Universitat Autònoma
de Barcelona

Colour constancy in natural images through colour naming and sensor sharpening

A dissertation submitted by **Javier Vázquez i Corral** at Universitat Autònoma de Barcelona to fulfil the degree of **Doctor of Philosophy**.

Bellaterra, March 16, 2011

Director	**Dr. Maria Vanrell i Martorell** Dep. Ciències de la Computació & Centre de Visió per Computador Universitat Autònoma de Barcelona
Co-director	**Prof. Graham D. Finlayson** School of Computing Sciences University of East Anglia
Thesis committee	**Dr. Elli Angelopoulou** Pattern Recognition Lab, Computer Science Department University of Erlangen-Nuremberg **Prof. Anya Hurlbert** Institute of Neuroscience University of Newcastle upon Tyne **Dr. Joost van de Weijer** Dep. Ciències de la Computació & Centre de Visió per Computador Universitat Autònoma de Barcelona

This document was typeset by the author using $\LaTeX\,2_\varepsilon$.

The research described in this book was carried out at the Computer Vision Center, Universitat Autònoma de Barcelona.

ISBN: 978-84-938351-1-8

Printed by

A mi abuela Carmen

THE SCIENTIFIC METHOD

Good theories
are those capable
of being disproved, Karl
Popper says. Like
that if I come
next week,
at the same time, sit
over my coffee
just exactly
there
where I looked up
and observed
you,
looking at me,
that I will find you,
again,
there,
and this time
have the courage
to smile.

Roald Hoffmann

Descubrí que no soy disciplinado por virtud, sino como reacción contra mi negligencia; que parezco generoso por encubrir mi mezquindad, que me paso de prudente por mal pensado, que soy conciliador para no sucumbir a mis cóleras reprimidas, que sólo soy puntual para que no se sepa cuán poco me importa el tiempo ajeno.

Gabriel García Márquez

Agraiments

Muchos años después, frente al tribunal de defensa de tesis, el doctorando Javier Vazquez-Corral había de recordar aquella mañana remota en que sus padres le llevaron por primera vez al colegio. Barcelona era entonces una gris ciudad preolímpica. Su mundo era tan reciente, que muchas cosas carecían de nombre para él, y para mencionarlas las tenía que señalar con el dedo.

Així podria començar, parafrasejant en gran García Márquez, la història dels meus anys d'estudiant, que ara toca a la seva fi. I com diuen en castellà que ”es de bien nacidos ser agradecidos” ara vull agrair a tots aquells que m'han ajudat en els darrers passos d'aquesta història.

Primer de tot vull agrair-li a la meva directora, na Maria Vanrell, la seva disponibilitat i paciència durant aquests poc més de quatre anys. I want also to thank my other supervisor Graham Finlayson for his guidance and support. Treballar amb caps així val molt la pena.

També vull agrair als companys del grup de color del centre de visió per computador el seu ajut quan era necessari. En Ramon per no prohibir-me l'entrada al seu despatx, n'Alejandro per respondre els dubtes més biològics, en Robert i les anècdotes de cada un a Norwich, en Xavi, n'Olivier, en Joost, n'Eduard, en Fahad, en Jordi, na Naila, na Shida, en David, en Jaime, n'Ahmed, en Marc i n'Ekain.

I am not forgetting about the Colour Lab at UEA. I am extremely graithful to David Connah, who has teached me the important things of the English culture and Roberto and our mutual support in these final moments of the thesis.I want also to thank Michael, Perla, Dome and Stuart for the teas and lunches we have spent together.

També vull agrair a la gent del CVC, en particular, na Montse i na Pilar, tota la feina que fan. A en Jose i n'Ivan per haver-me fet riure aquests darrers mesos. Dintre del CVC no em puc oblidar dels companys de despatx que he tingut en algun moment: Murad, Bhaskar, Ferran, Jaume, Joan, Carles, Ferran, Mohammad, Alicia i Marçal. Aquests tres darrers, juntament amb en David, en Jordi i en Nacho els vull també agrair els dinars a la Giralda que hem fet durant aquest temps.

A mig camí entre la feina i l'oci, entre Barcelona i Norwich, vull agrair a en Francesc les cerveses i els partits del Barça vistos en anglès. Seguint a Norwich, també vull donar les gràcies a en Toni i na Susana. Como no, debo acordarme también de David y Eva, vascos conquistando tierras inglesas. I am not forgetting Daniela, Rowena, Osama, the people from the Greenman Walking Club, the people from the CCP football team (particularly Rich, Seb, Mathias and Rob for the beers together), and the usual assistants of Grad's Bar and Tapas.

Tornant a la terra, primer de tot agrair als amics ”de sempre” la capacitat d'escoltar-me aquells dies que tot sortia malament: en Jose, en David, en Carlos, en Miguel, na Marta i na Raquel. No em puc oblidar de la gent del grup (veure imatge 1.7) pel seu acompanyament durant tot aquest temps. També en Xavi, na Miriam i en Ferran, i el grup jove de la parròquia. Al trio de francès: Anna, Emma i Maria els vull agrair les sortides per Gràcia. Als antics companys de Mates, particularment a en Victor, per continuar-me suportant. I a n'Alba per les xerrades catalano-valencianes.

Un cop fet això, només em queda agrair a la meva familia tot el seu suport. En especial als que dinem cada diumenge a casa: tío Manolo, abuela Carmen, Néstor, Yasmina, Marta, papá i mamá. Esta tesis es por y para vosotros. Os quiero.

Resum

El color és la conseqüència de la interacció de tres propietats físiques: la llum incident, la reflectància de l'objecte i la sensibilitat del sensor. La llum incident varia en condicions naturals, de manera que el coneixement de l' il·luminant que es troba a l'escena és una qüestió important en el color computacional. Una manera de tractar aquest problema en condicions de calibratge és seguint tres passos: 1) la construcció d'una base de sensors de banda estreta per poder aplicar el model diagonal, 2) la construcció del conjunt factible d'il·luminants, i 3) la definició de criteris per seleccionar el millor il·luminant. En aquest treball ens centrem en la introducció de criteris de percepció ,més concretament enfocant-nos en el primer i tercer pas.

Per fer front a l'etapa de selecció de l'il·luminant, nosaltres postulem que les categories bàsiques de color poden ser utilitzades com a categories d'ancoratge per recuperar el millor il·luminant. Aquestes categories de color han evolucionat per a codificar les estadístiques pertinents color natural. Per tant la imatge recuperada és la millor representació de l'escena etiquetada amb els termes bàsics de color. Es demostra amb diversos experiments com aquest criteri de selecció assoleix l'estat de l'art en la constància del color computacional. A més d'aquest resultat, demostrem psicofísicament que l'error angular que s'utilitza habitualment en la constància del color no es correlaciona amb les preferències humanes, i es proposa una nova avaluació basada en la percepció.

L'aplicació del criteri de selecció definit en aquesta tesi es basa enormement en l'ús d'un model diagonal per al canvi d'il·luminant. Per tant, la segona aportació es centra en la creació d'una base adequada de sensors de banda estreta de cara a representar les imatges naturals. Ens proposem utilitzar la tècnica de l'spectral sharpening per calcular una base única de banda estreta optimitzada per a representar a un ampli conjunt de reflectàncies naturals sota il·luminació natural i donades a partir de la base dels cons humans. Els sensors proposats permeten predir colors únics i les dades de World Color Survey independentment de l'il·luminant utilitzant una funció compacta de singularitat. A més, s'estudien les diferents famílies de sensors de banda estreta per reduir al mínim diferents mesures de percepció. Aquest estudi ens porta a ampliar el procediment de spherical sampling de 3D a 6D.

Diverses línies de recerca resten obertes, com ara, la mesura dels efectes de la utilització dels sensors de banda estreta en el mètode de selecció d'il·luminant proposat, o la inserció d' informació espacial contextual per millorar la hipòtesi de categories. Finalment, explorar com els sensors individuals es poden ajustar als colors d'una escena.

Abstract

Colour is derived from three physical properties: incident light, object reflectance and sensor sensitivities. Incident light varies under natural conditions; hence, recovering scene illuminant is an important issue in computational colour. One way to deal with this problem under calibrated conditions is by following three steps, 1) building a narrow-band sensor basis to accomplish the diagonal model, 2) building a feasible set of illuminants, and 3) defining criteria to select the best illuminant. In this work we focus on colour constancy for natural images by introducing perceptual criteria in the first and third stages.

To deal with the illuminant selection step, we hypothesize that basic colour categories can be used as anchor categories to recover the best illuminant. These colour names are related to how the human visual system has evolved to encode relevant natural colour statistics. Therefore the recovered image provides the best representation of the scene labelled with the basic colour terms. We demonstrate with several experiments how this selection criterion achieves current state-of-art results in computational colour constancy. In addition to this result, we psychophysically prove that usual angular error used in colour constancy does not correlate with human preferences, and we propose a new perceptual colour constancy evaluation.

The implementation of this selection criterion strongly relies on the use of a diagonal model for illuminant change. Then, the second contribution focuses on building an appropriate narrow-band sensor basis to represent natural images. We propose to use the spectral sharpening technique to compute a unique narrow-band basis optimized to represent a large set of natural reflectances under natural illuminants and given in the basis of human cones. The proposed sensors allow predicting unique hues and the World colour Survey data independently of the illuminant by using a compact singularity function. Additionally, we studied different families of sharp sensors to minimize different perceptual measures. This study brought us to extend the spherical sampling procedure from 3D to 6D.

Several research lines remain still open, such as, measuring the effects of using the computed sharp sensors on the category hypothesis; or inserting spatial contextual information to improve category hypothesis. Finally, to explore how individual sensors can be adjusted to the colours in a scene.

Contents

1 Introduction **1**
1.1 Colour science 1
1.2 Image formation 5
1.3 Colour constancy 6
1.4 Computational colour constancy 7
1.5 Colour naming 9
1.6 Thesis outline 10

2 Computational colour constancy: a review **11**
2.1 Providing a diagonal model 11
  2.1.1 Sensor-based spectral sharpening 12
  2.1.2 Perfect sharpening 13
  2.1.3 Data-based sharpening for a pair of illuminants 14
  2.1.4 Measurement tensor method 15
  2.1.5 Spherical sampling 16
  2.1.6 Other approaches 17
  2.1.7 Usual error measure 17
2.2 Building the feasible set 18
2.3 Illuminant selection 19
2.4 Usual error measure 21

I Part 1: Illuminant selection **23**

3 Colour naming for colour constancy **25**
3.1 Introduction 25
3.2 Basic term categories 27
3.3 Category methods 28
  3.3.1 Mathematical formulation 30
  3.3.2 Fast implementation 31
  3.3.3 Illuminant selection 32
3.4 Experiments 33
3.5 Results and discussion 35
3.6 Conclusions and further work 38

4 Psychophysical evaluation of colour constancy algorithms **41**
4.1 Introduction 41
4.2 Experimental setup 42

4.2.1 A new image dataset 44
4.2.2 Selected colour constancy algorithms 45
4.3 Results 47
4.4 Perceptual performance evaluation 51
4.5 Conclusion 53

II Part 2: Sensor sharpening 55

5 A new sensor basis to predict colour names and unique hues 57
5.1 Introduction 57
5.2 Philipona and O'Regan biological model 59
5.3 Philipona and O'Regan singularity index 60
5.4 Spectral sharpening to generalise the biological model 61
5.5 A new compact singularity index 63
5.6 Results 65
5.6.1 Comparison of PO's model vs SS model 65
5.6.2 Comparison PO's SI vs our CSI 66
5.6.3 Invariant reflectance properties on natural images 69
5.7 Conclusions 73

6 Experiments using spherical sampling 75
6.1 Introduction 75
6.2 Experiment 1: Minimising CIE Delta E 76
6.3 Experiment 2: Maximising colour ratio stability 78
6.4 Experiment 3: Multispectral case 81
6.5 Conclusions 82

7 Conclusions and further work 85
7.1 Conclusions 85
7.2 Further work 86

A Appendix: Lovisolo and Da Silva sphere sampling 89

Bibliography 91

List of Publications 101

List of Tables

3.1 Angular error on the different datasets. 36
3.2 Angular error limit by selecting the best solution during the combination on the different datasets. 37
3.3 Top-down approach. 38

4.1 Buttons codification. 48
4.2 Correlation between each observer and mean observer. 49
4.3 Results of the experiment in the 1-to-1 comparison. 49
4.4 Experiment results in a general comparison. 50
4.5 Results using Thurstone's Law of Comparative Judgement. 50
4.6 Results using Bradley-Terry ordinal pairwise comparison analysis. 51
4.7 Results using Thurstone's law of comparative judgement binary pairwise comparison analysis. 51
4.8 Angular error for the different methods on 415 images of the dataset. 53
4.9 Estimated perceptual angular error for the different methods on 415 images of the dataset. 54

5.1 Unique hues found in the different experiments and the PO prediction. EOS means End Of Spectrum. 66
5.2 Unique hues located by our indices applied directly to the sensors versus the Philipona index applied to the sensors. EOS means End Of Spectrum. 70

6.1 Results of the first experiment in the three different datasets (measure: mean Δ_ϵ). 78
6.2 Angular distance between the sensors mentioned and the sensors found by spherical sampling in the first experiment. 79
6.3 Results of the second experiment in the three different datasets. 79
6.4 Angular distance between the sensors mentioned and the sensors found by spherical sampling in the second experiment. 80
6.5 Results of the third experiment in the three different datasets. 83
6.6 Angular distance between the sensors mentioned and the sensors found by spherical sampling in the third experiment. 83

List of Figures

1.1 Colour information allows us to distinguish the danger 2
1.2 Visible light compared with the rest of electromagnetic information. 2
1.3 Different cone and rod responses over the spectrum. 3
1.4 Opponent process scheme. 4
1.5 (a) CIEXYZ colour matching functions (b) CIEXYZ chromaticity diagram (right). 5
1.6 Dichromatic model scheme. 6
1.7 Colour constancy effect. 7
1.8 Munsell chips used by Berlin and Kay. 9
1.9 Colour naming data 10

2.1 (a) Sigma Foveon D10 sensors, (b) Narrow-band sensors, (c) Sharp sensors. . 13
2.2 Relation between points in the sphere and sensors 17

3.1 Example of colour names categories 28
3.2 Example of our method 29
3.3 Different feasible solutions for the same scene providing different explanations of that scene. 33
3.4 Examples from a Real-World dataset 35
3.5 Examples from the Controlled indoor dataset 36

