings (Atick et al., 1993; Webster & Mollon, 1994, 1997).

Twer & MacLeod (2001), in their *pleistochrome* model, suggest that such tunings could possibly be represented by split coding of non-linear responses found in cells in the macaque parvocellular pathway (B. Lee, Pokorny, Martin, Valbergt, & Smith, 1990). They propose a narrow operating range for these parvocellular responses which is determined by the distribution of colour in natural polychromatic stimuli. Goda, Koida, & Komatsu (2009) also report optimal response functions closely following those predicted by the *pleistochrome* model in populations of LGN neurones corresponding to colour statistics from natural scenes. Interestingly, Twer & MacLeod (2001) also point out that the non-linearity in the responses of cells from the magnocellular pathway (B. Lee et al., 1990) fitted using the same *pleistochrome* model is much more saturated, indicating that while the parvocellular pathway is optimised to minimise perceptual error in the detection of the colour of natural stimuli, the magnocellular pathway is optimised for the detection of boundaries (hence the steeper non-linearity).

Next, let us examine the proposition that the area of the chromatic discrimination ellipse increases with increasing distance from the illuminant chromaticity. Figure 4.49 shows the area of the chromatic ellipses plotted against the distance of the stimuli from the scene illuminant. Of course, since the current set of experiments were not originally designed to investigate the effect of the adaptation, the sampling of the points is not the same for all conditions. Even so, the curves indicate that the minima of the area of the chromatic ellipses do not coincide with the illuminant chromaticity ($x = 0$). This is true for both the textures and the uniform patches. It is also interesting to note that for uniform patches, the inter-observer variability decreases close to these minima. For the textures, the minima are close to the $\pm 3\sigma$ range of the major axis of a Gaussian distribution fitted to the colour distribution of the textures, except for observers MH, CN and CT in the fluorescent lighting condition.

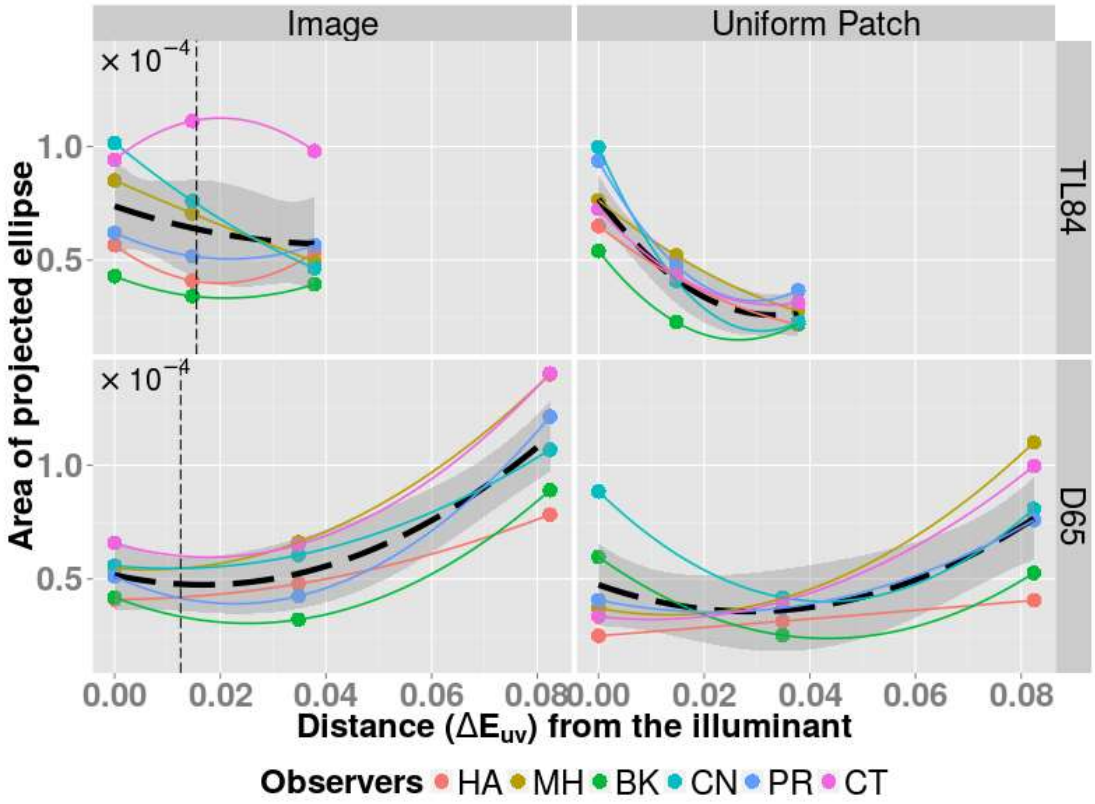


Figure 4.49: Area of the chromaticity ellipses as a function of distance from the illuminant chromaticity. Each point represents an average measurement over three repetitions for each of the six observers. In each condition, observer-wise fits using a quadratic function are also plotted. The graph also contains an average quadratic fit (thicker black line) with 95% confidence regions. A vertical dashed-line denotes the $\pm 3\sigma$ range of the major axis of a Gaussian distribution fitted to the colour distribution of the textures.

If one makes similar plots for the luminance (Figure 4.50), most of the average fitted quadratic curves show a very low curvature, almost resolving into straight lines. We also notice that the luminance thresholds are indeed higher under fluorescent TL84 lighting than for artificial daylight for most observers. This is true both for textures and uniform patches. In general, the inter-observer variance for luminance is much lesser than for the area of the chromatic ellipses.

As a result of these variations for luminance and chromatic projections, the volumes of the discrimination ellipsoids show a slight curvature which is more readily identifiable for uniform patches than for textures (Figure 4.51).

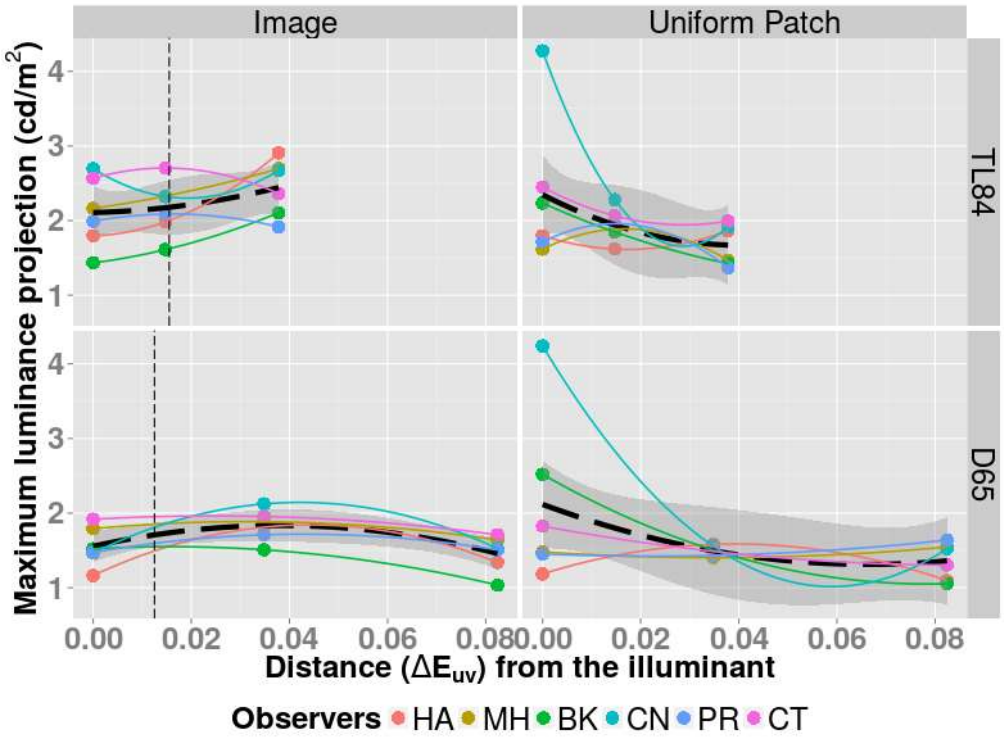


Figure 4.50: Luminance projections of discrimination ellipsoids as a function of distance from the illuminant chromaticity. The symbols and lines represent the same aesthetics as in Figure 4.49.

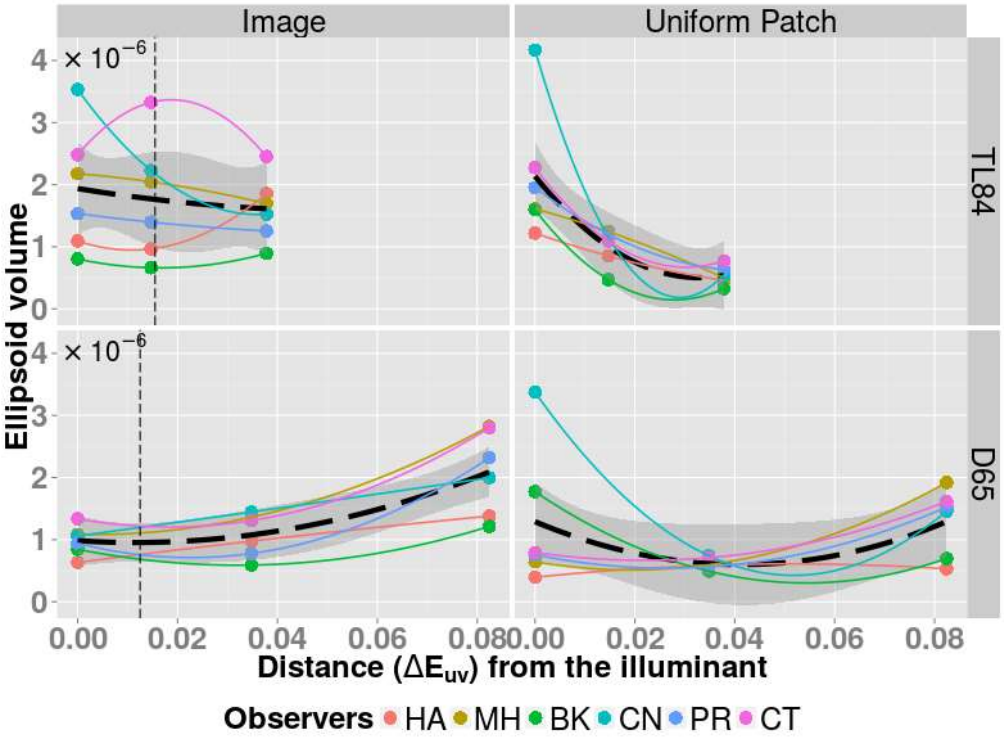


Figure 4.51: Luminance projections of discrimination ellipsoids as a function of distance from the illuminant chromaticity. The symbols and lines represent the same aesthetics as in Figure 4.49.

4.8 Conclusion

Experiment 1a estimated the discrimination thresholds for simulated skin patches from two ethnicities, Caucasian and Oriental, in 3-D colour space. A large amount of data using 18 participants was collected and these estimations may be used to define colour difference metrics for skin specific stimuli, and to design acceptability criteria for skin prosthetics.

The subsequent experiments investigated the effect of illuminant and relative adaptation (distance and direction from the ambient illumination chromaticity) on the discrimination thresholds for textures and uniform patches. It was found that, for both textures and uniform stimuli, the discrimination was worse in fluorescent lighting compared to simulated daylight. It was also observed that the area of the projected chromatic ellipses first decreased and then increased with distance from the illuminant chromaticity. Furthermore, the orientation of the chromatic ellipses for texture stimuli away from the illuminant chromaticity were found to be influenced by the direction of the relative adaptation. This effect was more pronounced for ecologically valid simulations of skin as compared to textures with the same relative colour distributions but different means. At the illuminant chromaticity, the average chromatic ellipses for textures were aligned along the axis of maximum colour variation of the stimulus, albeit with a large inter-observer variability.

Chapter 5

Scope and contribution

The aim of the research presented in this thesis was to contribute towards the development of an appearance model for natural polychromatic stimuli. This was done in two steps. The first part of the thesis critically examines two important pseudo-invariants which form the basis of most colour appearance models – the achromatic locus (Chapter 2) and the unique hues (Chapter 3). The second part (Chapter 4) presents measurements of discrimination thresholds for an important natural texture – human skin, and investigates how these discrimination thresholds change depending on the illumination condition and properties of the colour distribution of the texture.

Chapter 2 reports measurements of achromatic settings made by 30 observers in three ambient illumination conditions – dark, simulated daylight and fluorescent lighting. The novel contribution of this chapter is a method of navigation through colour space which is based on unique hue directions. The results of this chapter show that compared to the conventional methods of colour adjustment which often employ movement along cardinal axes of the colour space, this novel unique-hues based navigation scheme results in a reduction in the variability of achromatic settings made by observers. Furthermore, the results also show that irrespective of the navigation scheme, the achromatic settings made by observers tend to closely lie along the daylight locus.

Chapter 3 is devoted to the modelling of chromatic adaptation using a large dataset of unique hue settings (Wuerger, 2013; Xiao et al., 2011) made by 185 observers in three illumination conditions – dark, simulated daylight and fluorescent lighting. The results show that a simple von Kries model operating on cone absorptions is the most efficient model in terms of the information theoretic Akaike criteria. No neural substrates have been reported for unique hues so far (Conway & Stoughton, 2009; Mollon, 2009;

Stoughton & Conway, 2008). While the analysis in this chapter does not go so far as to propose a novel mechanism, the results support the possibility that the mechanisms for UR and UG are non-linear in terms of cone absorptions. Another contribution of this chapter is the analysis of the dataset itself. Although this dataset has previously been used to model age related effects (Wuerger, 2013) and to evaluate the performance of the CIECAM02 colour appearance model (Xiao et al., 2013, 2015, 2011), no extensive analysis of the data has yet been reported. The analysis in this chapter shows that the distribution of hue angles of the unique hues is mostly unimodal (except the unique green in fluorescent lighting), and that the hue angles vary with the luminance of the settings.

In addition to these two chapters, work was also done on two related studies. The first of these (Xiao et al., 2015) investigates the performance of CIECAM02 colour appearance model on NCS unique hue data. Significant differences were found in the UY and UB lines. The contributions to this paper were in terms of statistical analysis, data interpretation and manuscript preparation. The second study (Makin & Chauhan, 2014) explores, in part, the performance of the Common Rate Controller hypothesis on the UY-UB line. The findings from this paper indicate that the same rate control mechanism which is used to estimate the position of an occluded moving target might be used to track the rate of colour change along the UY-UB line. Contributions made to this paper were in the form of experimental design and setup, and partial manuscript preparation. These two studies are not included in this thesis and have only been cited as references.