4.1 Images regularly selected in the experiment as natural (left) versus images hardly ever selected (right). 43
4.2 Experiment schedule 44
4.3 Image dataset under D65 illuminant. 45
4.4 Camera and grey sphere setup. 45
4.5 Reflectance of the paint used on the ball. 46
4.6 Comparison to the mean observer (black line). 49
4.7 Estimated perceptual angular error (between method estimations and preferred illuminants). 52
4.8 Angular error plot for the different methods on 415 images of the dataset. . . 53

5.1 General scheme of the PO biological model 60
5.2 Sharp sensors predicting psychophysical data. 63
5.3 Prediction of the PO singularity index on the WCS data through the hue axis (a) PO model results (b) sharp sensors. 67

5.4 Prediction of the PO singularity index on the WCS data through the hue-value axis (a) PO model results (b) sharp sensors results. 67
5.5 Prediction of the PO singularity index on the unique hues data on the x,y chromaticity plane (a) PO model results (b) sharp sensors results. 68
5.6 Prediction of the PO singularity index on the unique hues data on the x,y plane plane (a) PO model results (b) sharp sensors results. 68
5.7 Comparison of the WCS data (background) versus: (a) the PO's index, (b) our index (right). 69
5.8 WCS Data (background) versus (a)our normalised index and (b) our simplified index. 70
5.9 CSI index in the x, y space: (a) surface plot, (b) contour plot. 70
5.10 Unique hues obtained by our different singularity indices 71
5.11 Reflection properties of some images of the Foster and Nascimento dataset (left) and their singularity index (right.) 72

6.1 Datasets used in this chapter 77
6.2 Original sensors -dotted lines- compared to spherical sampling sensors -solid lines- for the second dataset in the first experiment. 78
6.3 Original sensors -dotted lines- compared to spherical sampling sensors -solid lines- for the second dataset in the second experiment. 80
6.4 (a) Basis functions for dataset 2 (b) Sharp functions for dataset 2. 82

A.1 Points sampled with the algorithm 90

Chapter 1

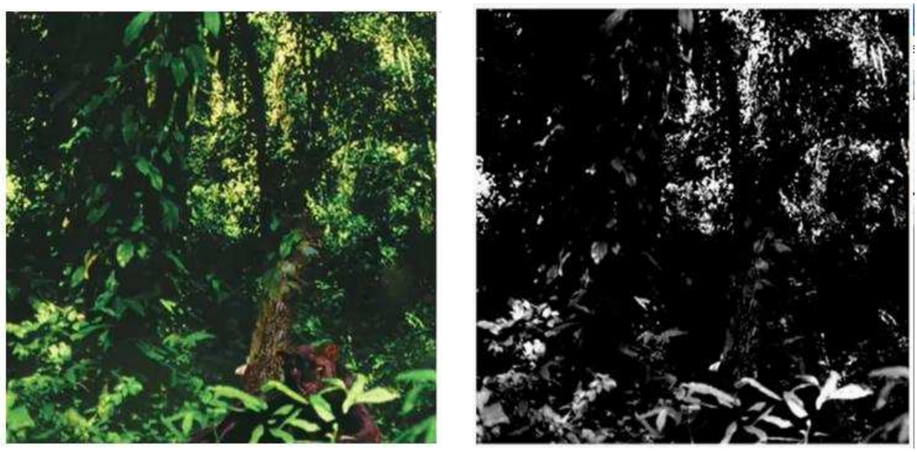
Introduction

The goal of this thesis is the study of computational colour constancy for natural images. Colour constancy is the ability of humans to perceive stable colours under different illuminants. Computationally, colour constancy is defined as the estimation of the scene illuminant given a single image that is an ill-posed problem. Colour constancy consists of estimating a 3×3 matrix which represents the illuminant. In order to reduce the complexity of the problem, colour constancy methods can be studied at three different stages that can be solved separately. The first stage is to build a narrow-band set of sensors to accomplish a diagonal model of illuminant change, which converts the original 3×3 matrix into a diagonal one, reducing in this way the number of unknowns to 3. Second stage consists of defining a set of possible solutions for the problem. Finally, last stage consists of estimating the best illuminant from the previous set. Particularly, in this work we tackle colour constancy problem by considering data derived from psychophysical colour naming experiments to the first and the third stages. The underlying idea is to use the information used in human representations as an evolutionary coding of natural statistics.

1.1 Colour science

Colour is a very important cue in our daily life and in our culture. Colour from an evolutionary point of view has allowed us to distinguish fruits in the forest or predators in the jungle. In Figure 1.1 we can see a colour image with a predator and the correspondent luminance image. Living in a world in greyscale would hinder our ability to interact with the environment. At night, when colour information disappears, it is more difficult to recognise dangers. From a cultural viewpoint, colour is inserted in many different communication codes. Green signals correlate with an allowance to proceed, while red are assigned to a prohibition.

The human visual system (from now on HVS) acts at different stages and not all of them are related to colour [95]. Here we roughly review the representation of colour from our eyes to our brain. The eye captures the light using different types of cells. Some cells react to different wavelengths, all of them located approximately between 400 nanometres



(a) (b)

Figure 1.1: (Image courtesy of Beau Lotto)(a) Original image (b) luminance image. Colour information allows us to distinguish the danger.

and 700 nanometres. We can see in Figure 1.2 that visible light is a very small part of the electromagnetic radiation, which range from gamma rays up to radio waves.

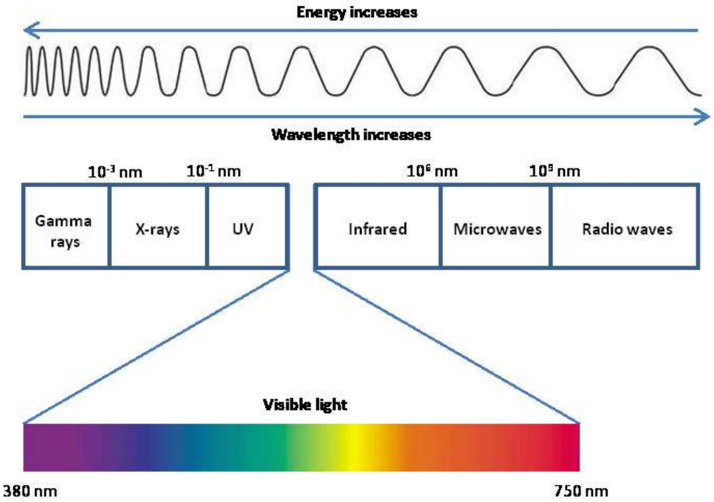


Figure 1.2: Visible light compared with the rest of electromagnetic information.

Light sensitive cells are located at the retina. Retina is usually explained as a tissue where images of the real world are projected. There are two different types of cells in the retina: the cones and the rods. Rods (R) are stimulated at low intensity light and get saturated before the cones. They basically extract intensity information and are usually not considered in colour perception. On the other hand, there are three different types of cones, reacting lo large (L), medium (M) and short (S) wavelengths. The spectral absorption of these cones depending on the wavelengths is shown in Figure 1.3. These different wavelengths are usually related for simplicity to the colours red for the L cones, green for the M cones and blue for the S cones.

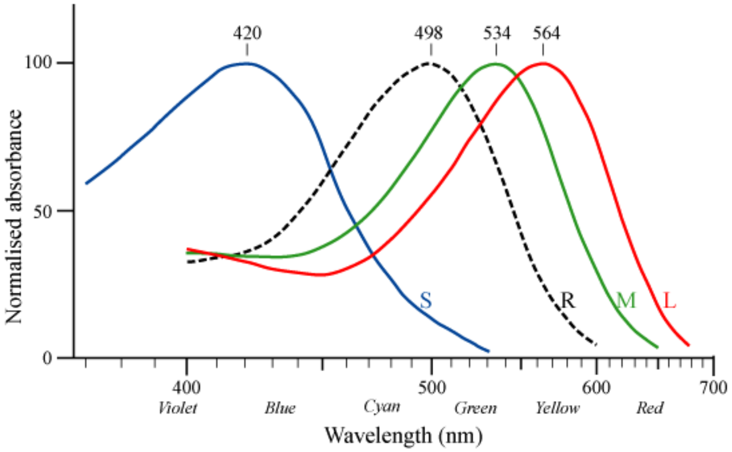


Figure 1.3: Different cone and rod responses over the spectrum.

Rod and cone cells are not the only cells dealing with colour. In the lateral geniculate nucleus (LGN) located in the thalamus (LGN is the primary processing center of our brain) is where the opponent cells are located [28]. These cells take the information given by the LMS cells and perform a centre-surround procedure. This step inserts an opponent coding acting as three channels: intensity, red versus green and blue versus yellow. A scheme of this combination is shown in Figure 1.4. Information from these opponent cells is later sent to the visual cortex. The full mechanisms for colour processing in the brain are not completely known yet.

Colorimetry is the field of physics that has mathematically specified colour perception. The beginning of this science can be related to the Newton observations [86], but, a big breakthrough appears with the matching experiments performed by Guild and Wright [56], [121]. These experiments were based on the trichromatic theory that states that colour can be modelled by only using three primaries [116] and considering the Grassman laws of colour addition [55]. A particular experiment was performed using a bipartite matching field of around two degrees of visual angle. The visual angle was selected in order to consider only the fovea, and therefore, to be able to discard any information coming from the rods. The experiment was defined as an observer-adjustable procedure: on one side of the field a test colour was projected and on the other side, an observer-adjustable colour was set. The experiment was performed by 7 observers [122]. From this experiment, the main institution in the field, the *Comission Internationale de l' Éclaraige* (CIE) defined the colour matching functions $\overline{r}, \overline{g}, \overline{b}$ for a standard observer. From these functions they define three different tristimulus values R,G, and B quantifying the amount of each colour perceived by a standard observer from a surface with reflectance $S(\lambda)$ and illuminant $E(\lambda)$. This space is known as

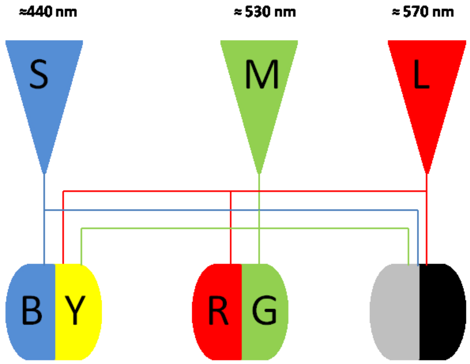


Figure 1.4: Opponent process scheme.

$CIERGB$ and is given by.

$$\begin{aligned} R &= \int E(\lambda)S(\lambda)\overline{r}(\lambda)d\lambda \\ G &= \int E(\lambda)S(\lambda)\overline{g}(\lambda)d\lambda \quad (1.1) \\ B &= \int E(\lambda)S(\lambda)\overline{b}(\lambda)d\lambda \end{aligned}$$

(1.2)

However, these colour matching functions have some problems, a major one being the existence of negative values. From this problem CIEXYZ space was derived. This space was committed to accomplish two specific goals: having all positive response in the colour-matching functions $\overline{x}, \overline{y}, \overline{z}$, and, having one of the colour-matching functions to be equal to the photopic response function $CIE\ V(\lambda)$. These functions are shown in Figure 1.5.(a) Then, the different tristimulues values for X,Y, and Z are given by

$$\begin{aligned} X &= \int E(\lambda)S(\lambda)\overline{x}(\lambda)d\lambda \\ Y &= \int E(\lambda)S(\lambda)\overline{y}(\lambda)d\lambda \quad (1.3) \\ Z &= \int E(\lambda)S(\lambda)\overline{z}(\lambda)d\lambda \end{aligned}$$

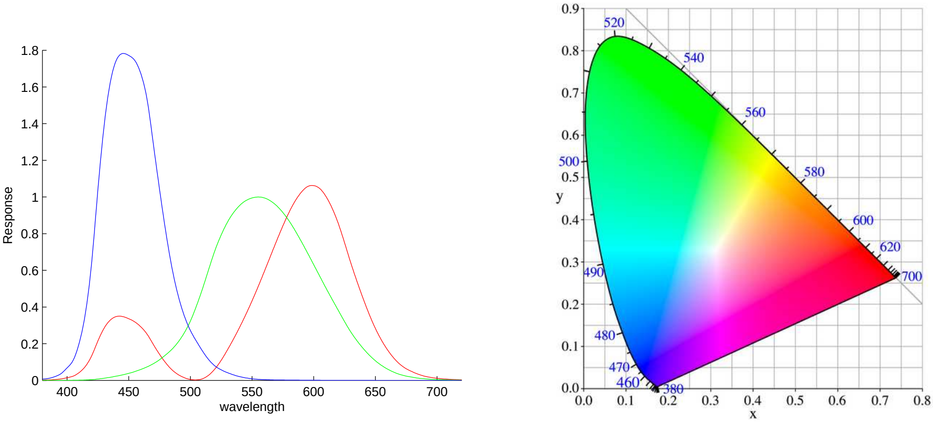
(1.4)

From this CIEXYZ space, a chromaticity space can be defined by projecting onto the $X + Y + Z = 1$ plane and avoiding one of these dimensions as

$$\begin{aligned} x &= \frac{X}{X+Y+Z} \\ y &= \frac{Y}{X+Y+Z} \quad (1.5) \end{aligned}$$

(1.6)

This chromaticity diagram (Figure 1.5.(b)) will represent a 2D projection of the gamut of all different colours humans can see, since $\overline{x}, \overline{y}, \overline{z}$ represent a standard observer.



(a) (b)

Figure 1.5: (a) CIEXYZ colour matching functions (b) CIEXYZ chromaticity diagram (right).

CIEXYZ space has been shown as a very useful space in terms of colour parametrisation. However, it has a main drawback: it does not correlate with human similarity perception. In order to counter this problem, CIE has been defining different spaces: CIELAB [98], CIECAM97 [81] and CIECAM02 [96]. All these spaces are perceptually uniform, this means, distances in the space correspond to perceptual distances. These spaces act as chromatic adaptation models since they require to know the colour of the illuminant, that is, the colour of the white point. Once the white point is obtained, these spaces are defined to be completely independent from original light source. Every space has been defined trying to overcome problems detected in the previous ones [31].

1.2 Image formation

Computer vision is a sub-field of artificial intelligence that tries to emulate human tasks via the definition of specific engines. These tasks can vary from object recognition, object tracking or object segmentation. In these tasks colour plays an important role, and then is a key feature to be used in computer vision. Computational colour relies on an image usually acquired with cameras based on three basic sensors. Trichromatic cameras capture the light reflected by the scene using three different sensors which react to different physical wavelengths. However, spectral information is not the only information that influences the image formation process, also the geometry of the different surfaces in the scene interact with the process.