The second, and main part of the thesis is a long three-part chapter (Chapter 4) which extends the idea of discrimination measurements to natural skin textures. Each of the three sub-chapters presents data from three related experiments. The first experiment reports thresholds in three illumination conditions for simulated skin images and corresponding mean uniform patches. The thresholds for the polychromatic stimuli are consistently higher, for both the chromatic, and the luminance projections. The results from this experiment can be used to optimise line-elements for describing skin texture differences in colour-space. The second and third experiments investigate how these

discrimination thresholds vary with the illumination condition and the mean location of the texture in colour space. The results show that the projected chromatic discrimination ellipses increase in area as a function of the distance between the illuminant chromaticity and the mean colour of the texture. A direct and practical use of these results can be in lighting design for environments which have specific (more often than not, maximal) skin discrimination requirements such as dermatology clinics, prosthetics colouring or printing labs, and workspaces for makeup artists.

There were two additional outputs from the methods developed for the series of experiments described in Chapter 4. First, the QUEST adaptive procedure efficiency was optimised and a solution for the optimisation parameter (Eq. 4.5) was derived. Second, as described in Section 4.3.1, contributions were also made to the development of an algorithm for reconstructing reflectance spectra for skin from three-channel colorimetric measurements (Xiao et al., 2014).

Altogether, this thesis provides a critical examination of two fundamental theoretical underpinnings of modern colour appearance models – the achromatic locus and the unique hues, and introduces results which can be used to extend these models to natural polychromatic textures, in particular, human skin. Chapter 2 has been published in the Journal of Vision (Chauhan et al., 2014). Chapter 3 and Chapter 4 are currently being prepared for submission.

Appendix 1.

Colour spaces

This section gives a brief mathematical description of the forward transformations for the $CIELAB$ and the $CIE\ 1976\ UCS$ colour spaces. $CIELAB$ is a 3-D luminance-chromaticity space that uses a reference white to normalise its coordinates, while the $CIE\ 1976\ UCS$ is an absolute chromaticity-only space (i.e., a 2-D plane) which does not employ any normalisations. Usually the $CIE\ 1976\ UCS$ chromaticity coordinates are supplemented by the luminance value to fully define the stimulus. Throughout this section, $\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ and $\boldsymbol{X^W} = \begin{bmatrix} X^W \\ Y^W \\ Z^W \end{bmatrix}$ denote the $CIE\ 1931\ XYZ$ tristimulus values (Eq. 1.1) of a uniform colour stimulus and a reference white respectively.

Appendix 1.1 CIELAB

The $CIELAB$ space is characterised by a cube-root non-linearity along each normalised channel. It describes colour appearance using three coordinates – L^*, a^* and b^*. L^* denotes the lightness (relative brightness with respect to the reference white) of the stimulus, while a^* and b^* denote the red-green and the yellow-blue opponency channels respectively. The equations for the three coordinates are given by

$$L^* = 116 f\left(\frac{Y}{Y^W}\right) - 16$$

$$a^* = 500\left[f\left(\frac{X}{X^W}\right) - f\left(\frac{Y}{Y^W}\right)\right] \qquad \text{Eq. A1.1}$$

$$b^* = 200\left[f\left(\frac{Y}{Y^W}\right) - f\left(\frac{Z}{Z^W}\right)\right]$$

Here, the function f represents the cubic non-linearity and is given by

$$f(x) = \begin{cases} \sqrt[3]{t}\,, & t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29}\,, & \text{otherwise} \end{cases} \qquad \text{Eq. A1.2}$$

For a given reference $\boldsymbol{X^W}$ the transformation is bijective, and thus, fully invertible.

$CIELAB$ also has an associated colour difference metric usually denoted by ΔE_{LAB}, which is simply the Euclidean distance between two $CIELAB$ defined colours. For two colours $\boldsymbol{L^1}$ and $\boldsymbol{L^2}$ defined in $CIELAB$ space, this is given by

$$\Delta E_{LAB} = \left\| \boldsymbol{L^1} - \boldsymbol{L^2} \right\|_2 \qquad \text{Eq. A1.3}$$

Appendix 1.2 CIE 1976 UCS

CIE (2004) describes the $CIE\ 1976\ UCS$ as an alternative to the $CIE\ 1931\ XYZ$ derived xy chromaticities with enhanced perceptual uniformity. It is an absolute (not normalised to a reference) chromaticity plane derived by a projective transformation of the $CIE\ 1931\ XYZ$ tristimulus coordinates. The two chromaticity coordinates are conventionally denoted by the letters u' and v'. For a given stimulus $\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$, they are given by

$$u' = \frac{4X}{X + 15Y + 3Z}$$
$$v' = \frac{9Y}{X + 15Y + 3Z} \qquad \text{Eq. A1.4}$$

Due to the projective nature of the transform it is not invertible and thus does not describe the stimulus fully. To complete the stimulus description and to make the

transform bijective, the CIE 1976 UCS chromaticities are often quoted alongside the absolute luminance of the stimulus, which is simply the Y coordinate of the CIE 1931 XYZ tristimulus values.

It is to be noted that the CIE 1976 UCS space is distinct from the $CIELUV$ space (also described in CIE, 2004), which is a relative space, akin to $CIELAB$. The absolute and yet perceptually relevant nature of CIE 1976 UCS makes it an ideal choice for comparing colours spanning across a set of illumination conditions.

Appendix 2.

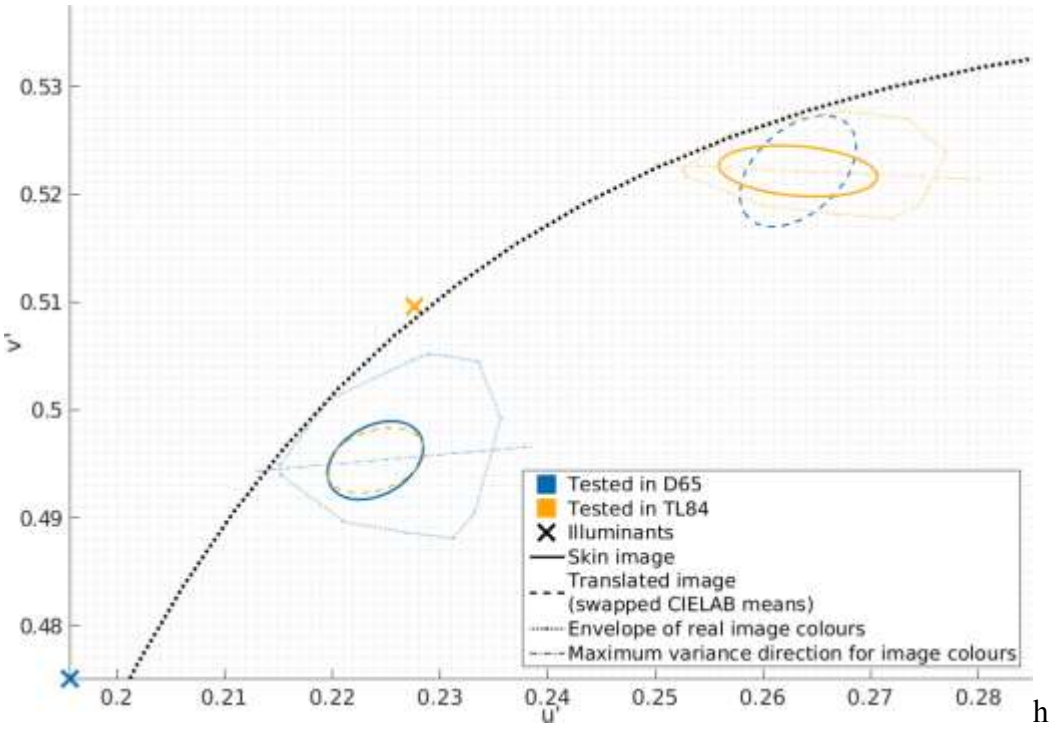
Projected ellipses for individual observers

This appendix provides individual observer chromatic ellipses for textures and uniform stimuli. Only observers who participated in all the experiments ($N = 6$) are shown. In Appendix 1.1, Experiment 1 (simulated skin) and Experiment 2 (swapped stimuli) are compared. Two plots are shown for each observer – the first plot shows results for polychromatic stimuli while the second plot shows results for the uniform mean $CIELAB$ patches. Appendix 1.2 provides a similar comparison between Experiment 1 and Experiment 3 (illuminant-chromaticity centred stimuli).

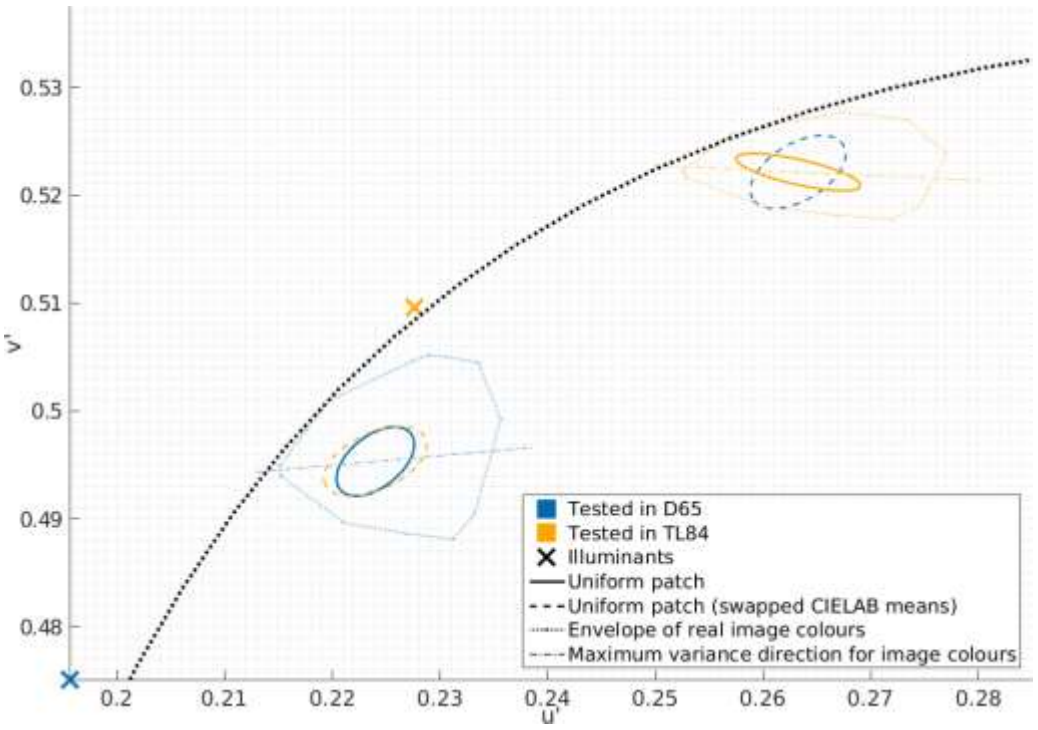
Appendix 2.1

Experiment 1 (simulated skin)
vs.
Experiment 2 (swapped means)

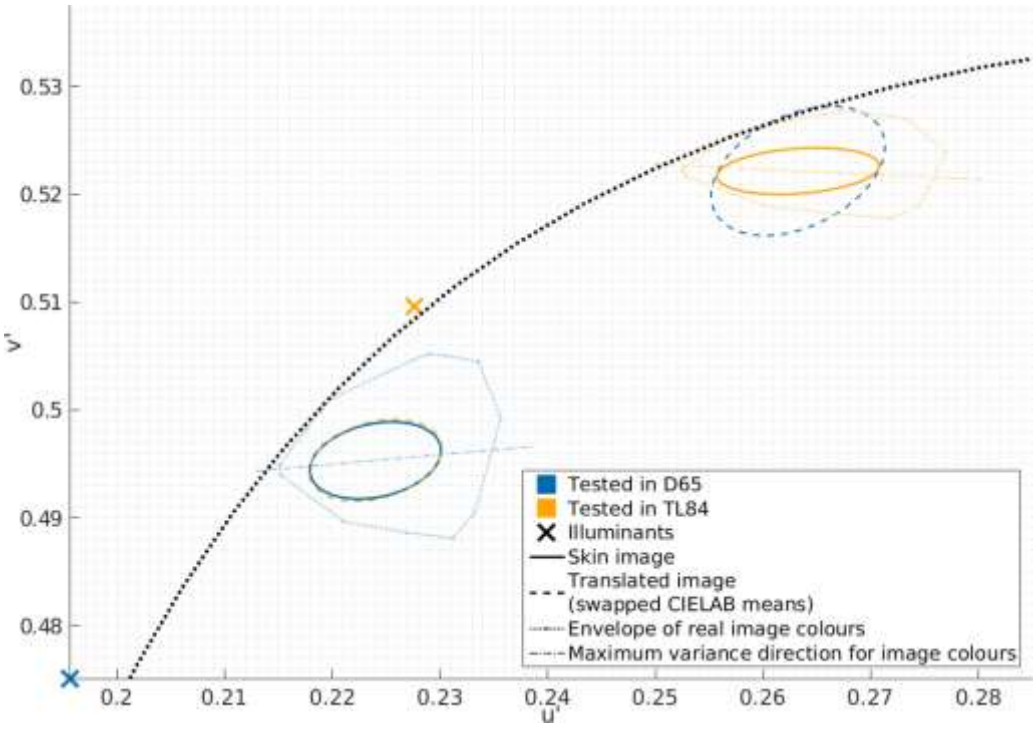
Observer: HA

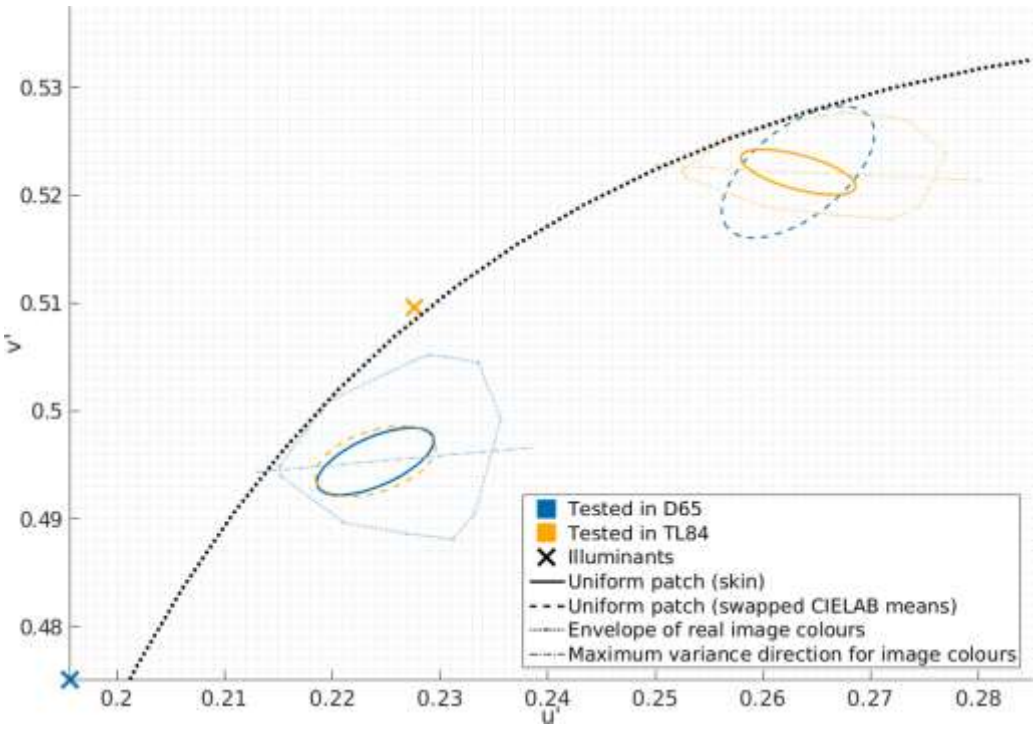


h

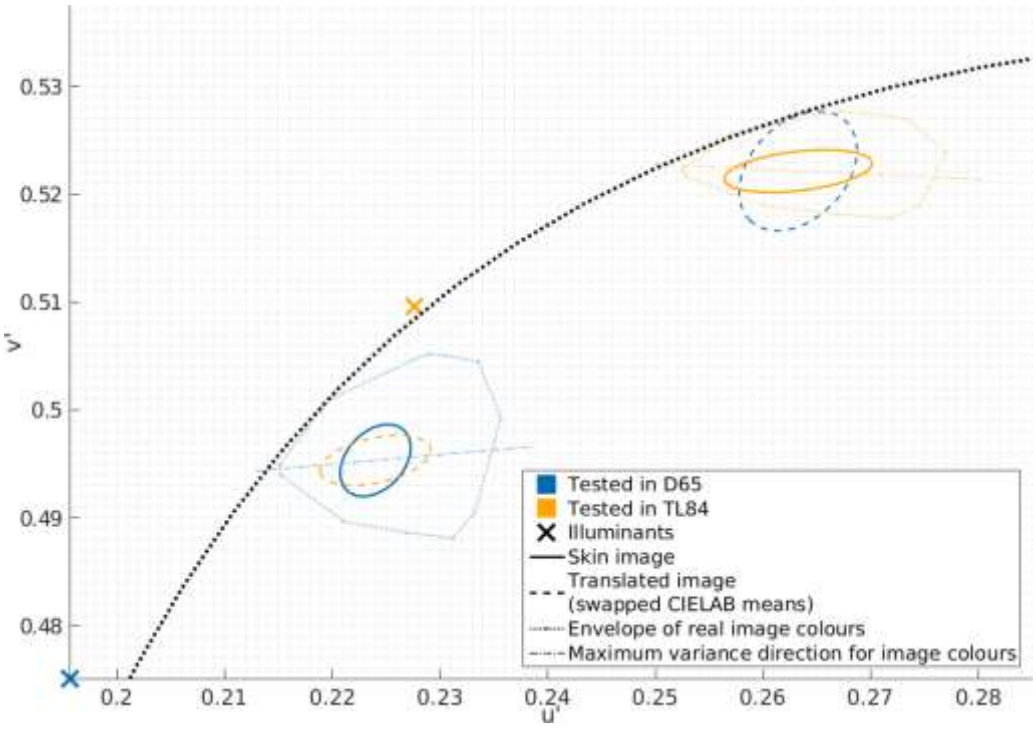


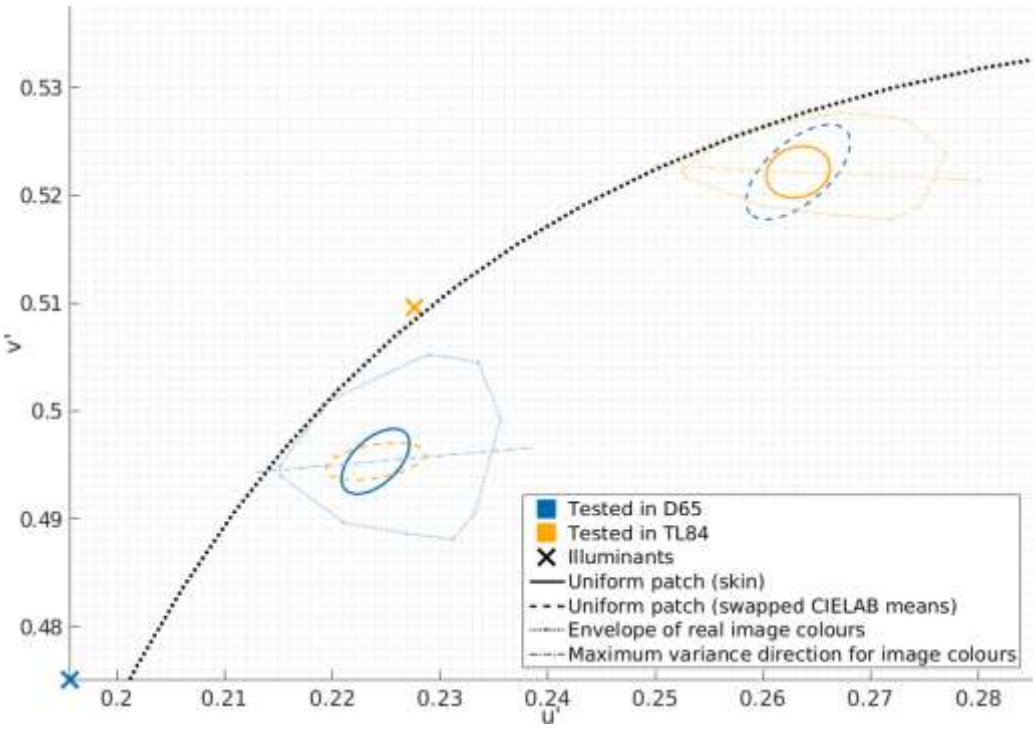
Observer: MH


0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

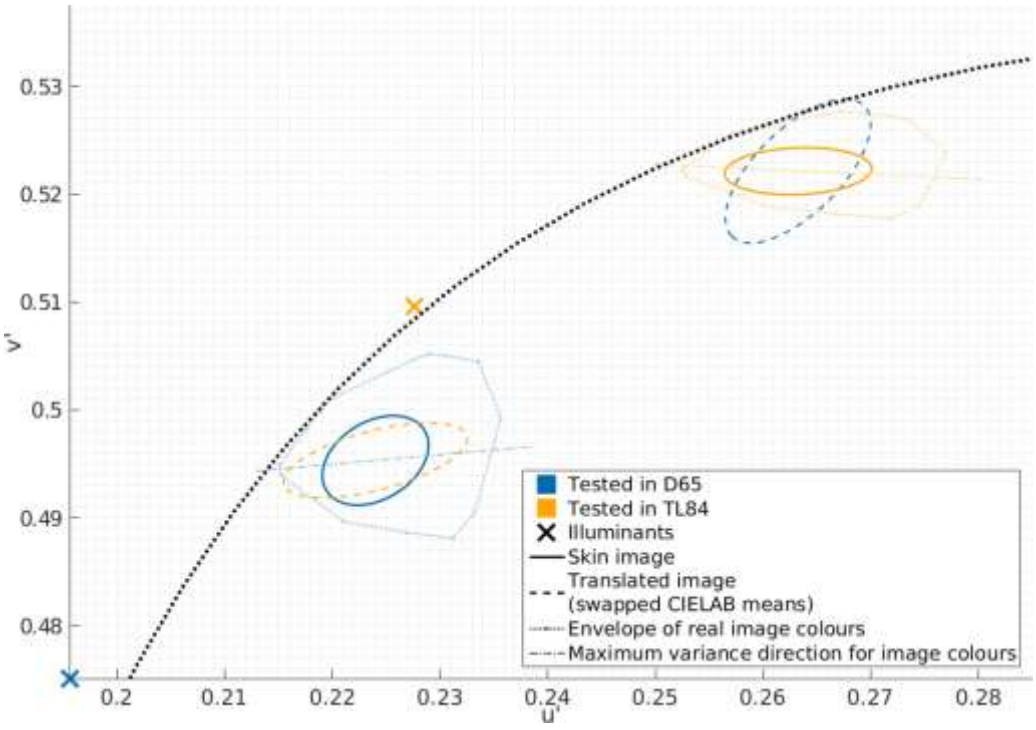

0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

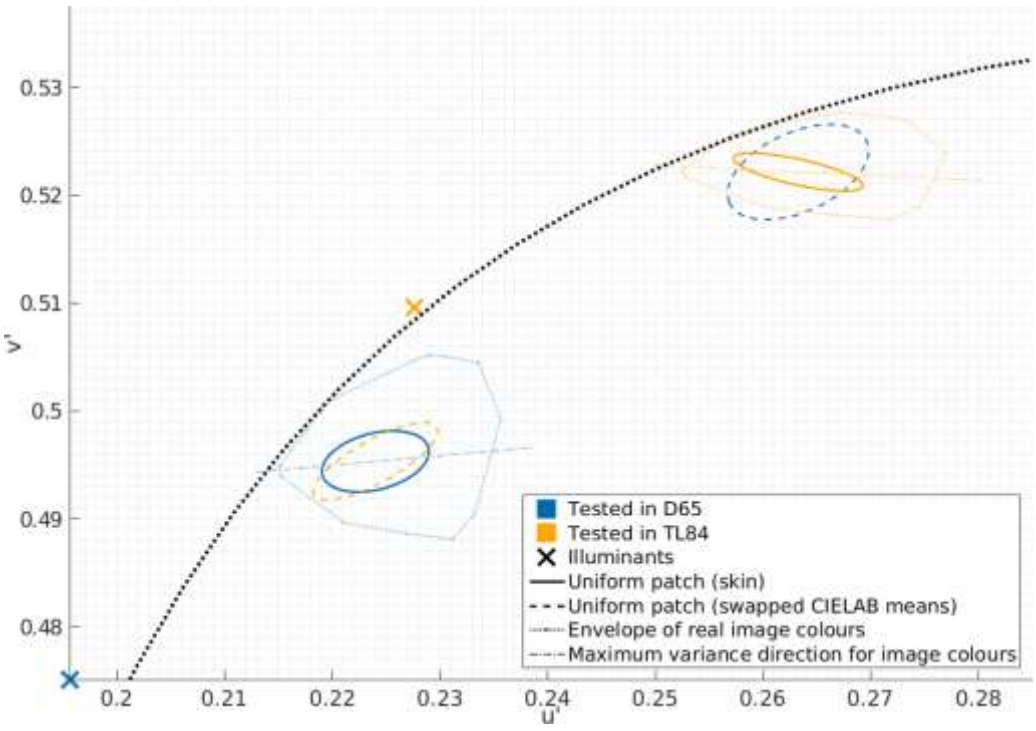
Observer: BK


0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

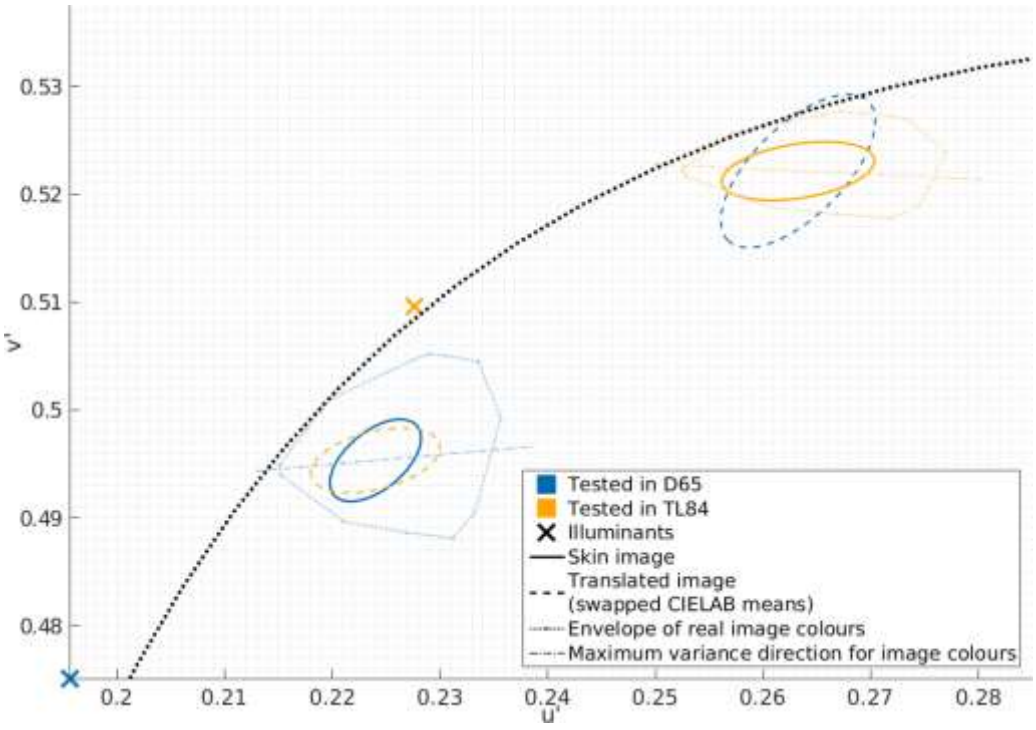

0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

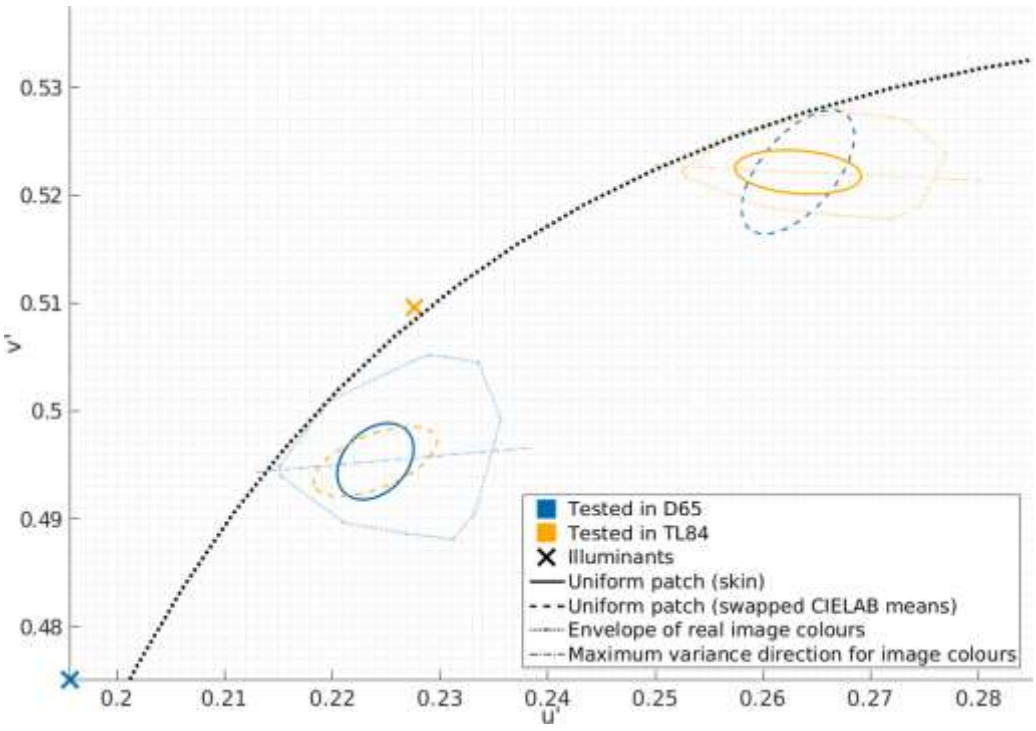
Observer: CN


0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

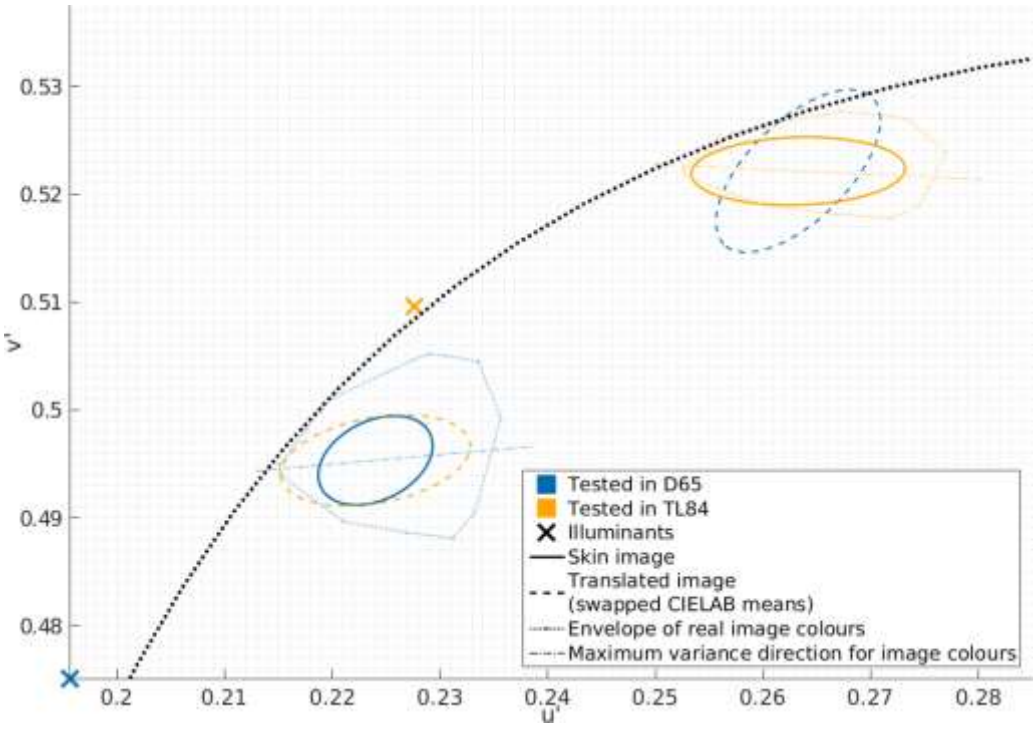