A model capturing all these effects was defined by Shafer in [101]. This dichromatic model describes the colour of a scene as a combination of diffuse and specular components. These components are defined by the surface of the object, denoted as surface albedo (diffuse component) and the colour of the specularities (specular component). Mathematically, the model is defined as

$$\rho_i = m_b(\underline{n}, \underline{s}) \int E(\lambda)C_b(\lambda)R_i(\lambda)d\lambda + m_s(\underline{n}, \underline{s}, \underline{v}) \int E(\lambda)C_s(\lambda)R_i(\lambda)d\lambda \qquad (1.7)$$

where $R_i(\lambda)$ represents the i-th sensor, $E(\lambda)$ is the illuminant spectrum, $C_b(\lambda)$ is the surface albedo, $C_s(\lambda)$ the specular colour. Finally, $\underline{n}$ is the surface patch normal, $\underline{s}$ the direction of the incident light and $\underline{v}$ is the direction of the viewer. A scheme of this model is found in Figure 1.6

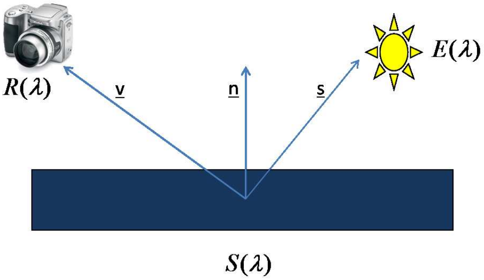


Figure 1.6: Dichromatic model scheme.

The main problem of the dichromatic model is its complexity. For this reason, for computer vision applications the geometry of the image and the specular component is usually discarded , focusing only in the information given by the albedo.

$$\rho_i = \int E(\lambda)S(\lambda)R_i(\lambda)d\lambda \qquad (1.8)$$

where colour representation only depends on the sensor sensitivities, the colour of the illuminant conditions and the colour of the surface.

1.3 Colour constancy

One of the main properties of the human visual system is its colour constancy ability. How colour constancy is achieved in HVS it is still not completely understood. It is usually defined as the tendency of objects to appear the same colour even under changing illumination [66]. Several studies widely agree that colour constancy is not based on a single mechanism, and these mechanisms spread over the different neural levels. In order to locate different mechanisms in the scheme, Hurlbert and Wolf [67] defined three different classes:

Sensory: These mechanisms only require a simple linear transformation of the photoreceptors. Some authors defend that colour constancy is mainly achieved at this level [113]. Main sensory mechanisms are reduced to: chromatic adaptation to the mean, chromatic adaptation to the variance and spatial contrast. This last mechanism is strongly supported by the results obtained by Foster *et al.* [45]. These results state that cone-photoreceptor ratio excitations produced by light from natural surfaces are statistically almost invariant under changes in illumination due to natural light. Therefore, they keep the cone-excitation ratios invariant from the surfaces across illuminant changes helping in solving colour constancy.

Perceptual: These mechanisms require prior segmentation in order to find relevant scene elements. Some perceptual cues are: 3D shapes [60], depth perception [118], mutual reflections, chroma variance, specularities [71] and scene movement [119]

Cognitive: These mechanisms require recognition of objects in the scene. In this class we have memory colour, that states that we tend to perceive an object's colour influenced by the colour stored in our memory. For example, we will see a banana as yellow, sky as blue or grass as green [87]. Memory colour was firstly introduced by Hering in the 19th century [61]. Another cognitive cue will be the conciousness of illuminant change [4]. However, conciousness of illuminant change will be in the border between cognitive and perceptual aspects, since it is related to the scene movement.

Experiments in colour constancy have been done using different paradigms. One of the most widely used is defined via the adjustment of a particular patch (called test patch) displayed or projected under different illuminant conditions with a specific background. The adjustment of the test patch can be done following different approaches [103]. *Asymmetric matching* consists of adjusting the test patch viewed under one illuminant to another patch in a second stimulus viewed under a different illuminant. Both stimulus can be seen simultaneously, that is, both appeared at the same time, or successively. *Achromatic matching* consists of transforming the colour of the test patch to an achromatic colour. *Colour naming* consists of converting the test patch onto one of a particular colour category.

1.4 Computational colour constancy

In this work we will focus on solving colour constancy from a computational point of view. As we introduced previously, colour constancy is often defined as the ability of perceiving colours stable under different lights. Therefore, we need to represent a banana yellow both at midday, when the illumination is mainly bluish and at sunset, when the illumination is reddish. Unfortunately, this information is not coped by a camera device. For example, in Figure 1.7 we can perceive that the t-shirt worn by the woman in the middle is perceived as orangish in the context of both images, although we can see that in isolated conditions the t-shirt it purely green under the turquoise illuminant.

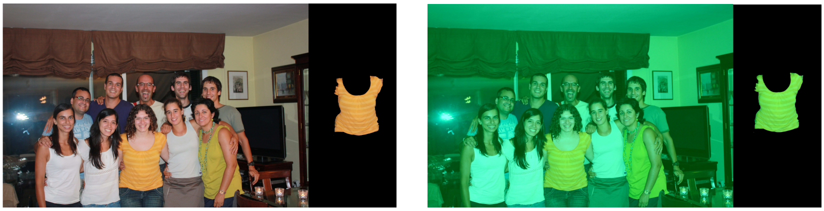


Figure 1.7: Colour constancy effect.

Computational colour constancy is an under-constrained problem and therefore it does not have a unique solution. From a mathematical point of view, let us suppose we have an object with reflectance $S(\lambda)$ viewed under two illuminants $E_1(\lambda)$, $E_2(\lambda)$ and captured by a camera with sensitivities $R_i(\lambda)$, $i \in \{1, 2, 3\}$. Then, the colours captured by the camera are denoted as $\underline{\rho}^1$ and $\underline{\rho}^2$, where their components are given by

$$\begin{aligned} \underline{\rho}_i^1 &= \int S(\lambda)E_1(\lambda)R_i(\lambda)d\lambda \\ \underline{\rho}_i^2 &= \int S(\lambda)E_2(\lambda)R_i(\lambda)d\lambda \qquad (1.9) \\ & \qquad (1.10) \end{aligned}$$

The relation between these two values is usually modelled by a 3×3 matrix. That means, we tackle it as a 9-dimensional problem.

$$\underline{\rho}^1 = \begin{pmatrix} \alpha & \delta & \epsilon \\ \zeta & \beta & \eta \\ \theta & \vartheta & \gamma \end{pmatrix} \cdot \underline{\rho}^2 \tag{1.11}$$

Equation 1.11 is crucial in order to solve computational colour constancy. Without loss of generality we can consider $E^1(\lambda)$ as the canonical illuminant. The canonical illuminant is the illuminant under which the colours appear in a trustworthy form (usually a white illuminant). Then, if we wish to discount the colour of the illuminant from the image of a scene it suffices to find the mapping that removes this colour. As an example, if an image is captured under bluish light then all the recorded sensor responses are biased in the blue direction and, in particular, a white surface will itself be bluish. If we can find the map that goes from the blue light to a white counterpart then applying this map will remove the colour cast. However, equation 1.11 has 9 components in a 3×3 matrix and colour constancy, viewed from this viewpoint, is still a 9-dimensional problem.

A lot of research have been devoted to solve colour constancy from many different perspectives but there is not still a widely accepted solution. In this thesis we propose to frame all colour constancy approaches as a three-stage process.

First step deals with the *computation of a new basis of sensors* in order to accomplish the so-called von Kries or diagonal model of illuminant change [120], [75]. This means, accomplishing that the illuminant change can be modelled as

$$\underline{\rho}^1 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \cdot \underline{\rho}^2 \tag{1.12}$$

Equation 1.12 reduces the complexity of the problem by supposing that illumination change is a process which operates in each sensor response channel independently. Then, the 3×3 original matrix is converted to a diagonal one greatly simplifying colour constancy computation.

Second step deals with the *definition of the set of plausible illuminants*. This set is called the feasible set. Plausible illuminants are determined by restrictions usually related to the colour values that the corrected image can take. For example the gamut of the colours in the output image should be inside the gamut of possible colours.

Finally, third step deals with the *selection of the best illuminant map from the feasible set*. Different methods have been designed to perform this selection ranging from statistical to physical approaches.

A review of different methods following this framework is found in section 2. We define this framework in order to locate the contributions of the thesis. In this thesis we have been working in the first and the third steps. In both cases we have developed our contributions by the use of colour naming data learned psychophysically.

1.5 Colour naming

Different psychophysical experiments have been proposed to explain the special responses to some particular colours in our brain. The most influential one related to colour naming is still the one performed by Berlin and Kay [12]. In this work, Berlin and Kay asked speakers of 20 different languages of industrialised cultures to name different patches of the Munsell book of colour. The selected patches consist of a set of Munsell reflectances spanning the gamut of typical colour names. It consists of 40 hues spanning the colour circle, each printed in 8 values. Those patches are usually shown as a 2D grid of colours where neighbouring patches have similar hue and saturation (see Figure 1.8). Then, they analysed those patches and extracted the conclusion that most evolved languages shared 11 colour terms. These terms are: *red, green, blue, yellow, orange, brown, pink* and *purple* as chromatic colours and *white, grey* and *black* as achromatic. Furthermore, they analysed written data for other 78 languages and also concluded that the number of colour terms is related to the evolution of the culture associated.

Some further research based on this topic have been done by Sturges *et al.* [106], Boynton *et al.* [17] and Benavente *et al.* [10] in order to locate the colours consistently named in a reference space.

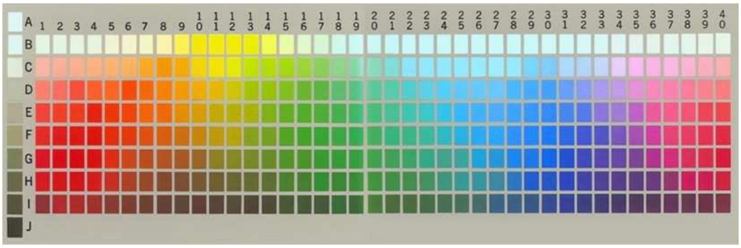
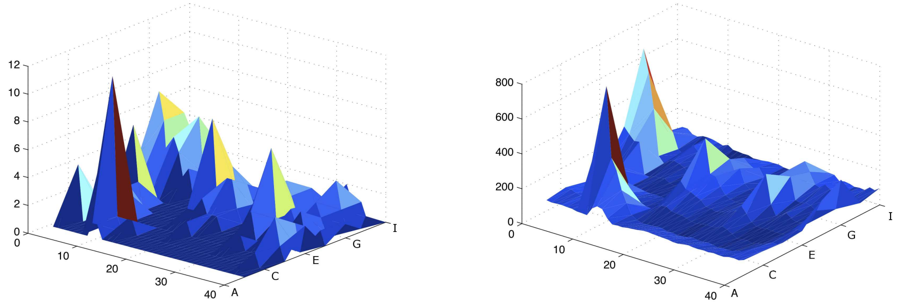


Figure 1.8: Munsell chips used by Berlin and Kay.

Berlin and Kay experiment has been further developed by Cook, Kay and Regier. Following the same procedure as that in the previous experiment, they launched the World Colour Survey (WCS) data experiment [1]. In this new experiment they analysed 110 unwritten languages, using 24 native speakers and 330 Munsell chips (40 of them are the Berlin and Kay ones). However, in this case, the experiment was not performed under controlled conditions. The results of this experiment show that four universal chromatic colours arise: *red, green, blue* and *yellow.* These colours also correspond to the unique hues found in some physiological studies [73], [110]. In Figure 1.9 we plot the difference between Berlin and Kay and WCS results.

Recently, new colour naming experiments are being done in on-line platforms [83], [82]. This new paradigm allows to reach people from very different scenarios which means uncon-



(a) (b)

Figure 1.9: (a) Berlin and Kay psychophysical data. (b)WCS psychophysical data.

trolled conditions onto the stimulus but a higher number of subjects.

1.6 Thesis outline

In this thesis, we focus on computational colour constancy for natural images. The main novelty in this work is the use of colour naming psychophysical data (see section 1.5) in order to deal with colour constancy for natural images. Particularly, two main contributions are presented. To report the work we have divided the thesis in two parts.

The first part of the thesis is devoted to the problem of Illuminant Selection. We define a new colour constancy method which uses the basic colour term categories defined by Berlin and Kay [12] as anchor points. This method will be explained in chapter 3. In chapter 4 we have explored on selection criteria with a psychophysical experiment to test if usual physical colour constancy measures adapt to human preferences.

The second part of the thesis, titled Sensor Sharpening, starts in chapter 5 where we prove that by using some particular sharp sensors we can predict both results from the WCS data and from the unique hues data. Moreover, the relation between these data and the sensors is based on an achromatic measure. Finally, in chapter 6 we show how by using the spherical sampling technique, we can find sharp sensors minimising a range of perceptual measures.

Chapter 2

Computational colour constancy: a review

In this section we review some of the most important works on the different stages of colour constancy introduced on the previous chapter and which are in the basis of our work. Firstly, we explain the different methods that have been used to provide sharper (more narrow-band) sensors that allow the use of a diagonal model for illuminant change. We review in depth four different sharpening methods and the spherical sampling technique that is a key step for final improvements. Secondly, we review different approaches to construct a feasible set of illuminants either if it is explicitly computed or implicitly used in different methods. Finally, different approaches defined to estimate the colour of the illuminant are surveyed. They are divided into statistical, probabilistic, physical and criteria-combination methods.

2.1 Providing a diagonal model

In the previous chapter we have outlined the von Kries law and the diagonal model for colour constancy. They state that the illuminant change can be modelled by a process which operates in each sensor response channel independently, that is,

$$\underline{\rho}^1 = D^{1,2}\underline{\rho}^2 \tag{2.1}$$

Then, the 3×3 original matrix is converted to a diagonal one. However, most of the camera sensors are very broad functions and the diagonal model does not perform properly in this case. Then, the question that arises if it exists a set of sensors accomplishing equation 2.1.

The narrow-band sensors are the answer to this question. The narrow-band sensors are such that for each wavelength only one of them is sensitive. Let us prove that under the narrow-band assumption equation 2.1 holds. To this end, we suppose the extreme case where sensors are only sensitive to a single wavelength (Dirac Delta functions). Then, we consider the wavelength λ_1 to which only the sensor $R_1(\lambda)$ is sensitive. In this case, equation 1.11 can be written as

$$\begin{pmatrix} E^1(\lambda_1)S(\lambda_1)R_1(\lambda_1) \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha & \delta & \epsilon \\ \zeta & \beta & \eta \\ \theta & \vartheta & \gamma \end{pmatrix} \cdot \begin{pmatrix} E^2(\lambda_1)S(\lambda_1)R_1(\lambda_1) \\ 0 \\ 0 \end{pmatrix} \tag{2.2}$$

.

From this equation we find the following three equalities

$$\begin{aligned} E^1(\lambda_1)S(\lambda_1)R_1(\lambda_1) &= \alpha \cdot E^2(\lambda_1)S(\lambda_1)R_1(\lambda_1) \\ 0 &= \zeta \cdot E^2(\lambda_1)S(\lambda_1)R_1(\lambda_1) \\ 0 &= \theta \cdot E^2(\lambda_1)S(\lambda_1)R_1(\lambda_1) \end{aligned} \tag{2.3}$$

.

Then,

$$\alpha = \frac{E^1(\lambda_1)}{E^2(\lambda_1)} \text{ , } \zeta = 0 \text{ and } \theta = 0$$

.

Applying the same procedure to the wavelengths λ_2 and λ_3 were the second and third sensors are sensitive we find $\delta = \epsilon = \eta = \vartheta = 0$ and

$$\beta = \frac{E^1(\lambda_2)}{E^2(\lambda_2)} \text{ and } \gamma = \frac{E^1(\lambda_3)}{E^2(\lambda_3)}$$

and then the diagonal model holds.

We have proven that in the case of a set of narrow-band sensors the diagonal model holds. But, how to convert the camera sensors into a set of sensors more narrow-band? The solution to this problem has come through the *spectral sharpening* technique that is based on the idea that there might exist a privileged linear combination of the camera sensors with respect to which Equation 2.1 works. We can call this approach *transformed diagonal mapping*. The remarkable and useful conclusion of the spectral sharpening work was that even for broad-band sensor systems a diagonal matrix model of illuminant change could be used. Mathematically, the equation defining a transformed diagonal mapping is:

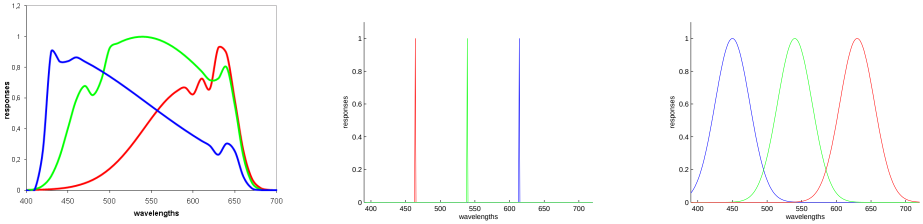
$$T\underline{p}_r^1 \approx D^{1,2} T\underline{p}_r^2 \tag{2.4}$$

where r is defining different surface reflectances, T is usually referred to as sharpening matrix and its respective sensors are referred to as sharp sensors [35]. This equation has been solved in different ways, each of them leading to a different matrix T, and then defining a different set of sharp sensors. In Figure 2.1 we can see (a) the Sigma Foveon D10 sensors, (b) a set of narrow-band sensors and (c) a set of sharp sensors.

In this chapter we will explain four different methods defined to compute the matrix T, and a final improvement in order to refine the results. We will also introduce some psychophysically measured matrices that can be used as sharpening matrices. Finally, we will define the usual error measure used to compare these methods.

2.1.1 Sensor-based spectral sharpening

Finlayson *et al.* [35] first defined the so-called Sensor-based spectral sharpening. The idea underlying this method is that a sensor $R(\lambda)$ can be sharpened in an interval $[\lambda_1, \lambda_2]$ denoting



(a) (b) (c)

Figure 2.1: (a) Sigma Foveon D10 sensors, (b) Narrow-band sensors, (c) Sharp sensors.

wavelengths. The resulted sharp sensor can be denoted as $R^t(\lambda)\underline{c}$, $\underline{c}$ being a coefficient vector. It can be found by minimising

$$\min(\int_{\Phi}[R(\lambda)^t\underline{c}]^2 d\lambda + \mu\{\int_{\omega}[R(\lambda)^t\underline{c}]^2 d\lambda - 1\}) \tag{2.5}$$

where ω is the visible spectrum, Φ denotes wavelengths outside $[\lambda_1, \lambda_2]$ and μ is a Lagrange multiplier that constrains the integral of the resulting sensor to be 1. In other words, they strengthen the percentage of the norm of sensor $R(\lambda)$ lying in the interval $[\lambda_1, \lambda_2]$ in relation to all the other intervals.

To solve this problem for all the spectra the authors propose to define K intervals where K is the number of sensors (in our case $K = 3$). These intervals do not intersect with each other and cover all the spectra. Then, the kth column of matrix T will be the one that minimises the equation 2.5 for this particular interval.

2.1.2 Perfect sharpening

Another approximation to tackle the problem is perfect sharpening. Perfect sharpening is based on the work of Marimont and Wandell [80] where they explain the idea of illuminant and reflectance dimensionality. A set of surface reflectances is n-dimensional if we can write any reflectance S of the set as a linear combination of n basis functions S_i

$$S(\lambda) = \sum_{i=1}^{n} \sigma_i S_i(\lambda) \tag{2.6}$$

Similarly, a set of illuminants is m-dimensional if we can define m basis functions E_i that accomplish

$$E(\lambda) = \sum_{i=1}^{m} \epsilon_i E_i(\lambda) \tag{2.7}$$

Based on this assumption, Finlayson *et al.* [33] have shown that if illuminants are two-dimensional and reflectances three-dimensional, or vice versa, spectral sharpening is perfect.

Let us suppose that the set of reflectances is three dimensional ($n = 3$) and the set of illuminants is two dimensional ($n = 2$). First, we need to remark, that under the assumptions stated above, a colour descriptor (Eq. 1.9) under a light E^1 can be written using a lighting matrix $\Lambda(\epsilon)$

$$\underline{p}^1 = \Lambda(\epsilon)\underline{\sigma} \tag{2.8}$$

$\underline{\sigma}$ being the vector obtained applying equation 2.6 to the surface and being $\Lambda(\epsilon)_{ij}$ a 3×3 matrix

$$\Lambda(\epsilon)_{ij} = \int S_i(\lambda)E^1(\lambda)R_j(\lambda) \tag{2.9}$$

with $E^1(\lambda)$ computed following equation 2.7.

Then, without loss of generality, we can define E^1 as the canonical light. As illuminants are two dimensional we need a second illuminant $E^2(\lambda)$ independent from the canonical one $E^1(\lambda)$ to span the space.

Associated with this illuminant will also be a new lighting matrix Λ^2. This second lighting matrix is some linear transform away from the first one, $\Lambda^2 = M\Lambda^1$, that is, $M = \Lambda^2[\Lambda^1]^{-1}$.

Therefore, as $E^2(\lambda)$ and $E^1(\lambda)$ span the space, any other lighting matrix will be a combination of Λ^1 and $M\Lambda^1$. For this reason, any colour descriptor under an illuminant $E^e(\lambda) = \alpha E^1(\lambda) + \beta E^2(\lambda)$ can be written as

$$\underline{p}^e = [\alpha I + \beta M]\Lambda^1\underline{\sigma} = [\alpha I + \beta M]\underline{p}^1 \tag{2.10}$$

where I is the identity matrix. Calculating the eigenvector decomposition of M

$$M = T^{-1}DT \tag{2.11}$$

and expressing the identity matrix in terms of T, $I = T^{-1}IT$, they rewrite equation 2.10 as a diagonal transform

$$T\underline{p}^e = [\alpha I + \beta D]T\underline{p}^1 \tag{2.12}$$

Finally, writing $\underline{p}^1$ in terms of $\underline{p}^e$

$$T\underline{p}^1 = [\alpha I + \beta D]^{-1}T\underline{p}^e \tag{2.13}$$

Finlayson *et al.* also shown in that paper that this theoretical perfect sharpening method works in practice.

2.1.3 Data-based sharpening for a pair of illuminants

Finlayson *et al.* [35] defined the data based sharpening which uses linear algebra methods to directly solve for T by minimising the residual error between a pair of illuminants. To this end, they define W^1 and W^2 as $3 \times N$ matrices containing the colour values for a set of N different reflectances under two different illuminants E_1 and E_2.

$$TW^1 \approx D^{1,2}TW^2 \tag{2.14}$$

where $D^{1,2}$ is a diagonal matrix. Then, they solve this equation for $D^{1,2}$ in a least-squares sense. This can be done by the Moore-Penrose inverse

$$D^{1,2} = TW^1\left[TW^2\right]^+ \tag{2.15}$$

where $[]^+$ represent the pseudoinverse (see [54]). Rearranging equation 2.15, they arrive to

$$T^{-1}D^{1,2}T = W^1[W^2]^+ \tag{2.16}$$

Therefore, if U and V, are respectively the eigenvector and eigenvalue decompositions of $W^1[W^2]^+$ then $T = U^{-1}$.

This method has a clear drawback: it only searches a solution for a pair of illuminants. This means that, if we change one of the illuminants, the matrix T will change. For this reason, Barnard *et al.* [6] proposed to introduce more flexibility onto the data-based sharpening, by averaging over a set of illuminants (not only one), and introducing a parameter to prioritise positivity.

2.1.4 Measurement tensor method

Chong *et al.* [25] , introduced a new data-based method which finds a matrix T for a complete set of illuminants at the same time. This method is based on the measurement tensor. The measurement tensor is defined as

$$M_{kij} = \int R_k(\lambda)E_i(\lambda)S_j(\lambda) \tag{2.17}$$

where $\{E_i\}_{i=1,\cdots,I}$ is a set of illuminants, $\{S_j\}_{j=1,\cdots,J}$ is a set of reflectances and $\{R_k\}_{k=1,\cdots,K}$ are sensors. To solve the problem, they slice the measurement tensor by constant k, finding different $I \times J$ matrices $\Gamma_k = M_{kij}$. After that, they prove:

- A measurement tensor supports colour constancy if and only if there exists a basis, that, for all k, Γ_k is a *rank* 1 matrix
- A measurement tensor supports colour constancy if and only if it is a rank 3 tensor

where an order 3 tensor τ is rank N if N is the smallest integer such that there exist vectors $\{\underline{a_n}, \underline{b_n}, \underline{c_n}\}_{n=1,\cdots,N}$ allowing decomposition as the sum of outer products

$$\tau = \sum_{n=1}^{N} \underline{c_n} \circ \underline{a_n} \circ \underline{b_n} \tag{2.18}$$

We can rewrite this equation with our measurement M_{kij} tensor as

$$M_{kij} = \sum_{n=1}^{N} C \circ A \circ B \tag{2.19}$$

Where the columns of A, B and C are composed by the different $\underline{a_n}$, $\underline{b_n}$, and $\underline{c_n}$ respectively and C is the colour basis and therefore, $T = C^{-1}$.

Then, the sharpening problem becomes equivalent to search for the M_{kij} rank 3 best approximation. In order to find this rank 3 approximation authors apply a Trilinear Alternate Least Squares (TALS) method [59] to equation 2.19. This method has the drawback of local convergence. Also, TALS method needs initialisation values for two of their three matrices, usually C being one of them.

2.1.5 Spherical sampling

Spherical sampling defines all the different sensors on a combinatorial solution. Furthermore, it also finds the geometrical relation between them. This is possible by defining a relation between each sensor and a point in the n-sphere which allows the method to refine sharpening matrices found by other methods.

We will define the method in an n-dimensional formulation (until now, it was usually defined for 3 dimensions [39]). The definition of this technique as n-dimensional is due to the fact that we will use it in chapter 6.

Spherical sampling can be defined as follows. Let us represent our sensors R as an $m \times n$ matrix where m is the wavelength sampling and n the number of sensors in an orthogonal basis. To this end, we can use the singular value decomposition (SVD) of these sensors

$$R = U \cdot \Sigma \cdot V^t \tag{2.20}$$

where U is an orthogonal matrix with dimension $m \times n$, Σ is a diagonal $n \times n$ matrix containing the singular values of matrix R and V^t is an orthogonal $n \times n$ matrix. Then, U is the basis we seek.

From this basis U, we can define a new set of sensors $\mathcal{R}$ $(m \times n)$, different from the original sensors R, by multiplying the basis by any linear transformation P $(n \times n)$, which simply consists of n sample points vectors, $\underline{p_1}, \cdots, \underline{p_n}$ located over the n-sphere. Then,

$$\mathcal{R} = UP, \quad P = [\underline{p_1}, \cdots, \underline{p_n}]; \tag{2.21}$$

We are interested in the relation between the original sensors R and the newly defined ones $\mathcal{R}$. Using equation 2.20 and 2.21 we find

$$\mathcal{R} = UP = U\Sigma V^t(\Sigma V^t)^{-1}P = R(\Sigma V^t)^{-1}P. \tag{2.22}$$

Therefore, relating this equation to the previous methods we have

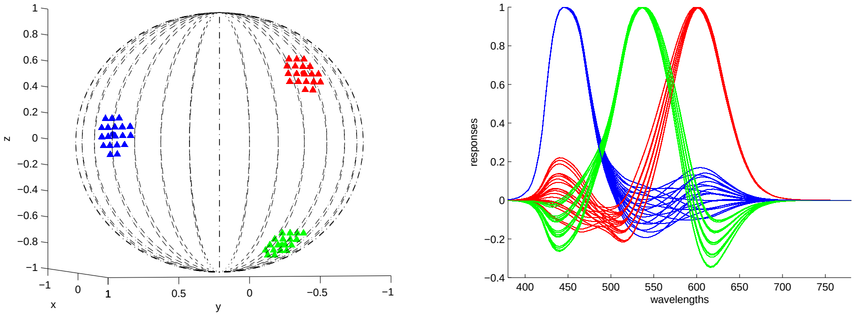
$$T = ((\Sigma V^t)^{-1}P)^t. \tag{2.23}$$

We can also rearrange this equation in order to relate a transformation matrix T with a set of points P over the sphere.

$$P = \Sigma V^t T^t. \tag{2.24}$$

These last two equations show that we can relate a matrix T found by using any previous methods to a set of points P located over the sphere. Then we can refine this matrix by sampling points around the set P and computing back their matrix T. An example of this for the case $n = 3$ is shown in Figure 2.2(a) where a set of points separated by a 10 degrees distance is plotted and in Figure 2.2(b) where their corresponding sensors are plotted.