0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

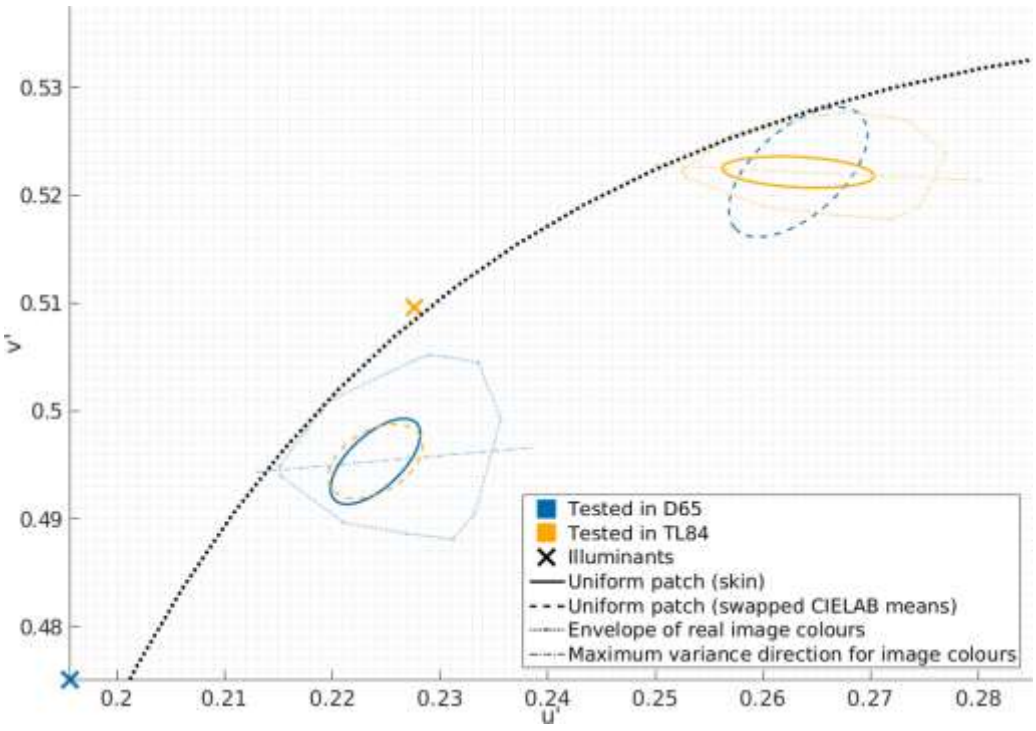
Observer: PR


0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours


0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

Observer: CT

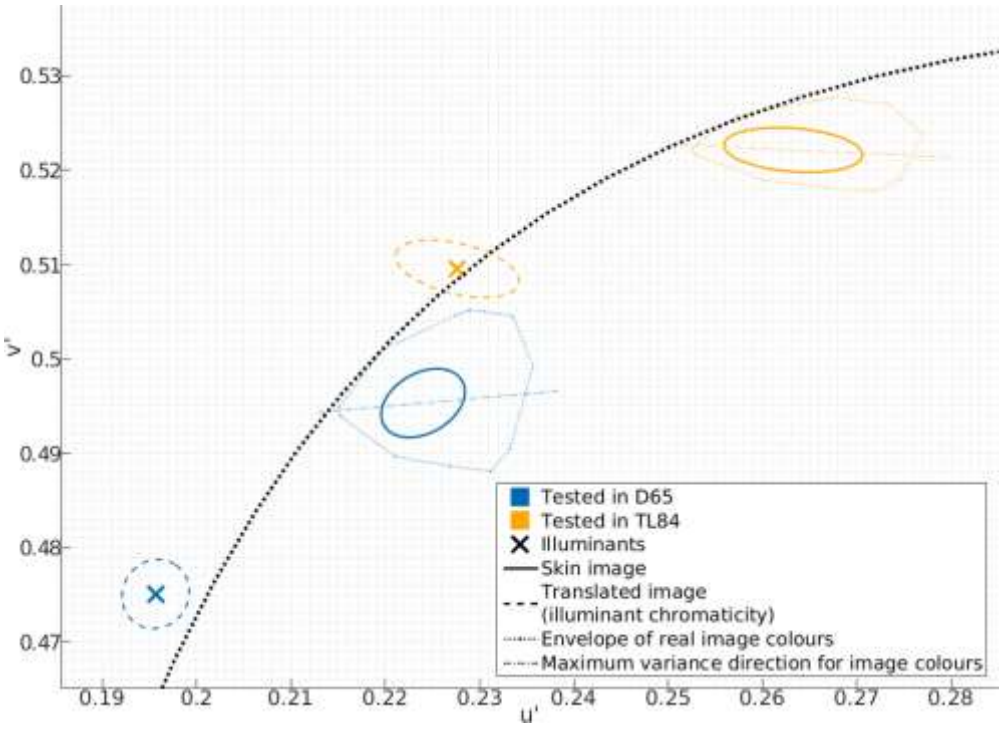

0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

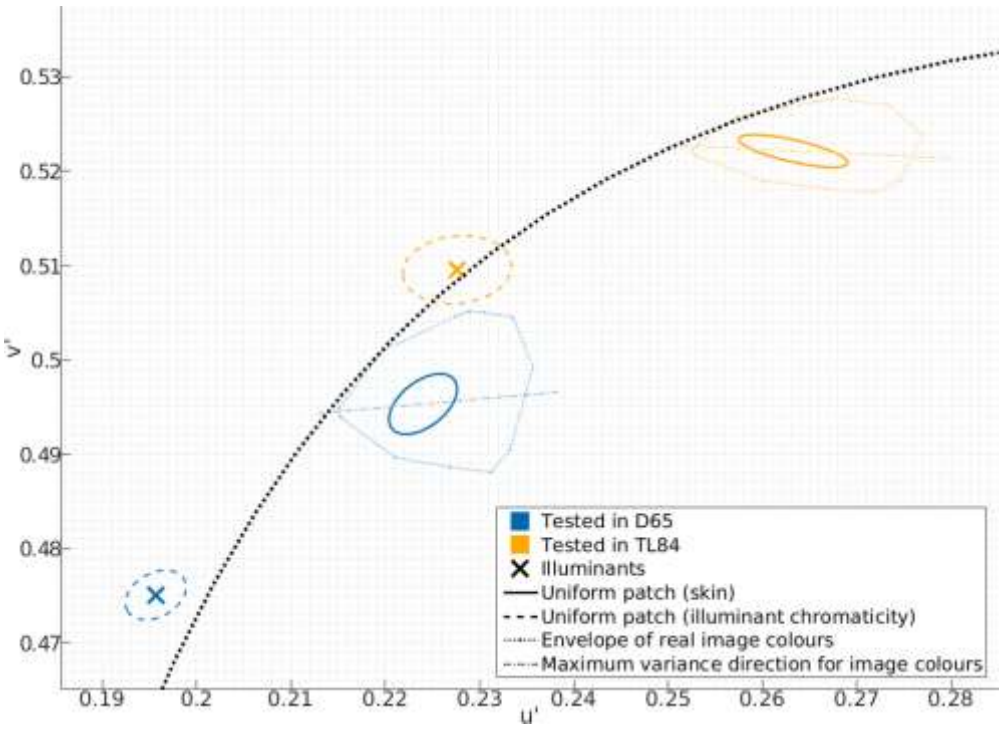

0.53
0.52
0.51
0.5
0.49
0.48
v'
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (swapped CIELAB means)
Envelope of real image colours
Maximum variance direction for image colours

Appendix 2.2

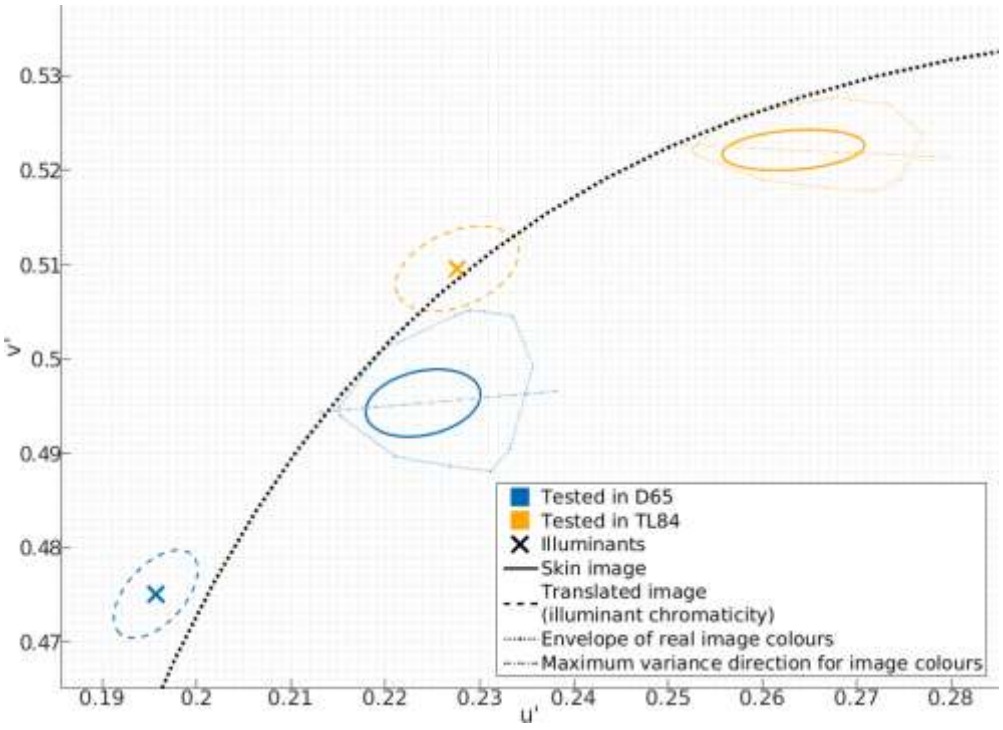
Experiment 1 (simulated skin)
vs.
Experiment 3 (illuminant-chromaticity)

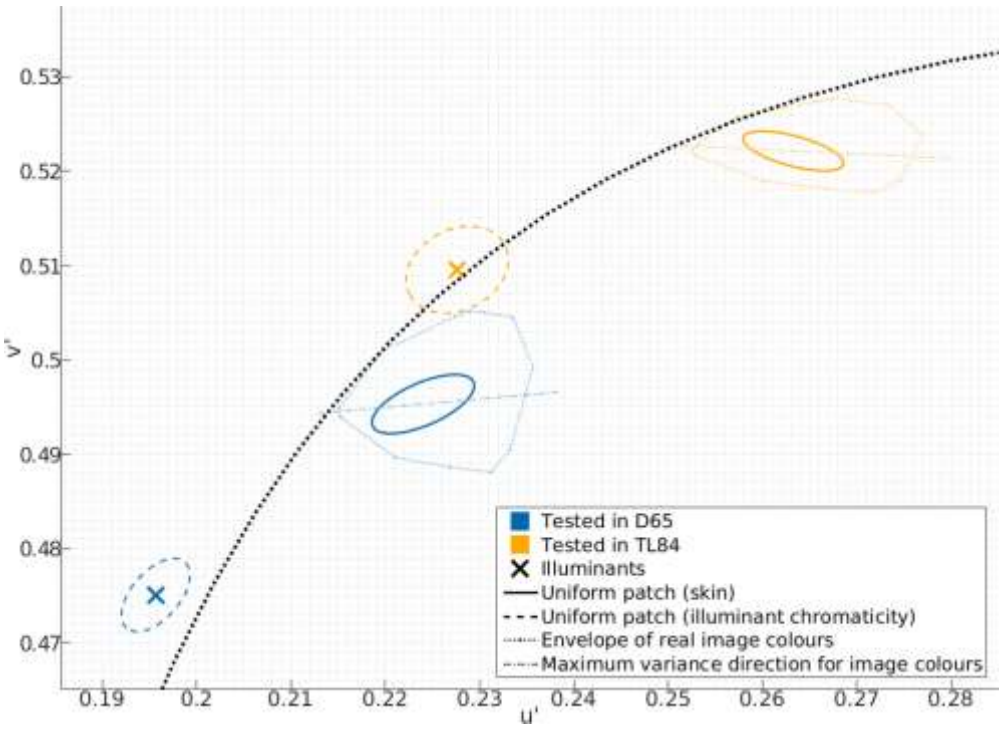
Observer: HA


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

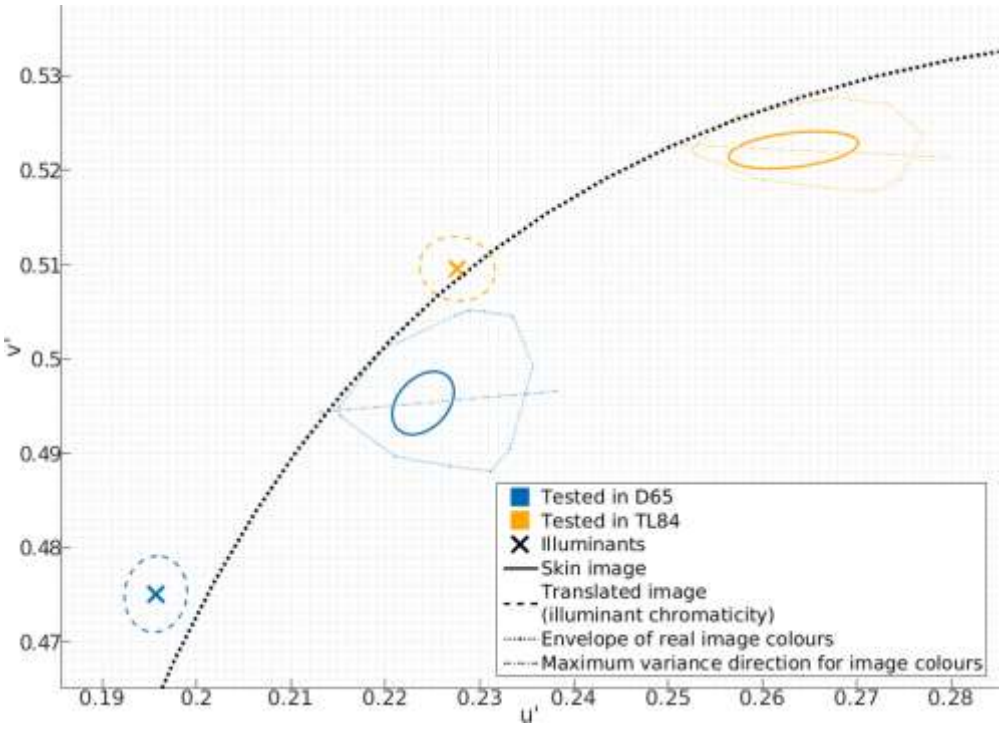

0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

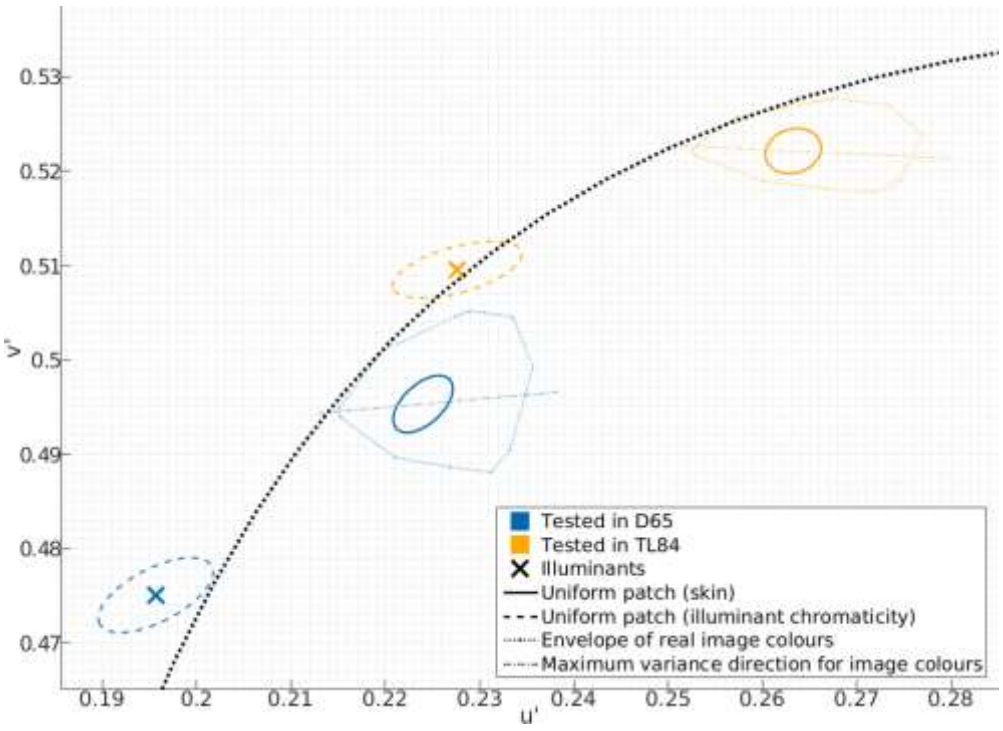
Observer: MH


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

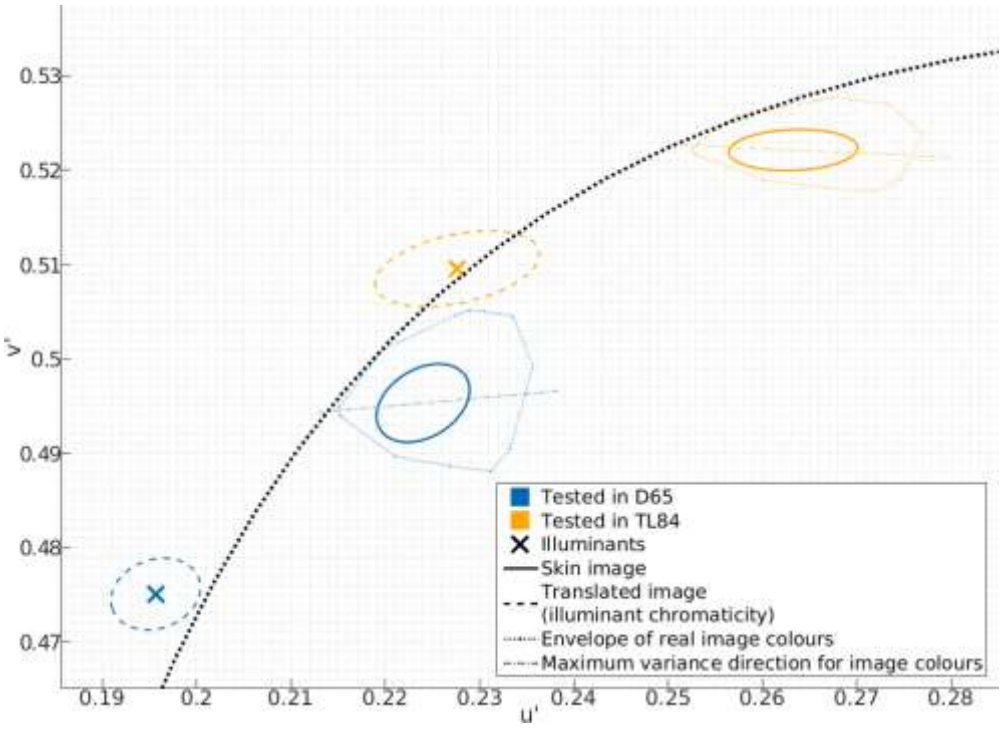

0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

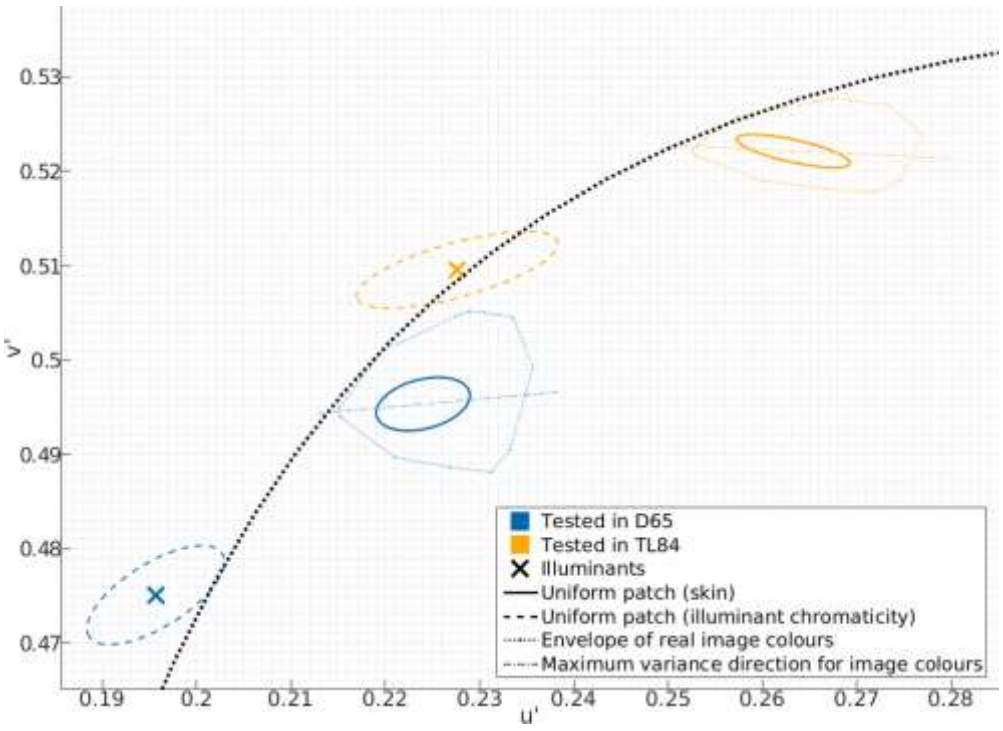
Observer: BK


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

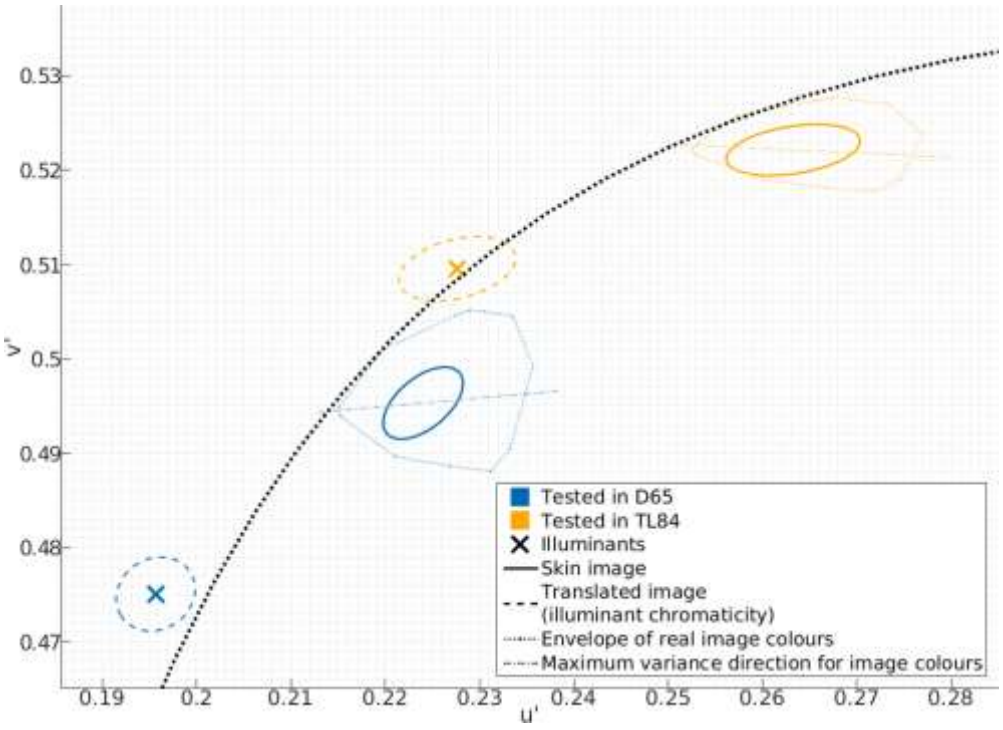

0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

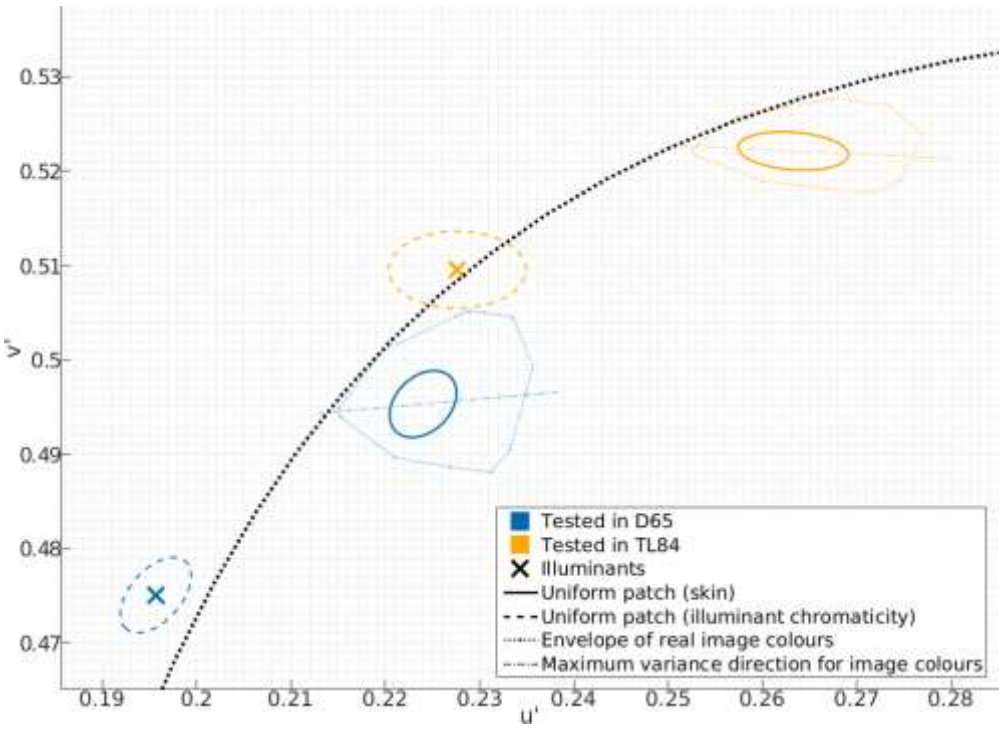
Observer: CN


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

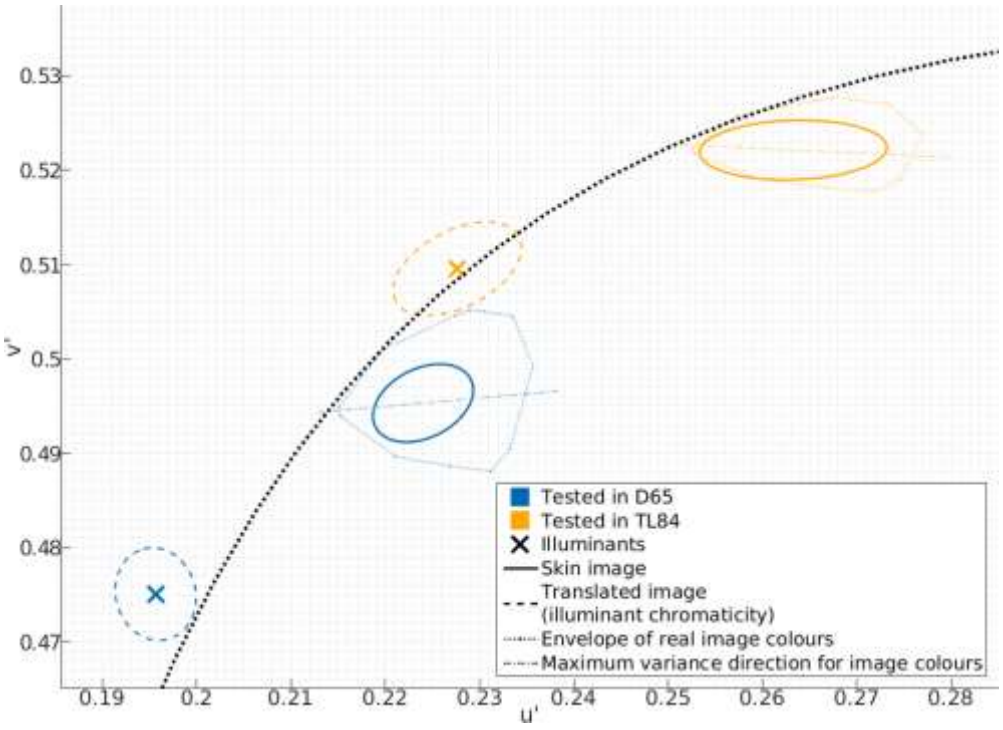

0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

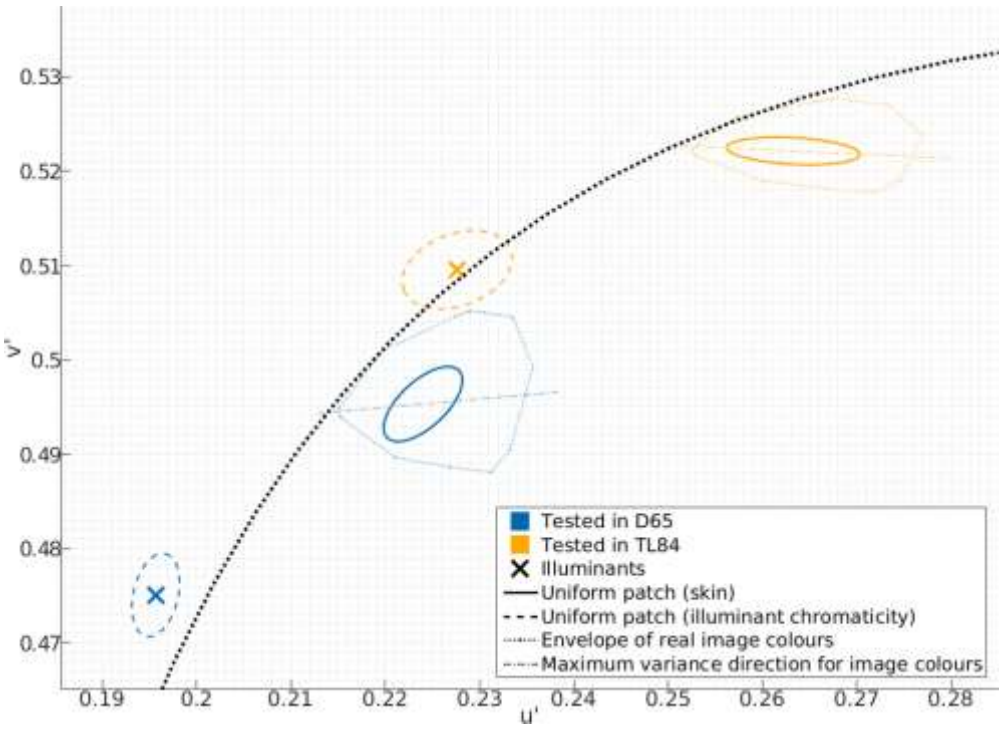
Observer: PR


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

Observer: CT


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Skin image
Translated image
(illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours


0.53
0.52
0.51
v'
0.5
0.49
0.48
0.47
0.19
0.2
0.21
0.22
0.23
0.24
0.25
0.26
0.27
0.28
u'
Tested in D65
Tested in TL84
Illuminants
Uniform patch (skin)
Uniform patch (illuminant chromaticity)
Envelope of real image colours
Maximum variance direction for image colours

References

Agahian, F., Amirshahi, S. A., & Amirshahi, S. H. (2008). Reconstruction of reflectance spectra using weighted principal component analysis. *Color Research & Application*, *33*(5), 360–371. http://doi.org/10.1002/col.20431

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. http://doi.org/10.1109/TAC.1974.1100705

Al-Harbi, F. A., Ayad, N. M., Saber, M. A., ArRejaie, A. S., & Morgano, S. M. (2015). Mechanical behavior and color change of facial prosthetic elastomers after outdoor weathering in a hot and humid climate. *The Journal of Prosthetic Dentistry*, *113*(2), 146–51. http://doi.org/10.1016/j.prosdent.2014.09.008

Atick, J., Li, Z., & Redlich, N. (1993). What does post-adaptation color appearance reveal about cortical color representation? *Vision Research*, *33*(1), 123–129. http://doi.org/10.1016/0042-6989(93)90065-5

Awange, J. L., Bae, K.-H., & Claessens, S. J. (2008). Procrustean solution of the 9-parameter transformation problem. *Earth, Planets and Space*, *60*(6), 529–537. http://doi.org/10.1186/BF03353115

Babaei, V., Amirshahi, S. H., & Agahian, F. (2011). Using weighted pseudo-inverse method for reconstruction of reflectance spectra and analyzing the dataset in terms of normality. *Color Research & Application*, *36*(4), 295–305. http://doi.org/10.1002/col.20613

Bar-Haim, Y., Saidel, T., & Yovel, G. (2009). The role of skin colour in face recognition. *Perception*, *38*(1), 145–148. http://doi.org/10.1068/p6307

Barlow, H., & Földiák, P. (1989). Adaptation and decorrelation in the cortex. In *The computing neuron* (pp. 54–72). CONF, Addison-Wesley Longman Publishing Co., Inc.

Bäuml, K.-H. (2002). Color appearance of spatial pattern: the role of increments and decrements. *Vision Research*, *42*(13), 1627–1637. http://doi.org/10.1016/S0042-6989(02)00096-2

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*, *31*(i10).

Bianco, S., & Schettini, R. (2010). Two new von Kries based chromatic adaptation transforms found by numerical optimization. *Color Research & Application*, *35*(3), 184–192. http://doi.org/10.1002/col.20573

Blackwell, T., & Buchsbaum, G. (1988). The effect of spatial and chromatic parameters on chromatic induction. *Color Research & Application*, *13*(3), 166–173. http://doi.org/10.1002/col.5080130309