The formulation defined in this section means that in order to find all the different sensors we only need to sample the different values P located over the n-sphere. This problem has been widely studied in the 3-dimensional case [94], while its general n-dimensional formulation (with $n > 3$) has not received as much attention. In this thesis we will use the method of Lovisolo and Da Silva [77], which is able to work in n-dimensional spaces. We mimic the explanation of this method in the appendix.



(a) (b)

Figure 2.2: Relation between points in the sphere and sensors (a) 10 degrees samplings of points over the sphere and (b) sensor sensitivities corresponding to the samplings in (a)

2.1.6 Other approaches

All the previous methods for sharp sensors have been defined in order to help in solving for colour constancy. But, sharpening matrices related to these methods are not unique. Other matrices have been derived from psychophysical experiments in order to do chromatic adaptation. Some examples are Bradford transform, Fairchild transform, CAT02 transform or HPE transform [31]. These matrices also represent sharp sensors as was proved by Finlayson and Susstrunk [39].

2.1.7 Usual error measure

To evaluate how well a particular sharpening method set works we will adopt the method presented in [35]. A reference light E^1 is chosen and the RGB values for a reference sensor set are calculated for a large set of reflectances $\underline{p}_r^{E^1}$. The reference sensor set could be the XYZ colour matching functions or the sensitivities of a camera. These RGB values, henceforth called descriptors, are the correct answer to colour constancy.

Given the corresponding RGB values under a second light E we find the diagonal matrix D that maps these back to the reference lighting conditions. Of course the diagonal model is not perfect and so we expect an error between each mapped RGB and its corresponding descriptor. This percentage error is calculated as:

$$Error = 100 \text{ x } \frac{\|\underline{p}_r^{E^1} - D\underline{p}_r^{E}\|}{\|\underline{p}_r^{E^1}\|} \tag{2.25}$$

In all our experiments the diagonal components of D are simply the ratio of the sensor responses of the descriptor for a perfect white diffuser divided by the same white surface viewed under the second light. Of course we can calculate an analogous error given a sharpening matrix T:

$$Error = 100 \text{ x } \frac{\|\underline{p}_r^{E^1} - T^{-1}DT\underline{p}_r^{E}\|}{\|\underline{p}_r^{E^1}\|} \tag{2.26}$$

Now D is again calculated as the ratio of RGB values for a white surface under the two lights but now with respect to sensors transformed by T. The above error is calculated for many surfaces under many lights.

2.2 Building the feasible set

The construction of a feasible set of solutions is a usual stage in computational constancy methods. The concept was firstly referred in the seminal work of Forsyth [42], where the idea is introduced as the basic goal to solve the problem. It copes with the set of all possible solutions, namely, the set of all possible mappings representing feasible illuminant changes. Depending on the dimension of the mappings the feasible set can be in a 3D space for the case of the diagonal model or a 9D space for full linear models.

The feasible set is explicitly computed in some methods such as *Color in Perspective* [32], *Edge Gamut Mapping* [53] or *Gamut Contrained Illuminant Estimation* [37], where their differences are derived from the differences in the assumptions that will constraint this feasible set. Constraints are either applied onto the set of illuminants (Forsyth [42]) or onto the set of likely reflectances (Finlayson [32]). Whatever assumption is used, the aim is always to reduce the size of the feasible sets. There are some works focusing on a efficient computation of the feasible set (Finlayson *et al.* [38]).

In some methods, such as *Grey-World* [21], *Shades-of-Grey* [40] or *Grey-Edge* [111] the feasible set is not explicitly computed since they are based on assumptions addressed to establish a selection criterion onto the set instead of trying to reduce its size. An important distinction when computing the feasible set is derived from the conditions we are working on.

Under calibrated conditions, that means, for the case where the camera sensitivities of the acquisition device are known, the sensor sensitivities are the main constraint in reducing the set of likely solutions. In this case the feasible set is defined by the set of illuminants that map the gamut of the original image inside the canonical gamut. Let $G(C)$ be the canonical gamut provided by all the reflectances, the canonical illuminant and the sensor sensitivities and I the original image represented as a $3 \times N$ matrix. Then the feasible set is defined as the set of different mappings $\underline{m}$ such as

$$FS = \{\underline{m} : G(\underline{m} \cdot I) \subset G(C)\} \tag{2.27}$$

where $G(\underline{m} \cdot I)$ is the gamut of the image corrected by the mapping $\underline{m}$ which is related to a specific illuminant change. In this case the $\cdot$ operation stands for the application of a 1×3 map onto all the image points. In the case of non-calibrated conditions, the feasible set is defined as

$$FS = \{\underline{e} : G(\underline{m} \cdot I) \in G(RGB)\} \tag{2.28}$$

where $G(RGB)$ is the complete sensor-dependent cube.

2.3 Illuminant selection

In section 2.1 we have reviewed different methods that might sharpen the sensors to achieve a more accurate use of the diagonal model. In section 2.2 we have reviewed the methods that explicitly build the feasible set of mappings. In this section we finish the review by introducing the rest of the methods organised according to the type of criteria they use to finally select an illuminant. We propose a taxonomy of methods or selection criteria that is divided in four groups: statistical, probabilistic, physical and those based on criteria-combination.

Statistical methods are based on basic criteria assuming some statistical properties on the image content. In this group we have the criteria used in gamut-mapping methods. The most visited approximations are *Maximum Volume* and *Average*. The former selects the illuminant $\hat{e}$ whose associated mapping maximises the image volume. Assuming the diagonal model this is equal to

$$\underline{\hat{m}} = \underline{m}_{\arg\max_i \underline{m}_{i,1} \cdot \underline{m}_{i,2} \cdot \underline{m}_{i,3}} \tag{2.29}$$

where in $\underline{m}_{i,k}$ i denotes the different mappings in the FS and k the different colour channels. On the other hand, the *Average* heuristic selects the illuminant associated to the average mapping of the FS.

$$\underline{\hat{m}} = mean(\underline{m}_{i,1}, \underline{m}_{i,2}, \underline{m}_{i,3}) \tag{2.30}$$

Other criteria similar to previous ones have been those related to the computation of a norm, usually L_2 and L_1.

In this group of methods we also have those methods that rely on properties of the scenes instead of on the properties of the corrected images. For example, *Grey-World* [21], which assumes that scenes are grey in mean. Given an image I as a $3 \times N$ matrix, where each row, I_i, represent a colour channel it can be denoted as $I = (I_1, I_2, I_3)^T$. The estimated illuminant is given by

$$\underline{\hat{e}} = (mean(I_1), mean(I_2), mean(I_3)) \tag{2.31}$$

A similar approach is the *max-RGB* method [75] that assumes the highest intensity values of the image as the white point. Thus, the estimated illuminant is given by

$$\underline{\hat{e}} = (\max(I_1), \max(I_2), \max(I_3)) \tag{2.32}$$

Another approach with a similar selection criterion is the *WhitePatch*. In this case the white point is equal to the *RGB* values of the image pixel having the maximum colour module.

A generalisation of *Grey-World* and *max-RGB* was defined in *Shades-of-Grey* [40], which generalises and outperforms the previous methods by using the Minkowski norm, Particularly, an illuminant $\underline{\hat{e}}$ is considered as the scene illuminant if it accomplishes

$$\left(\frac{\int I^p dx}{\int dx} \right)^{\frac{1}{p}} = k \cdot \underline{\hat{e}} \tag{2.33}$$

where k is a constant. In particular, $p = 1$ is the Grey-World method, and $p = \infty$ is the Max-RGB algorithm.

Grey-Edge [111] extends *Shades-of-Grey.* In this case, the image derivatives are added to the Minkowski norm. This generalisation improvies previous results. It is based on a similar formulation

$$\left(\int \left|\frac{\delta I^p dx}{\delta dx}\right|\right)^{\frac{1}{p}} | = k \cdot \underline{\hat{e}} \tag{2.34}$$

Probabilistic methods are those based on the assumption that colour constancy can be solved in the frame of a probability space, $(\Omega, \mathcal{F}, P)$, where Ω is the feasible set of illuminants, $\mathcal{F}$ is the σ-algebra composed by elements for each illuminant individually, and

$$P(\underline{e}|I) : \Omega \longrightarrow [0, 1]$$

is the probability of having illuminant $\underline{e}$ given the image I.

Then, in these methods prior information is used to select the illuminant. For example, Bayesian Colour Constancy [19] uses as prior information the fact that each surface can be modelled as a truncated multivariate normal distribution. One of the main problems of this work is its complexity, since it depends on the number of surfaces in the scene. Particularly, in this paper they suppose 8 surfaces. The increase of this value also increases the complexity of the algorithm. A generalisation for this method can be found in [51].

Voting methods, such as Illumination by Voting [100] which use the Probabilistic Hough transform, compute the conditional probability of all the illuminants given a RGB value of the image. That is,

$$\underline{\hat{e}} = \arg\max_{\underline{e}} \sum_{\underline{p} \in RGB_{im}} Pr(\underline{e}|\underline{p}) \tag{2.35}$$

Color-by-Correlation [36] uses as prior information the probability of a colour appears under one particular illuminant. In particular, they learn the probability of a surface to appear under a particular illuminant and they build a correlation matrix with this information. This method works in a chromaticity space and the illuminant is selected as the one that accomplishes

$$\underline{c}_e = thresh2(thresh(chist(C_{im}))^t M_{Bayes})C_{ill}) \tag{2.36}$$

where

C_{im} are the image chromaticities.

C_{ill} are the chromaticities for the set of selected illuminants.

M_{Bayes} is the correlation matrix.

$chist(x)$ is an histogram in 2-Dimensions.

$thres(x)$ is a binarisation to 1 for $x > 0$.

$thres2(h)$ returns a vector with 0 in all positions except the one where the global maximum is achieved, in this case it returns 1.

Notation provided in equation 2.36 is proposed in the same paper [36] as a general framework for all computational colour constancy methods.

Physics-based methods. Whilst all the previous reviewed methods are based on the Lambertian simplification, here we have grouped the methods that assume the full dichromatic model of colour, and therefore, they also consider the reflection component.

Some examples of these physical methods are found in Funt *et al.* [49] that uses mutual reflections to solve the problem, Klinker *et al.* [70] that uses both highlights and shading and the method of Lee [76] which uses specular highlights to correct the illuminant.

Criteria combination Recently, some researchers have dealt with the idea of combining different colour constancy methods by selecting the method that better perform for each image. First attempt was defined by van de Weijer *et al.* [112]. In this paper, the selection of the colour constancy solution is related to an image annotation task. The method is based on Probabilistic Latent Semantic Analysis [62].

Gijsenij *et al.* [52] have recently proposed a method based on the Weibull statistics of the natural images. That is, the statistics of the image determine which method is used. Other authors, such as Bianco *et al.* [14], [15] have used different classification techniques, such as decision forests, in order to select the appropriate method.

Finally, other methods have used a voted-for procedure in an image by dividing it in multiple patches. These patches are defined to better accomplish the physical properties of the different methods. An example of these methods is the one defined by Riess *et al* in [97].

2.4 Usual error measure

The usual error measure used to evaluate colour constancy algorithms is the angular error. The angular error calculates the angular distance between the physical (real) illuminant and the estimated illuminant. This measure does not take into account intensity information, that is, it only measures chromaticity differences independently of the brightness. Let $\underline{p}_w$ be the physical white point (illuminant), and let $\widehat{\underline{p}}_w$ be the estimated white point. Then, we define the angular error between these two illuminants as

$$e_{ang} = acos\left(\frac{\underline{p}_w \cdot \widehat{\underline{p}}_w}{\|\underline{p}_w\|\|\widehat{\underline{p}}_w\|}\right) \tag{2.37}$$

To evaluate the performance of different methods on large datasets, either the root mean square error, the median or the mean on the angular error is computed. In this work we mainly use the root mean square of the angular error since is the measure that punishes the outliers.

Part I

Part 1: Illuminant selection

Chapter 3

Colour naming for colour constancy

Finding colour representations which are stable to illuminant changes is still an open problem in computer vision. Until now most approaches have been based on physical constraints or statistical assumptions derived from the scene, while very little attention has been paid to the effects that selected illuminants have on the final colour image representation.

The novelty of this chapter is to propose perceptual constraints that are computed on the corrected images. We define the *category hypothesis*, which weights the set of feasible illuminants according to their ability to map the corrected image onto specific colours. Here we choose these colours as the universal colour categories related to basic linguistic terms which have been psychophysically measured. These colour categories encode natural colour statistics, and their relevance across different cultures is indicated by the fact that they have received a common colour name.

From this category hypothesis we propose a fast implementation that allows the sampling of a large set of illuminants. Experiments prove that our method rivals current state-of-art performance without the need for training algorithmic parameters. Additionally, the method can be used as a framework to insert top-down information from other sources, thus opening further research directions in solving for color constancy.

3.1 Introduction

Colour is derived from three components: the reflectance of the object, the sensitivity of cones, and the illuminant spectra. Of these components, the illuminant spectrum is the least stable. Illumination changes depending on different aspects: time of the day (daybreak, midday, sunset), or indoor/outdoor situations, for example. Thus the problem for computer vision is that the colour of an object depends on the light under which we are looking at it. The human visual system solves this problem thanks to the so-called color constancy property [65]. This property allows humans to identify the colour of an object independently of the colour of the light source. Colour constancy is important for human vision, since colour is a visual cue that helps in solving different visions tasks such as tracking, object recognition or categorisation. Therefore, several computational methods have tried to simulate human color constancy abilities to stabilise machine color representations.

None of the computational approaches defined in section 2.3 have introduced perceptual constraints. Consequently, very little attention has been paid to how the selected illuminant affects the perception of the content of the corrected image. Evidence derived from experimental psychology on natural images gives support to the conclusion that several different perceptual mechanisms contribute to achieve constant images [65]. Different mechanisms based on different visual cues such as the local and global contrast [74], [43], highlights [64], mutual reflections [71], categorical or naming stability [88] and memory colour of known objects [57], [58] are responsible for the almost perfect behaviour of the human constancy system. In this thesis we focus on the definition of a colour constancy method that considers the perceptual effects of categorisation on the corrected image.

In this chapter we concentrate on the naming stability cue. We propose the *naming hypothesis* as a criterion to constrain the feasible illuminants. We propose to use the capability of categorising, or assigning basic colour names, in the corrected image as the basis for weighting all feasible illuminants. In this sense, preferred illuminants will produce a colour categorised image with useful properties for further recognition tasks. Moreover, our process can be justified as it produces an image labelled with the colour categories that encode natural colour statistics which have evolved as relevant across different cultures by receiving a common colour name. The existence of the basic colour category terms was noted for the first time by Berlin and Kay [12], who recorded 11 basic terms. These basic terms were lately measured by Boynton and Olson [17] in psychophysical experiments.