Bompas, A., Powell, G., & Sumner, P. (2013). Systematic biases in adult color perception persist despite lifelong information sufficient to calibrate them. *Journal of Vision*, *13*(1), 19. http://doi.org/10.1167/13.1.19

Brainard, D. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America A*, *15*(2), 307. http://doi.org/10.1364/JOSAA.15.000307

Brainard, D., & Ishigami, K. (1995). Factors influencing the appearance of CRT colors. In *3rd Color and Imaging Conference Final Program and Proceedings* (pp. 62–66). Society for Imaging Science and Technology.

Brainard, D., & Wandell, B. (1992). Asymmetric color matching: how color appearance depends on the illuminant. *Journal of the Optical Society of America A*, *9*(9), 1433. http://doi.org/10.1364/JOSAA.9.001433

Brown, W. R. J. (1951). The Influence of Luminance Level on Visual Sensitivity to Color Differences. *Journal of the Optical Society of America*, *41*(10), 684. http://doi.org/10.1364/JOSA.41.000684

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, *33*(2), 261–304. JOUR. http://doi.org/10.1177/0049124104268644

Calabria, A., & Fairchild, M. (2001). Herding CATs: A Comparison of Linear Chromatic-Adaptation Transforms for CIECAM97s. *Color and Imaging Conference*, *2001*(1), 174–178. article.

Changizi, M. A., Zhang, Q., & Shimojo, S. (2006). Bare skin, blood and the evolution of primate colour vision. *Biology Letters*, *2*(2), 217–21. http://doi.org/10.1098/rsbl.2006.0440

Chauhan, T., Perales, E., Xiao, K., Hird, E., Karatzas, D., & Wuerger, S. (2014). The achromatic locus: effect of navigation direction in color space. *Journal of Vision*, *14*(1), 1–11. http://doi.org/10.1167/14.1.25

Chen, T. F., Baranoski, G. V. G., Kimmel, B. W., & Miranda, E. (2015). Hyperspectral Modeling of Skin Appearance. *ACM Transactions on Graphics*, *34*(3), 1–14. http://doi.org/10.1145/2701416

Chong, H. Y., Gortler, S. J., & Zickler, T. (2007). The von Kries Hypothesis and a Basis for Color Constancy. In *2007 IEEE 11th International Conference on Computer Vision* (pp. 1–8). IEEE. http://doi.org/10.1109/ICCV.2007.4409102

CIE. (2004). *15.3: 2004 Colorimetry*. *Vienna, Austria: CIE Central Bureau*. JOUR.

Conway, B. R., & Stoughton, C. M. (2009). Response: Towards a neural representation for unique hues. *Current Biology*, *19*(11), R442–R443. http://doi.org/10.1016/j.cub.2009.04.056

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. JOUR.

Crichton, S., Pichat, J., Mackiewicz, M., Tian, G.-Y., & Hurlbert, A. (2012). Skin chromaticity gamuts for illumination recovery. In *Conference on Colour in Graphics, Imaging, and Vision* (Vol. 2012, pp. 266–271). CONF, Society for Imaging Science and Technology.

Cula, O. G., Dana, K. J., Murphy, F. P., & Rao, B. K. (2004). Bidirectional imaging

and modeling of skin texture. *IEEE Transactions on Bio-Medical Engineering*, *51*(12), 2148–59. http://doi.org/10.1109/TBME.2004.836520

Danilova, M., & Mollon, J. (2010). Parafoveal color discrimination: a chromaticity locus of enhanced discrimination. *Journal of Vision*, *10*(1), 4.1-9. http://doi.org/10.1167/10.1.4

Delahunt, P., & Brainard, D. (2004). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision*, *4*(2), 1. http://doi.org/10.1167/4.2.1

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology*, *357*(1), 241–265. http://doi.org/10.1113/jphysiol.1984.sp015499

Doerschner, K., Boyaci, H., & Maloney, L. T. (2004). Human observers compensate for secondary illumination originating in nearby chromatic surfaces. *Journal of Vision*, *4*(2), 3. http://doi.org/10.1167/4.2.3

Donner, C., Weyrich, T., D'Eon, E., Ramamoorthi, R., & Rusinkiewicz, S. (2008). A layered, heterogeneous reflectance model for acquiring and rendering human skin. *ACM Transactions on Graphics (TOG)*, *27*(5), 140. JOUR.