Using the category hypothesis, we propose a computational approach that is a probabilistic method similar to illuminant voting [100] or colour by correlation [36], but with two essential novelties that we list below.

Firstly, the method gives a compact framework that allows prior-knowledge from learnt-colour categories to be easily introduced. Illuminant selection is done through the *Category hypothesis*, which is defined as the preference of illuminants that assign colour categories in the corrected images. In particular, we want to stress that this new algorithm can also be seen as a generalisation of simpler methods, such as, *WhitePatch* where we only consider the white category. This opens up a new way of generalising simple methods to allow greater complexity (i.e. not only by increasing their statistical complexity).

Secondly, we present a fast algorithm that builds a weighted feasible set for a fine sampling of the feasible illuminants. This fast algorithm can also be seen as a fast implementation of the Colour by Correlation approach [36] for the 3D case [8] in the particular case of a diagonal model of illuminant change. This fast algorithm requires the representation of the weighted feasible set in logarithm space. This in turn improves the illuminant selection step, since multiple solutions can be easily considered using a compact representation.

To evaluate the performance of the proposed approach, we compare our results with the existing state-of-the-art in terms of how well the illuminant is estimated. The results suggest that our approach achieves the performance of the other methods, whilst also incorporating the advantages mentioned above.

The chapter has been organised as follows. In section 3.2 we explain the basic colour term categories. Afterwards, in section 3.3 we introduce the category hypothesis, and we report the results compared to other current methods in section 3.4.

3.2 Basic term categories

Basic colour term categories were first defined by Berlin and Kay [12], and they were deduced from a large anthropological study based on speakers of 20 different languages and specific documentation from a further 78 languages. They concluded that the universal basic colour terms defined in most evolved languages are *white, black, red, green, yellow, blue, brown, purple, orange, pink* and *grey.* In subsequent works, psychophysical experiments have generated data that allow these basic categories to be specified accurately [106], [17], [10]. These datasets give 11 categories where colours have been labelled with a unique name. They are obtained from the averaged judgements given by all subjects in the experiment.

Basic colour categories are derived from anthropological and psychophysical experiments that bring us to the conclusion that relevant colours are those that receive a common colour name across different cultures. A similar conclusion about the relevance of these specific colour categories has also been derived from a biological model of the human colour sensors [92]. This work provides strong evidence that colour coding in human vision favours these colour categories. There are evidences that basic colour terms are likely to be encoding fundamental natural colour statistics [123]. That makes sense in an evolutionary theory as they would capture the most relevant information to survive.

In this work we make use of a mapping of these categories provided by Benavente-*et al.* [10]. The first row in Figure 3.1 shows the chromaticity of the convex-hull of these mapped colours at three different levels of intensity in the CIELab space. These polyhedron contain the parts of the colour space that are judged as pure colours (or focal colours); i.e. those colours named with a unique basic term. We will use these sets of colours as the anchor categories that will determine the corrected images. These sets are the focal points (F_i) of the corresponding colour. We use the CIELab space for Figure 3.1 for explanatory purposes.

In order to also encode common changes of these colours in real scenes, such as those in shadowed areas or textured surfaces, or even colours reproduced in man-made objects, we are going to experiment with some extensions of these basic categories, whilst not extending them beyond the convex-hull of the basic terms. Therefore, we define our categories depending on the distance to the focal points, whilst constraining them to remain inside the Convex Hull of the focal terms. Thus, a category C_i^β is defined as

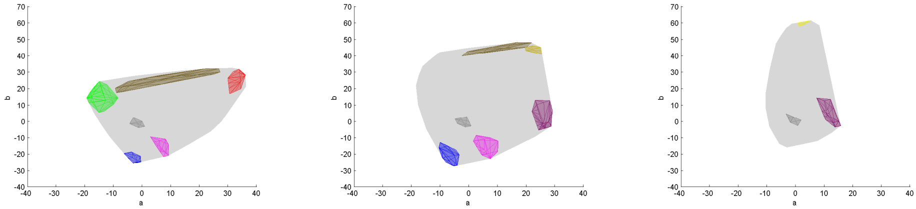
$$C_i^\beta = \{\underline{p} : \mathrm{d}(\underline{p}, F_i) < \beta, \underline{p} \in CH(F)\} \tag{3.1}$$

where $\underline{p}$ is a point in RGB space, $F = \{F_i\}_{i=1:11}$ is the set of focal colours presented in [10], CH represents the convex hull of a set of points and d refers to the Euclidean distance.

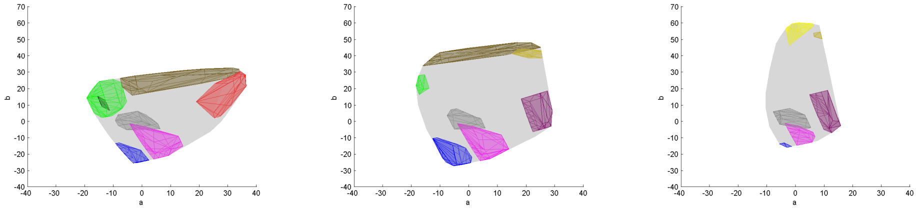
Then, from these equations, we are able to define a family of category sets by changing the β value. In Figure 3.1 we show some examples for these sets, where the first row represents the original basic categories ($\beta = 0$) as horizontal cross-sections in Lab space ($L = 25$, $L = 45$, and $L = 65$), and the second and third rows represent two different sets, $\beta = 10$ and $\beta = 20$ respectively. The grey background in all the different plots represents the global convex hull, which is the growing limit. To discretise category membership we will use a characteristic function defined as:

$$\mathcal{X}_{C_i^\beta}(\underline{p}) = \begin{cases} 1 & \text{if} \qquad \underline{p} \in C_i^\beta, \underline{p} \notin C_{j,j<i}^\beta \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

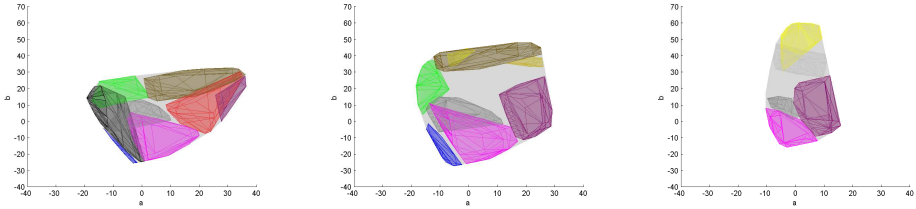
where $C^\beta = \{C_i^\beta\}_{i=1:11}$, i encodes each one of the eleven basic terms, namely {*white, black,*



(a) $L = 25$, $\beta = 0$ (b) $L = 45$, $\beta = 0$ (c) $L = 65$, $\beta = 0$



(d) $L = 25$, $\beta = 10$ (e) $L = 45$, $\beta = 10$ (f) $L = 65$, $\beta = 10$



(g) $L = 25$, $\beta = 20$ (h) $L = 45$, $\beta = 20$ (i) $L = 65$, $\beta = 20$

Figure 3.1: (a) colour name categories with luminance 25 in Lab space (b) colour name categories with luminance 45 in Lab space (c) colour name categories with luminance 65 in Lab space, (d), (e) and (f) first extension of the categories. (g),(h) and (i) second extension.

red, green, yellow, blue, brown, purple, orange, pink and *grey*} and $\underline{p}$ is a colour representation vector.

3.3 Category methods

We base our approach on the idea that colour constancy aims to produce corrected images where important contents are stable. We refer to these important contents as basic colour categories. These anchor categories constitute prior knowledge that is useful for general image understanding. Therefore we seek to correct images towards a new representation where these basic categories are anchors. This idea is formulated in the following hypothesis for colour constancy:

Category Hypothesis: *Feasible illuminants can be weighted according to their ability to anchor the colours of an image to basic colour categories.*

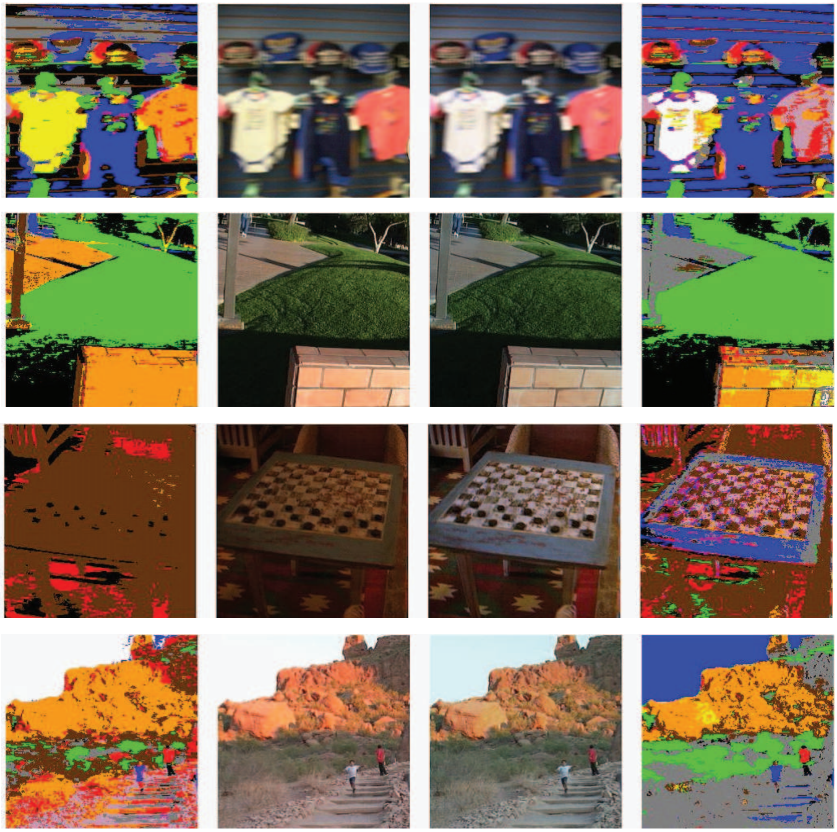


Figure 3.2: categorised original image (left), original image (center-left), corrected image (center-right), categorised corrected image(right).

Thus, we will call Category Methods those that, applying this hypothesis, compute a weighted feasible illuminant set according to the set of anchor categories being used, and select one of them that allows us to obtain a corrected image whose colours falls into these categories.

In Figure 3.2 we show some examples of the results provided by the proposed hypothesis using the basic colour terms categories. The original images are shown in the second column, while the first column presents the categorisation of these images. In the third column we give the corrected images and their corrected categorisation is given in the fourth column. Hence, from the first and the fourth column we can see how colour categorisation is changed, from the original to the corrected image, towards a more colourful image representation that in turn makes it more stable (e.g. sky is blue, the road is grey). Clearly, our proposal is simply a bottom-up approach that pursues a corrected, or more stable, image that needs further processing for full image understanding.

We will now explain our method in three parts: first, we will define the general mathematical formulation; secondly, we will explain the fast implementation of this mathematical formulation; and finally, we will explain the illuminant selection criteria.

3.3.1 Mathematical formulation

Let us define $P(\underline{e}|I)$ as the probability of having illuminant $\underline{e}$ in image I. This is approximated as

$$P(\underline{e}|I) \approx \frac{f(\underline{e})}{\sum_{\underline{\check{e}} \in FS} f(\underline{\check{e}})} = k_1 \cdot f(\underline{e}) \tag{3.3}$$

where FS is the feasible set of illuminants (in the C-Rule sense, considering as canonical gamut the whole RGB cube) and the function $f(\underline{e})$ is defined in a voting procedure in the same manner as Sapiro in [100]. This voting function is defined as

$$f(\underline{e}) = \sum_{\underline{p} \in RGB_I} P(\underline{e}|\underline{p}) \tag{3.4}$$

where RGB_I represents the different colours appearing in the image, and $P(\underline{e}|\underline{p})$ is the probability of having illuminant $\underline{e}$ given colour $\underline{p}$ in the image. This probability is defined to follow the category hypothesis introduced earlier, thus

$$P(\underline{e}|\underline{p}) = P(\underline{e}|\underline{p}, C^\beta) = \frac{\sum_{C_i^\beta \in C^\beta} \mathcal{X}_{C_i^\beta}(\underline{p} \cdot \mathrm{diag}(\underline{e})^{-1})}{\sum_{C_i^\beta \in C^\beta} \sum_{\underline{q} \in RGB} (\mathcal{X}_{C_i^\beta}(\underline{q}))} \tag{3.5}$$

quantifies the ability of illuminant $\underline{e}$ to categorise colour $\underline{p}$ in the set of anchor categories denoted as C^β, and is normalised by the total amount of nameable colours. $\mathcal{X}_{C_i^\beta}(\underline{x})$, defined in equation 3.2, is responsible for counting the number of colours falling in each one of the categories for the specific illuminant.

To simplify the previous formulation, the denominator in equation 3.5 is substituted by a constant

$$k_2 = 1/ \sum_{C_i^\beta \in C^\beta} \sum_{\underline{q} \in RGB} (\mathcal{X}_{C_i^\beta}(\underline{q})) \tag{3.6}$$

and we therefore rewrite $P(\underline{e}|I)$ as

$$P(\underline{e}|I) \approx k_1 \cdot k_2 \sum_{\underline{p} \in RGB_I} \sum_{C_i^\beta \in C^\beta} \mathcal{X}_{C_i^\beta}(\underline{p} \cdot \mathrm{diag}(\underline{e})^{-1}). \tag{3.7}$$

We want to highlight here, that this compact formulation could be used for a different set of categories than those used in this paper. Indeed, existing colour constancy methods can be incorporated within this framework. For instance, using white as a unique category means that the method acts as a White-Patch algorithm, while taking all possible colour values for a certain device as different categories behaves like the Color-by-Correlation [36] solution in the diagonal case for a 3D colour space.