Dunn, F. A., Lankheet, M. J., & Rieke, F. (2007). Light adaptation in cone vision involves switching between receptor and post-receptor sites. *Nature*, *449*(7162), 603–606. http://doi.org/10.1038/nature06150

Ekroll, V., Faul, F., Niederee, R., & Richter, E. (2002). The natural center of chromaticity space is not always achromatic: A new look at color induction. *Proceedings of the National Academy of Sciences*, *99*(20), 13352–13356. http://doi.org/10.1073/pnas.192216699

Fairchild, M. (2013). *Color Appearance Models*. Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/9781118653128

Fairchild, M., & Johnson, G. (2004). iCAM framework for image appearance, differences, and quality. *Journal of Electronic Imaging*, *13*(1), 126. http://doi.org/10.1117/1.1635368

Fine, L., & Dent, H. D. (1978). Color and its application in maxillofacial prosthetics. *The Journal of Prosthetic Dentistry*, *39*(2), 188–192. http://doi.org/10.1016/S0022-3913(78)80019-5

Fink, B., Grammer, K., & Thornhill, R. (2001). Human (Homo sapiens) facial attractiveness in relation to skin texture and color. *Journal of Comparative Psychology*, *115*(1), 92–99.

Fink, B., Matts, P. J., Klingenberg, H., Kuntze, S., Weege, B., & Grammer, K. (2008). Visual attention to variation in female facial skin color distribution. *Journal of Cosmetic Dermatology*, *7*(2), 155–61. http://doi.org/10.1111/j.1473-2165.2008.00382.x

Finlayson, G., Drew, M., & Funt, B. (1993). Diagonal transforms suffice for color constancy. *Computer Vision, 1993. Proceedings., Fourth International Conference on*, *11*(11), 164–171. http://doi.org/10.1109/ICCV.1993.378223

Finlayson, G., & Süsstrunk, S. (2000). Spectral Sharpening and the Bradford Transform. *Colour Image Science (CIS 2000)*.

Fisher, R. (1953). Dispersion on a Sphere. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *217*(1130), 295–305. http://doi.org/10.1098/rspa.1953.0064

Geisler, W. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, *59*, 167–92. http://doi.org/10.1146/annurev.psych.58.110405.085632

Giard, F., & Guitton, M. J. (2010). Beauty or realism: The dimensions of skin from cognitive sciences to computer graphics. *Computers in Human Behavior*, *26*(6), 1748–1752. http://doi.org/10.1016/j.chb.2010.07.001

Giesel, M., Hansen, T., & Gegenfurtner, K. (2009). The discrimination of chromatic textures. *Journal of Vision*, *9*(2009), 11.1-28. http://doi.org/10.1167/9.9.11

Goda, N., Koida, K., & Komatsu, H. (2009). Colour Representation in Lateral Geniculate Nucleus and Natural Colour Distributions. In A. Trémeau, R. Schettini, & S. Tominaga (Eds.), *Lecture Notes in Computer Science* (pp. 23–30). CHAP, Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-03265-3_3

Golz, J. (2010). Colour constancy: Influence of viewing behaviour on grey settings. *Perception*, *39*(5), 606–619. http://doi.org/10.1068/p6052

Golz, J., & MacLeod, D. (2002). Influence of scene statistics on colour constancy. *Nature*, *415*(6872), 637–40. http://doi.org/10.1038/415637a

Granzier, J., Toscani, M., & Gegenfurtner, K. (2012). Role of eye movements in chromatic induction. *Journal of the Optical Society of America A*, *29*(2), A353. http://doi.org/10.1364/JOSAA.29.00A353

Grassmann, H. (1853). Zur Theorie der Farbenmischung. *Annalen Der Physik*, *165*(5), 69–84. JOUR. http://doi.org/10.1002/andp.18531650505

Hansen, T., Giesel, M., & Gegenfurtner, K. (2008). Chromatic discrimination of natural objects. *Journal of Vision*, *8*(1), 1–19. http://doi.org/10.1167/8.1.2

Hansen, T., Walter, S., & Gegenfurtner, K. (2007). Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision*, *7*(4), 2–2. http://doi.org/10.1167/7.4.2

Harrison, D., & Kanji, G. K. (1988). The development of analysis of variance for circular data. *Journal of Applied Statistics*, *15*(2), 197–223. http://doi.org/10.1080/02664768800000026

Helmholtz, H. von. (1867). *Handbuch der physiologischen Optik* (Vol. 9). BOOK, Voss.

Helson, H., & Michels, W. C. (1948). The Effect of Chromatic Adaptation on Achromaticity. *Journal of the Optical Society of America*, *38*(12), 1025. http://doi.org/10.1364/JOSA.38.001025

Henley, S., & Fairchild, M. (2000). Quantifying Mixed Adaptation in Cross-Media Color Reproduction. *Color and Imaging Conference*, *2000*(1), 305–310. article.

Hering, E. (1920). *Outlines of a Theory of the Light Sense. Translated from German by L. Hurvich & D. Jameson 1964*. BOOK, Cambridge: Harvard Univ. Press.

Hinks, D., Cárdenas, L., Kuehni, R., & Shamey, R. (2007). Unique-hue stimulus selection using Munsell color chips. *Journal of the Optical Society of America A*, *24*(10), 3371. http://doi.org/10.1364/JOSAA.24.003371

Hungerford, E., Beatty, M. W., Marx, D. B., Simetich, B., & Wee, A. G. (2013). Coverage error of commercial skin pigments as compared to human facial skin tones. *Journal of Dentistry*, *41*(11), 986–91. http://doi.org/10.1016/j.jdent.2013.07.010

Hunt, R. W. G. (1991). Revised colour-appearance model for related and unrelated colours. *Color Research & Application*, *16*(3), 146–165. http://doi.org/10.1002/col.5080160306

Hurvich, C., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297–307. http://doi.org/10.1093/biomet/76.2.297

Hurvich, L., Jameson, D., & Cohen, J. (1968). The experimental determination of unique green in the spectrum. *Perception & Psychophysics*, *4*(2), 65–68. http://doi.org/10.3758/BF03209508

Jameson, D., & Hurvich, L. (1964). Theory of brightness and color contrast in human vision. *Vision Research*, *4*(1–2), 135–154. http://doi.org/10.1016/0042-6989(64)90037-9

Jameson, D., Hurvich, L. M., & Varner, F. D. (1979). Receptoral and Postreceptoral Visual Processes in Recovery from Chromatic Adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(6), 3034–3038. JOUR. Retrieved from http://www.jstor.org/stable/69911

Jammalamadaka, S. R., & Sengupta, A. (2001). *Topics in Circular Statistics*. World Scientific.

Jimenez, J., Sundstedt, V., & Gutierrez, D. (2009). Screen-space perceptual rendering of human skin. *ACM Transactions on Applied Perception (TAP)*, *6*(4), 23. JOUR.

Jones, M. J., & Rehg, J. M. (2002). Statistical color models with application to skin detection. *International Journal of Computer Vision*, *46*(1), 81–96. http://doi.org/10.1023/A:1013200319198

Kaiser, P., & Boynton, R. (1996). *Human color vision*. BOOK, Washington, DC: Optical Society of America.

Katoh, N., Nakabayashi, K., Ito, M., & Ohno, S. (1998). Effect of ambient light on the color appearance of softcopy images: Mixed chromatic adaptation for self-luminous displays. *Journal of Electronic Imaging*, *7*(4), 794. http://doi.org/10.1117/1.482665

Kaur, P., Dana, K. J., & Cula, G. O. (2015). From photography to microbiology: Eigenbiome models for skin appearance. In *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 1–10). IEEE. http://doi.org/10.1109/CVPRW.2015.7301310

Khemis, A., Kaiafa, A., Queille-Roussel, C., Duteil, L., & Ortonne, J. P. (2007).

Evaluation of efficacy and safety of rucinol serum in patients with melasma: a randomized controlled trial. *The British Journal of Dermatology*, *156*(5), 997–1004. http://doi.org/10.1111/j.1365-2133.2007.07814.x

Kraft, J. M., & Brainard, D. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences*, *96*(1), 307–312. http://doi.org/10.1073/pnas.96.1.307

Krauskopf, J., & Gegenfurtner, K. (1992). Color discrimination and adaptation. *Vision Research*, *32*(11), 2165–2175. http://doi.org/10.1016/0042-6989(92)90077-V

Krishnaswamy, A., & Baranoski, G. (2004). A biophysically-based spectral model of light interaction with human skin. *Computer Graphics Forum*, *23*(3 SPEC. ISS.), 331–340. http://doi.org/10.1111/j.1467-8659.2004.00764.x

Kuehni, R. (2005). Focal Color Variability and Unique Hue Stimulus Variability. *Journal of Cognition and Culture*, *5*(3), 409–426. http://doi.org/10.1163/156853705774648554

Kuehni, R. (2014). Unique hues and their stimuli-state of the art. *Color Research & Application*, *39*(3), 279–287. http://doi.org/10.1002/col.21793

Kulikowski, J., Daugirdiene, A., Panorgias, A., Stanikunas, R., Vaitkevicius, H., & Murray, I. J. (2012). Systematic violations of von Kries rule reveal its limitations for explaining color and lightness constancy. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *29*(2), A275-89. http://doi.org/10.1364/JOSAA.29.00A275

Kuriki, I. (2006). The loci of achromatic points in a real environment under various illuminant chromaticities. *Vision Research*, *46*(19), 3055–3066. http://doi.org/10.1016/j.visres.2006.03.012

Lee, B., Pokorny, J., Martin, P., Valbergt, A., & Smith, V. (1990). Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America A*, *7*(12), 2223. http://doi.org/10.1364/JOSAA.7.002223

Lee, R. J., Dawson, K. A., & Smithson, H. E. (2012). Slow updating of the achromatic point after a change in illumination. *Journal of Vision*, *12*(1), 19–19. http://doi.org/10.1167/12.1.19

MacAdam, D. (1942). Visual Sensitivities to Color Differences in Daylight. *Journal of the Optical Society of America*, *32*(5), 247. http://doi.org/10.1364/JOSA.32.000247

MacAdam, D. (1956). Chromatic Adaptation*. *Journal of the Optical Society of America*, *46*(7), 500. http://doi.org/10.1364/JOSA.46.000500

MacAdam, D. (1961). A nonlinear hypothesis for chromatic adaptation. *Vision Research*, *1*(1–2), 9–41. http://doi.org/10.1016/0042-6989(61)90020-7

MacAdam, D. (1963). Chromatic Adaptation II Nonlinear Hypothesis*. *Journal of the Optical Society of America*, *53*(12), 1441. http://doi.org/10.1364/JOSA.53.001441

MacLeod, D., & Boynton, R. (1979). Chromaticity diagram showing cone excitation

by stimuli of equal luminance. *Journal of the Optical Society of America*, *69*(8), 1183. http://doi.org/10.1364/JOSA.69.001183

Makin, A., & Chauhan, T. (2014). Memory-guided tracking through physical space and feature space. *Journal of Vision*, *14*(13), 10–10. http://doi.org/10.1167/14.13.10

Maxwell, J. C. (1860). On the theory of compound colours, and the relations of the colours of the spectrum. *Philosophical Transactions of the Royal Society of London*, *150*, 57–84. JOUR.

McMillan, G. P., Hanson, T. E., Saunders, G., & Gallun, F. J. (2013). A two-component circular regression model for repeated measures auditory localization data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *62*(4), 515–534. http://doi.org/10.1111/rssc.12004

Melgosa, M., Hita, E., Poza, A. J., Alman, D. H., & Berns, R. S. (1997). Suprathreshold color-difference ellipsoids for surface colors. *Color Research & Application*, *22*(3), 148–155. http://doi.org/10.1002/(SICI)1520-6378(199706)22:3<148::AID-COL3>3.0.CO;2-R

Melgosa, M., Pérez, M. M., El Moraghi, A., & Hita, E. (1999). Color discrimination results from a CRT device: Influence of luminance. *Color Research & Application*, *24*(1), 38–44. http://doi.org/10.1002/(SICI)1520-6378(199902)24:1<38::AID-COL8>3.0.CO;2-I

Mollon, J. (2006). Monge: The Verriest Lecture, Lyon, July 2005. *Visual Neuroscience*, *23*(3–4), 297–309. http://doi.org/10.1017/S0952523806233479

Mollon, J. (2009). A neural basis for unique hues? *Current Biology*, *19*(11), R441-2-3. http://doi.org/10.1016/j.cub.2009.05.008

Mollon, J., & Jordan, G. (1997). On the nature of unique hues. In C. Dickinson, I. Murray, & D. Carden (Eds.), *John Dalton's Colour Vision Legacy* (pp. 381–392). JOUR, Taylor and Francis.

Montag, E. D., & Berns, R. S. (2000). Lightness dependencies and the effect of texture on suprathreshold lightness tolerances. *Color Research & Application*, *25*(4), 241–249. http://doi.org/10.1002/1520-6378(200008)25:4<241::AID-COL4>3.0.CO;2-E

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, *4*(2), 61–64. JOUR.

Moroney, N., Fairchild, M., Hunt, R. W. G., Li, C., Luo, M. R., & Newman, T. (2002). The CIECAM02 Color Appearance Model. In *Color Imaging Conference* (pp. 23–27).

Nascimento, S., Ferreira, F., & Foster, D. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A*, *19*(8), 1484. http://doi.org/10.1364/JOSAA.19.001484

Nayatani, Y., Takahama, K., & Sobagaki, H. (1981). Formulation of a Nonlinear Model of Chromatic Adaptation. *Color Research & Application*, *6*(3), 161–171. http://doi.org/10.1002/col.5080060308

Newton, I. (1730). *Opticks, Or a Treatise of the Reflections, Refractions, Inflections and Colours of Light* (Vol. 678). BOOK, William Innys.

Paláncz, B., Zaletnyik, P., Awange, J. L., & Heck, B. (2010). Extension of the ABC-Procrustes algorithm for 3D affine coordinate transformation. *Earth, Planets and Space*, *62*(11), 857–862. http://doi.org/10.5047/eps.2010.10.004

Paravina, R. D., Majkic, G., Del Mar Perez, M., & Kiat-Amnuay, S. (2009). Color difference thresholds of maxillofacial skin replications. *Journal of Prosthodontics : Official Journal of the American College of Prosthodontists*, *18*(7), 618–25. http://doi.org/10.1111/j.1532-849X.2009.00465.x

Philipona, D. L., & Regan, J. K. O. (2006). Singularities in Reflection Properties. *Visual Neuroscience*, *23*(3–4), 331–339.

Poirson, A., & Wandell, B. (1990a). Task-dependent color discrimination. *Journal of the Optical Society of America A*, *7*(4), 776. http://doi.org/10.1364/JOSAA.7.000776

Poirson, A., & Wandell, B. (1990b). The ellipsoidal representation of spectral sensitivity. *Vision Research*, *30*(4), 647–652. http://doi.org/10.1016/0042-6989(90)90075-V

Poirson, A., Wandell, B., Varner, D., & Brainard, D. (1990). Surface characterizations of color thresholds. *Journal of the Optical Society of America A*, *7*(4), 783. http://doi.org/10.1364/JOSAA.7.000783

Purdy, D. (1931). Spectral Hue as a Function of Intensity. *The American Journal of Psychology*, *43*(4), 541–559. http://doi.org/10.2307/1415157

Regan, B. C., Julliot, C., Simmen, B., Viénot, F., Charles-Dominique, P., & Mollon, J. (2001). Fruits, foliage and the evolution of primate colour vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1407), 229–83. http://doi.org/10.1098/rstb.2000.0773

Regan, B. C., Reffin, J. P., & Mollon, J. (1994). Luminance Noise and the Rapid-Determination of Discrimination Ellipses in Color Deficiency. *Vision Research*, *34*(10), 1279–1299. http://doi.org/10.1016/0042-6989(94)90203-8

Robbins, H., & Monro, S. (1951). A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, *22*(3), 400–407. JOUR.

Rubin, M. L. (1961). Spectral Hue Loci of Normal and Anomalous Trichromates*. *American Journal of Ophthalmology*, *52*(2), 166–172. http://doi.org/10.1016/0002-9394(61)91112-6

Schefrin, B. E., & Werner, J. S. (1990). Loci of spectral unique hues throughout the life span. *Journal of the Optical Society of America A*, *7*(2), 305. http://doi.org/10.1364/JOSAA.7.000305

Schefrin, B. E., & Werner, J. S. (1993). Age-related changes in the color appearance of broadband surfaces. *Color Research & Application*, *18*(6), 380–389. http://doi.org/10.1002/col.5080180606

Schrödinger, E. (1920). Grundlinien einer Theorie der Farbenmetrik im Tagessehen. *Annalen Der Physik*, *368*(21), 397–426.

http://doi.org/10.1002/andp.19203682102

Shen, H.-L., Cai, P.-Q., Shao, S.-J., & Xin, J. (2007). Reflectance reconstruction for multispectral imaging by adaptive Wiener estimation. *Optics Express*, *15*(23), 15545–15554. JOUR.

Shevell, S. K. (1978). The dual role of chromatic backgrounds in color perception. *Vision Research*, *18*(12), 1649–1661. http://doi.org/10.1016/0042-6989(78)90257-2

Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, *15*(2), 161–171. http://doi.org/10.1016/0042-6989(75)90203-5

Smithson, H. (2014). S-cone psychophysics. *Visual Neuroscience*, *31*(2), 211–225. http://doi.org/10.1017/S0952523814000030

Stephen, I. D., Law Smith, M. J., Stirrat, M. R., & Perrett, D. I. (2009). Facial Skin Coloration Affects Perceived Health of Human Faces. *International Journal of Primatology*, *30*(6), 845–857. http://doi.org/10.1007/s10764-009-9380-z

Stephens, M. A. (1969). Multi-sample tests for the Fisher distribution for directions. *Biometrika*, *56*(1), 169–181. http://doi.org/10.1093/biomet/56.1.169

Stevens, J. (2009). *Applied multivariate statistics for the social sciences.* (5th ed.). Taylor & Francis.

Stiles, W. S. (1946). A modified Helmholtz line-element in brightness-colour space. *Proceedings of the Physical Society*, *58*(1), 41–65. http://doi.org/10.1088/0959-5309/58/1/305

Stiles, W. S., & Burch, J. M. (1955). Interim Report to the Commission Internationale de l'Eclairage, Zurich, 1955, on the National Physical Laboratory's Investigation of Colour-matching (1955). *Optica Acta: International Journal of Optics*, *2*(4), 168–181. http://doi.org/10.1080/713821039

Stockman, A., & Sharpe, L. T. (2000). The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, *40*(13), 1711–1737. http://doi.org/10.1016/s0042-6989(00)00021-3

Stockman, A., & Sharpe, L. T. (2006). Into the twilight zone: the complexities of mesopic vision and luminous efficiency. *Ophthalmic and Physiological Optics*, *26*(3), 225–239. http://doi.org/10.1111/j.1475-1313.2006.00325.x

Stoughton, C. M., & Conway, B. R. (2008). Neural basis for unique hues. *Current Biology*, *18*(16), R698-9. http://doi.org/10.1016/j.cub.2008.06.018

Sugiura, N. (1978). Further analysts of the data by akaike' s information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, *7*(1), 13–26. http://doi.org/10.1080/03610927808827599

Tailby, C., Solomon, S. G., & Lennie, P. (2008). Functional asymmetries in visual pathways carrying S-cone signals in macaque. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(15), 4078–87. http://doi.org/10.1523/JNEUROSCI.5338-07.2008

Tan, K. W., & Stephen, I. D. (2013). Colour detection thresholds in faces and colour patches. *Perception*, *42*(7), 733–41.

Tangkijviwat, U., Rattanakasamsuk, K., & Shinoda, H. (2010). Color preference affected by mode of color appearance. *Color Research & Application*, *35*(1), 50–61. JOUR. http://doi.org/http://dx.doi.org/10.1002/col.20536

Taylor, M. (1971). On the efficiency of psychophysical measurement. *The Journal of the Acoustical Society of America*, *49*(2), Suppl 2:505-508. http://doi.org/10.1121/1.1912379

Taylor, M., & Creelman, D. (1967). PEST: Efficient Estimates on Probability Functions. *The Journal of the Acoustical Society of America*, *41*(4A), 782. http://doi.org/10.1121/1.1910407

Thomson, M., & Foster, D. (1997). Role of second- and third-order statistics in the discriminability of natural images. *Journal of the Optical Society of America A*, *14*(9), 2081. http://doi.org/10.1364/JOSAA.14.002081

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, *35*(17), 2503–2522. http://doi.org/10.1016/0042-6989(95)00016-X

Tsumura, N., Ojima, N., Sato, K., Shiraishi, M., Shimizu, H., Nabeshima, H., … Miyake, Y. (2003). Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. In *ACM SIGGRAPH 2003 Papers on - SIGGRAPH '03* (Vol. 22, p. 770). New York, New York, USA: ACM Press. http://doi.org/10.1145/1201775.882344

Twer, T., & MacLeod, D. (2001). Optimal nonlinear codes for the perception of natural colours. *Network: Computation in Neural Systems*, *12*(3), 395–407. http://doi.org/10.1080/net.12.3.395.407

Valberg, A. (1971). A method for the precise determination of achromatic colours including white. *Vision Research*, *11*(2), 157–160. http://doi.org/10.1016/0042-6989(71)90231-8

Valberg, A. (2001). Unique hues: an old problem for a new generation. *Vision Research*, *41*(13), 1645–1657. http://doi.org/10.1016/S0042-6989(01)00041-4

Vazquez-Corral, J., O'Regan, K., Vanrell, M., & Finlayson, G. (2012). A new spectrally sharpened sensor basis to predict color naming, unique hues, and hue cancellation. *Journal of Vision*, *12*(6), 7. JOUR.

Vezhnevets, V., Sazonov, V., & Andreeva, A. (2003). A survey on pixel-based skin color detection techniques. In *Proc. Graphicon* (Vol. 3, pp. 85–92). inproceedings.

Volbrecht, V. J., Nerger, J. L., & Harlow, C. E. (1997). The Bimodality of Unique Green Revisited. *Vision Research*, *37*(4), 407–416. http://doi.org/10.1016/S0042-6989(96)00158-7

von Kries, J. (1905). Die gesichtsempfindungen. *Handbuch Der Physiologie Des Menschen*, *3*, 109–282. JOUR.

Vos, J. J., & Walraven, P. L. (1972). An analytical description of the line element in the zone-fluctuation model of colour vision—I. Basic concepts. *Vision Research*,

12(8), 1327–1344. http://doi.org/10.1016/0042-6989(72)90181-2

Walraven, J. (1976). Discounting the background—the missing link in the explanation of chromatic induction. *Vision Research*, *16*(3), 289–295. http://doi.org/10.1016/0042-6989(76)90112-7

Walraven, J., & Werner, J. S. (1991). The invariance of unique white; a possible implication for normalizing cone action spectra. *Vision Research*, *31*(12), 2185–2193. http://doi.org/10.1016/0042-6989(91)90171-Z

Watson, A., & Pelli, D. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120. http://doi.org/10.3758/BF03202828

Watson, G., & Williams, E. (1956). On the Construction of Significance Tests on the Circle and the Sphere. *Biometrika*, *43*(3/4), 344–352. JOUR. http://doi.org/10.2307/2332913

Webster, M., Miyahara, E., Malkoc, G., & Raker, V. (2000). Variations in normal color vision II Unique hues. *Journal of the Optical Society of America A*, *17*(9), 1545. http://doi.org/10.1364/JOSAA.17.001545

Webster, M., & Mollon, J. (1991). Changes in colour appearance following post-receptoral adaptation. *Nature*, *349*(6306), 235–238. http://doi.org/10.1038/349235a0

Webster, M., & Mollon, J. (1994). The influence of contrast adaptation on color appearance. *Vision Research*, *34*(15), 1993–2020. http://doi.org/10.1016/0042-6989(94)90028-0

Webster, M., & Mollon, J. (1997). Adaptation and the color statistics of natural images. *Vision Research*, *37*(23), 3283–3298. http://doi.org/10.1016/S0042-6989(97)00125-9

Weibull, W. (1951). A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics*, *13*, 293–297. Retrieved from http://ci.nii.ac.jp/naid/10004008665/en/

Welbourne, L. E., Thompson, P. G., Wade, A. R., & Morland, A. B. (2013). The distribution of unique green wavelengths and its relationship to macular pigment density. *Journal of Vision*, *13*(8), 15. http://doi.org/10.1167/13.8.15

West, G. (1979). Color perception and the limits of color constancy. *Journal of Mathematical Biology*, *8*(1), 47–53. http://doi.org/10.1007/BF00280585

West, G., & Brill, M. H. (1982). Necessary and sufficient conditions for Von Kries chromatic adaptation to give color constancy. *Journal of Mathematical Biology*, *15*(2), 249–258. http://doi.org/10.1007/BF00275077

Wetherill, B. (1966). *Sequential methods in statistics*. book, London: Methuen. Retrieved from http://cds.cern.ch/record/104933

Witzel, C., Valkova, H., Hansen, T., & Gegenfurtner, K. (2011). Object knowledge modulates colour appearance. *I-Perception*, *2*(1), 13–49. http://doi.org/10.1068/i0396

Worthey, J. A., & Brill, M. H. (1986). Heuristic analysis of von Kries color constancy.

Journal of the Optical Society of America A, *3*(10), 1708. http://doi.org/10.1364/JOSAA.3.001708

Wuerger, S. (1996). Color Appearance Changes Resulting from Iso-luminant Chromatic Adaptation. *Vision Research*, *36*(19), 3107–3118. http://doi.org/10.1016/0042-6989(96)00057-0

Wuerger, S. (2013). Colour Constancy Across the Life Span: Evidence for Compensatory Mechanisms. *PLoS ONE*, *8*(5), e63921. http://doi.org/10.1371/journal.pone.0063921

Wuerger, S., Atkinson, P., & Cropper, S. (2005). The cone inputs to the unique-hue mechanisms. *Vision Research*, *45*(25–26), 3210–3223. http://doi.org/10.1016/j.visres.2005.06.016

Wuerger, S., Maloney, L., & Krauskopf, J. (1995). Proximity judgments in color space: Tests of a Euclidean color geometry. *Vision Research*, *35*(6), 827–835. http://doi.org/10.1016/0042-6989(94)00170-Q

Wyszecki, G., & Stiles, W. S. (2000). *Color science : concepts and methods, quantitative data, and formulae*. John Wiley & Sons.

Xiao, K., Fu, C., Mylonas, D., Karatzas, D., & Wuerger, S. (2013). Unique hue data for colour appearance models. Part II: Chromatic adaptation transform. *Color Research & Application*, *38*(1), 22–29. http://doi.org/10.1002/col.20725

Xiao, K., Liao, N., Zardawi, F., Liu, H., Noort, R. Van, Yang, Z., … Yates, J. M. (2012). Investigation of Chinese skin colour and appearance for skin colour reproduction. *Chinese Optics Letters*, *10*(8), 83301.

Xiao, K., Pointer, M., Cui, G., Chauhan, T., & Wuerger, S. (2015). Unique hue data for colour appearance models. Part III: Comparison with NCS unique hues. *Color Research & Application*, *40*(3), 256–263. http://doi.org/10.1002/col.21898

Xiao, K., Qin, Z., Chauhan, T., Li, C., & Wuerger, S. (2014). Principal component analysis for skin reflectance reconstruction. In *Color and Imaging Conference* (Vol. 2014, pp. 146–150). CONF, Society for Imaging Science and Technology.

Xiao, K., Wuerger, S., Fu, C., & Karatzas, D. (2011). Unique hue data for colour appearance models. Part I: Loci of unique hues and hue uniformity. *Color Research & Application*, *36*(5), 316–323. http://doi.org/10.1002/col.20637

Yebra, A., García, J. A., Nieves, J. L., & Romero, J. (2001). Chromatic discrimination in relation to luminance level. *Color Research & Application*, *26*(2), 123–131.

Young, T. (1802). The Bakerian lecture: On the theory of light and colours. *Philosophical Transactions of the Royal Society of London*, *92*, 12–48. JOUR.