3.3.2 Fast implementation

The main problem of this formulation is its cpu time, which is large due to the double summation term. Therefore, in order to reach a fast implementation of the proposed voting approach, we reformulate equation (3.7) by reordering sums and obtaining

$$P(\underline{e}|I) \approx k_1 \cdot k_2 \cdot \sum_{C_i^\beta \in C^\beta} \sum_{\underline{p} \in RGB_I} \mathcal{X}_{C_i^\beta}(\underline{p} \cdot \mathrm{diag}(\underline{e})^{-1}) \tag{3.8}$$

in this way, the inner summation is equivalent to a product of two functions histn and $\mathcal{X}_{C_i}^\beta$, where histn is the normalised histogram of the image I and $\mathcal{X}_{C_i^\beta}$ is the characteristic function of a category C_i^β. Both functions are defined over the complete RGB domain which allows the reformulation of the previous equation as

$$P(\underline{e}|I) \approx k_1 \cdot k_2 \cdot \sum_{C_i^\beta \in C^\beta} \sum_{\underline{r} \in RGB} \mathrm{histn}(\underline{r} \cdot \mathrm{diag}(\underline{e})^{-1}) \cdot \mathcal{X}_{C_i^\beta}(\underline{r}). \tag{3.9}$$

Note that from now on, the inner summation is over the set of possible RGB values instead of over the values appearing in the image.

At this point we propose to estimate this probability by removing constants k_1 and k_2 and introducing a log monotonic function in the image domain. This implies that

$$\begin{aligned} P(\underline{e}|I) &\approx k_1 \cdot k_2 \cdot \hat{P}(\underline{e}|I) \\ &\propto \hat{P}(\underline{e}|I) \\ = \sum_{C_i^\beta \in C^\beta} \sum_{\underline{r} \in RGB} &\widehat{\mathrm{histn}}(\log(\underline{r} \cdot \mathrm{diag}(\underline{e})^{-1})) \cdot \hat{\mathcal{X}}_{C_i^\beta}(\log(\underline{r})) \end{aligned} \tag{3.10}$$

where the membership function and the histogram function have been redefined in log space as $\hat{\mathcal{X}}_{C_i^\beta}(\underline{r}) = \mathcal{X}_{C_i^\beta}(\exp(\underline{r}))$ and $\widehat{\mathrm{histn}}(\underline{x}) = \mathrm{histn}(\exp(\underline{x}))$. Furthermore, considering that taking logarithms transforms products into additions, we can write

$$\begin{aligned} &\hat{P}(\underline{e}|I) = \\ \sum_{C_i^\beta \in C^\beta} \sum_{\underline{r} \in RGB} &\widehat{\mathrm{histn}}(\log(\underline{r}) + \mathrm{diag}(\log(\underline{e}))^{-1}) \cdot \hat{\mathcal{X}}_{C_i^\beta}(\log(\underline{r})) \end{aligned} \tag{3.11}$$

which brings us to compute a linear correlation of two functions

$$\hat{P}(\underline{e}|I) = \sum_{C_i^\beta \in C^\beta} (\widehat{\mathrm{histn}} * \hat{\mathcal{X}}_{C_i^\beta})(\underline{e}) \tag{3.12}$$

that can be computed in the Fourier space as a simple product of functions. Using the *Fast Fourier Transform (FFT)* this can be done with a computational cost $O(n^3 \log(n))$.

3.3.3 Illuminant selection

In the foregoing sections we defined a computational framework that provides a weighted set of feasible solutions. The proposed algorithm assigns different probabilities to all plausible illuminants accordingly with the category hypothesis. The next step is to select the most relevant illuminant by using some specific criterion. To evaluate the performance of the hypothesis we set up experiments with two different criteria: i) selecting the illuminant with the maximum probability, which is the most common approach in probabilistic methods; and ii) selecting the illuminant by combining our feasible solutions with solutions provided by other methods which are based on a complementary hypothesis. In this way we can evaluate whether the category hypothesis can be improved by combining it with, for example, an edge-based hypothesis. This combination criterion can be seamlessly integrated within the proposed algorithm, which is another advantage of this framework. The use of a global convolution in the log-RGB space is the basis that allows the probabilities for a large sample of illuminants within the feasible set to be calculated, and allows us to work directly with these probabilities.

Using a maximum criterion we can formulate Category Correlation methods (heretofore CaC) to deliver a unique solution, which is given by

$$\underline{\hat{e}} = \arg \max_{\underline{e} \in FS} P(\underline{e}|I) \tag{3.13}$$

where $\underline{\hat{e}}$ is the estimated illuminant for the scene based on equation (3.3).

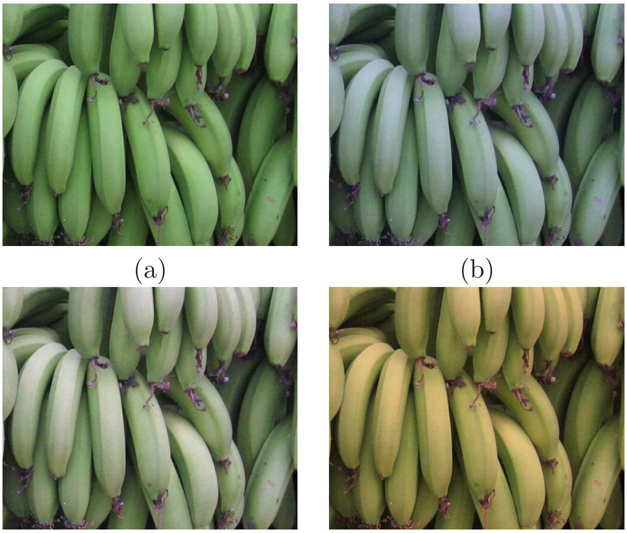
Using a combination criterion we are assuming that our weighted feasible set is providing different plausible explanations of the corrected image. For instance, in some particular images such as the bananas shown in Figure 3.3, we can see that disambiguating the scene illuminant from the object reflectances is an unsolvable problem. In this case most of the solutions in the feasible set could be equally plausible since they could correspond to different ripeness of the fruit or different illuminants. The four images in Figure 3.3 have been obtained from a clustering with standard k-means with four classes onto the feasible set and extracting the illuminant with maximum probability as the representative of each cluster. In this case, the original image was close to the green bananas given in solution (a).

Accordingly with the previous observation we can state that working with multiple solutions can be an improvement to classical constancy approaches. One of the strengths of our method relies on the fact that a large sample of likely illuminants has already been computed. In this way we can extract multiple solutions by directly thresholding onto the weighted feasible set. Then, a multiple solution set for a given image I is given by

$$S^{\alpha} = \{\underline{e} \in FS : P(\underline{e}|I) > \alpha\}, \tag{3.14}$$

which denotes the set of illuminants having a probability higher than α. Providing multiple solutions allows us to delegate the final selection either to other visual processes with contextual information or to other top-down selective tasks. This approach has been used in [112] where an illuminant is selected to improve a scene recognition task from a variety of solutions from different constancy methods (and after a learning step). There are also other methods selecting a unique solution from a set of precomputed ones [14], [15]. These last methods use classification techniques such as decision forest to this end.

Here in this work, we propose a criterion that estimates the best illuminant by selecting the solution from $S^{\alpha} = \{S_i\}_{i=1,\cdots,n}$ that is the most voted-for by solutions derived from other



(a) (b) (c) (d)

Figure 3.3: Different feasible solutions for the same scene providing different explanations of that scene.

methods based on different hypotheses and which are denoted as $\{T_j\}_{j=1,\cdots,m}$. Formally, we select the most voted-for illuminant by computing

$$\underline{\hat{e}} = S_{\arg\max_i \#\{v_j \in \underline{v} : v_j = i\}} \tag{3.15}$$

where $\underline{v} = \{v_j\}_{j=1,\cdots,m}$ encodes the solution of S^α that is closest to a solution in $\{T_j\}_{j=1,\cdots,m}$, and

$$v_j = \arg\min_i \operatorname{ang}(S_i, T_j) \tag{3.16}$$

where ang is the angular error distance between two given illuminants.

With this criterion we select an illuminant which has a high probability based on our own hypothesis and is reinforced by being close to the solutions provided by other hypotheses.

3.4 Experiments

To evaluate our hypothesis we have run our method under different parameters varying the category sets and the selection criteria. We have used two different datasets and we have compared our results with current state-of-the-art.

We denote our method as CaC_{sc}^{β} where sc is the selection criterion used and β refers to category threshold. The selection criterion will be m for selection based on maximum probability and c for a combined selection. For both selection criteria the value of β takes one out of four possible values: 0 (in order to use the basic categories), 10, 20 and 400.

This last value has been defined in order to select the complete convex hull (grey polygon in Figure 3.1). In all the experiments our methods have worked with a $\log RGB$ cube of 50 bins, that implies a sampling of 50^3 different illuminants.

Specifically for the combined criterion, we have selected our solutions by setting $\alpha = 0.95 \cdot \max(P(e/I))$. We have combined these solutions with 24 solutions coming from different applications of the grey-edge hypothesis. We have used a wide range of statistical combinations of this hypothesis by fixing the following parameters $p = 1, 6, 11, 16$, $\sigma = 1, 3$ and $n = 0, 1, 2$ where p is the Minkowski Norm, σ the smoothness parameter and n the differentiation order

We compare our method with previous approaches. These methods are divided in two groups, calibrated and no-calibrated methods. The first group includes C-Rule (maximum volume (GM-MV) and average (GM-AVE)) [42] and Gamut Constrained illuminant estimation (GCIE) [37]. In the second group we compare with Grey-Edge [111], Shades of grey [40], Max-RGB [75], Grey-World [21], Color-by-Correlation [36] and Neural Networks [22].

We have run the Grey-Edge algorithm provided by the author [111], and have considered the following set of parameters: $0 \leq n \leq 2$, $0 \leq \sigma \leq 5$, $0 \leq p \leq 15$. For Shades-of-Gray the values are $0 \leq \sigma \leq 5$, $0 \leq p \leq 15$. For the training of these two methods, we used 33% of the images to set the parameters, and we applied these parameters to the rest of the images. In this way, independence between training and testing sets is preserved.

Same experiments have been applied to two different images datasets that we list below:

Dataset 1. *Real-World Images* This dataset created by Ciurea and Funt [26] is composed by images captured with a grey sphere in the image field of view. This sphere allows to estimate the scene illuminant. In our experiments the ball has been excluded in order to avoid any influence in the results. This image dataset is gamma corrected, therefore we have removed this correction considering $\gamma = 2.2$ that is the typical value used in sRGB devices. Furthermore, since this dataset was recorded by a video-camera, all the image scenes within each of the 15 scenarios present a high correlation of contents. To avoid the effects derived from this fact we have followed a similar procedure from previous reported experiments. In particular, we have used the frames extracted in [13], that are the biggest independent image dataset that can be extracted from the Ciurea-Funt dataset. The total amount of images is 1135, but with a different number of images for each scenario. Both, for Grey-Edge and Shades-of-Grey we have used 5 scenarios for training and 10 scenarios for testing.

Dataset 2. *Controlled Indoor scenes:* This dataset created at Simon Fraser University [9] is composed by 321 indoor images. It consists of 31 scenes captured under 11 different conditions, totalling 321 images. This dataset is formed by raw images, therefore no gamma correction is needed. In this experiment we trained both Grey-Edge and Shades-of-Grey by using 10 scenes for training and 21 for test.

In order to analyse if category hypothesis brings us to meaningful solutions we have used the root mean square (RMS) of the angular error between the solution and the known scene illuminant. Low RMS error rates implies that images are generally corrected in a good direction. In this work we present a first step in the introduction of colour categories to constraint colour constancy, therefore we have based our analysis on how corrected solutions fall onto the basic term categories.

3.5 Results and discussion

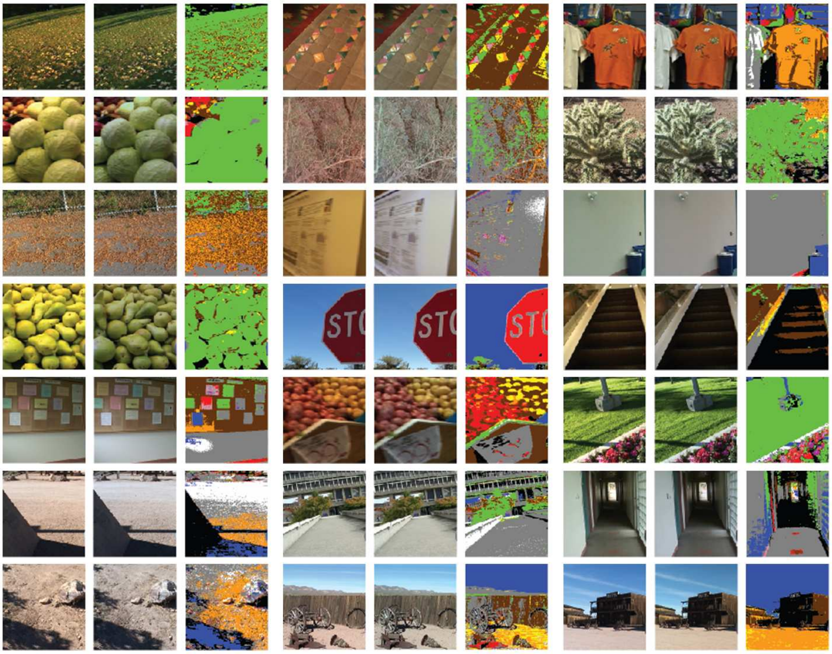


Figure 3.4: Examples from a Real-World dataset. Original image (left), corrected image (center), categorised corrected image (right).

The results obtained from the aforementioned experiments are summarised in Table 3.1. The results are divided into three parts: our results, uncalibrated methods and calibrated methods in this order. First rows of the table are related to our method. In particular, from the first two rows we can observe that our method achieves state-of-the art results by using a completely new hypothesis and with no training step that could set parameters to the dataset content. In these first two rows we applied the basic method CaC^0 that simply uses the focal colours of the 11 basic colour categories. The combination criteria does not introduce critical changes on the performance. In subsequent rows we study the effect on the performance of our method when changing the basic categories.

In the second part of the table we report the performance of different uncalibrated methods. From those methods, we have reported the results on the two datasets for the ones where we could run the code, and for the rest of methods (Neural Networks [22] and Color-by-Correlation [36]) we report the results provided in the literature that were just on dataset 3. For the case of calibrated methods we report the results for GM-MV and GM-AVE [42] computed by us, and we have transcribed from previous works the results for GCIE-11 and GCIE-87 [37]. A clear advantage is shown by calibrated methods which use the information derived from knowing camera sensitivities.

Before to analyse the results obtained when changing size of the basic colour categories, it is worth to note an important observation provided by experiments not reported here.

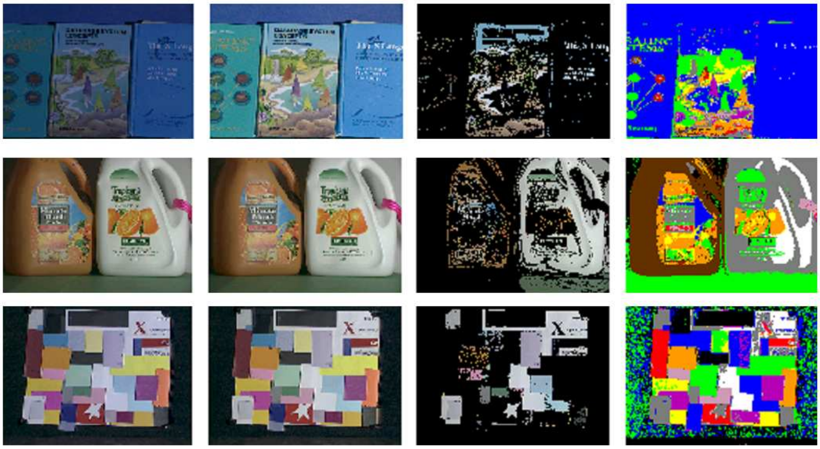


Figure 3.5: Controlled indoor dataset. Original image (left), corrected image by proposed method (center-left), points weighting the selected illuminant (center-right), categorisation of corrected images (right)

Method	Dataset 1	Dataset 2
$\mathbf{CaC^0_m}$	**14.57°**	**8.82°**
$\mathbf{CaC^0_c}$	**14.63°**	**8.19°**
CaC^{10}_m	**14.43°**	8.29°
CaC^{10}_c	14.55°	7.66°
CaC^{20}_m	14.72°	7.34°
CaC^{20}_c	14.74°	7.23°
CaC^{400}_m	14.76°	7.23°
CaC^{400}_c	14.79°	**7.05°**
Grey-Edge	14.62°	8.56°
Shades-of-Grey	14.77°	8.73°
Max-RGB	15.89°	11.76°
Grey-World	15.97°	13.56°
no-correction	20.32°	19.64°
Colour by Correlation	-	10.09°
Neural Networks	-	11.04°
GCIE 87 lights	-	7.11°
GCIE 11 lights	-	6.88°
GM-MV	-	6.89°
GM-AVE	-	6.86°

Table 3.1: Angular error on the different datasets.

Method	Dataset 1	Dataset 2
$CaC^{0}_{p_b}$	11.91°	7.12°
$CaC^{10}_{p_b}$	11.53°	6.27°
$CaC^{20}_{p_b}$	11.81°	5.70°
$CaC^{400}_{p_b}$	11.99°	5.54°

Table 3.2: Angular error limit by selecting the best solution during the combination on the different datasets.

Increasing the size of categories beyond the convex-hull of the basic colour categories conveys in a high decrease of performance. This fact supports the idea of the adequacy of the basic colour terms as centred anchors for a good adaptation to the most common image content.

As we can see from the results, for the case of a big real-world dataset (dataset 1) best results are obtained with the smallest categories CaC^0 and CaC^{10}. This result agrees with the general hypothesis of the method that defends that basic colour categories encode natural colour statistics, that agrees with the content of dataset 1, which is large dataset formed by a majority of natural images.

For the indoor dataset (dataset 2), the best results are achieved when we use the biggest sizes of categories, that is, the full convex hull of the colour categories. This fact can be explained by the high amount of non-natural and non-basic colours such as turquoise or other intermediate colours appearing in big areas of the images. In general, man-made objects present the property of having random colours (maybe non basic) in big homogeneous surfaces (non-textured). The size of the basic colour categories usually agrees with their texture appearance, a big green category correlates with highly textured green areas in natural vegetation, while yellow and red correspond with small category volumes correlating with a less frequent apparition in natural environments. Big homogeneous areas in man-made objects imply histograms with sharp peaks which lead to an increase in the number of solutions that can achieve a high weight. This obviously implies an increase in the error measure. It is for this last reason that the combination criteria works very well in this dataset. Many different interpretations are plausible, therefore the use of different cues becomes more important. We can see how CaC^{400}_c reaches almost the level of calibrated methods when the categories are adapted to the dataset content.

Apart from the shown results, we want to outline a further advantage derived from the method. The estimated illuminant provides us with an annotated image that informs about which parts of the images have been selected as anchors and with which colour. In Figure 3.5 we show some results of CaC^0_m, e.g, using basic categories and maximum selection, for images in dataset 1. From left to right, first column shows the original image, second column corresponds to the corrected image and third column display the categorised image. In Figure 3.2 we show a similar example for dataset 2 with the same basic method. In this case, the first and second columns show the original and corrected images respectively. Third column shows the points that have been annotated with basic names in the selected solution. Finally, the fourth column presents the categorisation of the corrected images with basic terms.

Forthwith, we also computed the performance bound we can obtain by improving the illuminant selection step in S^α. We want to emphasise again that all the images selected in this set were highly categorised with basic colours due to our selection of the value α.

The results for these performance bounds are shown in Table 3.2. These results reinforce our hypothesis since they prove that a proper solution is included into the set of higher categorised images.

The proposed method paves the way for further research related to the introduction of top-down knowledge from the image content that can additionally constrain the number of solutions and consequently allow a higher performance. By top-down knowledge we refer to further processes on the image content that can provide clues to select which are the best colour categories and even where they should be located in the image. Additional visual cues informing about the existence of a tree in the image applies for specific locations to find green colour in the image. To evaluate the effects of this kind of top-down knowledge onto the performance of our method, we have done one further experiment, that is reported in Table 3.3.

In this experiment we have applied a pre-computation step that has provided the basic colour categories appearing in the image under the canonical illuminant. In this way, this specific set of categories has been used to apply the basic algorithms to each image. In Table 3.3 we show the results of estimating the illuminant by selecting the maximum probability from the feasible set built using specific categories for each image. We can see that by introducing information from other top-down visual processes the increase on the performance is substantial.

Method	Dataset 1	Dataset 2
CaC^0_{TD}	11.51°	7.70°

Table 3.3: Top-down approach.

3.6 Conclusions and further work

The main novelty of this work is the definition of a new hypothesis for colour constancy that relies on a set of reflectances, or colour categories, that encode relevant colour information in natural scenes. These categories are those that receive a name across different languages and cultures. These colours are distributed around the achromatic reflectances and we hypothesise that they can act as anchors for image correction.

We propose a colour constancy method that estimates the best illuminant according to its ability to label image points with these basic colour categories. We use representatives for these categories obtained from psychophysical experiments. The method we propose builds a set of feasible illuminants that are weighted according to the hypothesis.

A fast implementation is easily defined by working in log-space. The proposed algorithm allows one to obtain a large sampling of the feasible solutions which is the basis for a useful framework. Having a set of multiple solutions provides different selection criteria and an open framework to introduce new cues from complementary visual processes.

We show that our methods achieve current state of art with some advantages. Our method is a purely bottom-up method providing a framework for further combination with complementary visual information. The method is based on general psychophysical data that can be modified depending on the application. Lastly, and most importantly, our results are achieved without the need for a training step, as is required in many other approaches.

The proposed method can be framed within the family of probabilistic methods that estimates the illuminant by voting. The method can be seen as a generalisation of previous approaches such as *WhitePatch*, which results from using a single achromatic category in our method, or *Color-by-Correlation* (for the $3D$ case) where categories are represented by the full set of reflectances used.

Further research is now possible to exploit the advantages of using the weighted feasible set. Complementary visual cues, or constraints derived from specific visual tasks, can provide further information for deciding on the final illuminant.

Chapter 4

Psychophysical evaluation of colour constancy algorithms

Performance evaluation for computational colour constancy algorithms has been usually done by comparing the angular error between the estimated chromaticity and the chromaticity of a canonical illuminant, which is highly dependent on the image dataset. However, recently, some works have used perceptual constraints to estimate illuminants; in this case selection is based on increasing the performance on the subsequent steps of the systems. In this chapter we propose a new performance measure, the perceptual angular error. It evaluates the performance of a colour constancy algorithm according to the perceptual preferences of humans, or naturalness (instead of the actual optimal solution) and is independent of the visual task. We show the results of a new psychophysical experiment comparing solutions from three different colour constancy algorithms. Our results show that in more than half of the judgements the preferred solution is not the one closest to the optimal solution. Our experiments were performed on a new dataset of images acquired with a calibrated camera with an attached neutral grey sphere, which better copes with the illuminant variations of the scene.

4.1 Introduction

Colour constancy is the ability of the human visual system to perceive a stable representation of colour despite illumination changes. Like other perceptual constancy capabilities of the visual system, colour constancy is crucial for succeeding in many ecologically relevant visual tasks such as food collection, detection of predators, etc. The importance of colour constancy in biological vision is mirrored in computer vision applications, where success in a wide range of visual tasks relies on achieving a high degree of illuminant invariance. In the last twenty years, research in computational colour constancy has tried to recover the illuminant of a scene from an acquired image.

This has been shown to be a mathematically ill-posed problem and therefore does not have a unique solution. A common computational approach to illuminant recovery (and colour constancy in general) is to produce a list of possible illuminants (feasible solutions) and then use some assumptions, based on the interactions of scene surfaces and illuminants to select the most appropriate solution among all possible illuminants. Comparison studies [5], [7] have ranked the performance of different algorithms, which usually depend on the properties of the image dataset and the statistical measures used for the evaluation. It is generally agreed that, although some algorithms may perform well in average, they may also perform poorly for specific images. This is the reason why some authors [63] have proposed a one-to-one evaluation of the algorithms on individual images. In this way, comparisons become more independent of the chosen image dataset. However, the general conclusion is that more research should be directed towards a combination of different methods, since the performance of a method usually depends on the type of scene it deals with [23]. In all these approaches, the evaluation of the performance of the algorithms has been based on computing the angular error between the selected solution and the actual solution that is provided by the computational colour constancy method.

Some recent proposals, like our section 3 or [109], [112], turn away from the usual approach and deal instead with multiple solutions delegating the selection of a unique solution to a subsequent step that depends on high-level, task-related interpretations, such as the ability to annotate the image content. In this example, the best solution would be the one giving the best semantic annotation of the image content. It is in this kind of approach where the need for a different evaluation emerges, since the performance depends on the visual task and this can lead to an inability to compare different methods. Hence, to be able to evaluate this performance and to compare it with other high-level methods, we propose to explore a new evaluation procedure.

In summary, the goal of this chapter is to show the results of a new psychophysical experiment following the lines of that presented in [115]. The previous results were confirmed, that is, humans do not chose the minimum angular error solution as the more natural. Furthermore, in this chapter we propose a new measure to reduce the gap between the error measure and the humans' preference. Our new experiment represents an improvement over the old one in that it considers the uncertainty level of the observer responses and it uses a new, improved image dataset. This new dataset has been built by using a neutral grey sphere attached to a calibrated camera to better estimate the illuminant of the scene. We have worked with the shades-of-grey [40] algorithm instead of CRule [42]. This decision has been made on the basis that CRule is calibrated, whereas the other algorithms are not. This chapter is divided as follows. In section 4.2 we present how the experiment has been driven. Afterwards, in section 4.3 we show the results. Later on, in section 4.4 a new perceptual measure to deal with the evaluation of colour constancy algorithms is presented. Finally, in section 4.5, we sum up the conclusions.

4.2 Experimental setup

Subjects were presented with a pair of images (each one a different colour constancy solution) on a CRT monitor and asked to select the image that seems ”most natural”. The term ”natural” was chosen not because it refers to natural objects but because it refers to natural viewing conditions, implying the least amount of digital manipulation or global perception of an illuminant. Figure 4.1 shows some exemplary pictures from the database. The pictures

Figure 4.1: Images regularly selected in the experiment as natural (left) versus images hardly ever selected (right).

on the left are examples of images selected as natural most of the time, while those on the right are examples of images hardly ever selected as natural.

The global schematics of the experiment are shown in Figure 4.2. We used a set of 83 images from a new image dataset that was built for this experiment (the image collection details are explained in section 4.2.1). The camera calibration allows us to obtain the CIE1931 XYZ values for each pixel and consequently, we converted 83 images from CIE XYZ space to CIE sRGB. Following this, we replaced the original illuminant by D65 using the chromaticity values of the grey sphere that was present in all image scenes.

From the original images, 5 new pictures were created by re-illuminating the scene with 5 different illuminants. To this end we have used the chromatic values of each illuminant (3 Planckians: 4000K, 7000K, 10000K, and two arbitrary illuminants: Greenish ($x = 0.3026$, $y = 0.3547$) and Purplish ($x = 0.2724$, $y = 0.2458$), totalling 415 images. Afterwards, the three colour constancy algorithms (Grey-World [21], Shades-of-Grey [40] and MaxName [115]) explained in section 4.2.2 were applied to the newly created images. Consequently, we obtain one solution per test image per algorithm, totalling 1245 different solutions. These solutions were converted back to CIE XYZ to be displayed on a calibrated CRT monitor (Viewsonic P227f, which was tested to confirm its uniformity across the screen surface) using a visual stimulus generator (Cambridge Research Systems ViSaGe). The monitor's white point chromaticity was ($x = 0.315$, $y = 0.341$) and its maximum luminance was 123.78 $Cd/m2$. The experiment was conducted in a dark room (i.e. the only light present in the room came from the monitor itself).

The experiment was conducted on 10 naïve observers recruited among university students and staff (none of the observers had previously seen the picture database). All observers were tested for normal colour vision using the Ishihara and the Farnsworth Dichotomous Test (D-15). Pairs of pictures (each obtained using one of two different colour constancy algorithms) were presented one on top of the other on a grey background ($31Cd/m2$). The order and position of the picture pairs was random. Each picture subtended 10.5 x 5.5 degrees to the

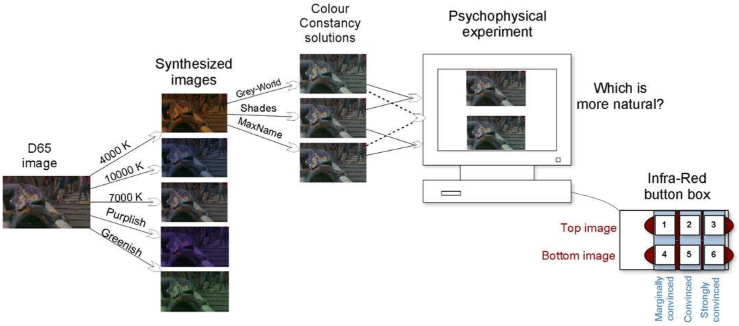


Figure 4.2: Experiment schedule

observer and was viewed from 146 cm. This brings us to 1245 pairs of observations per observer. No influence on picture (top or bottom) position in the observers' decision was found.

For each presentation, observers were asked to select the picture that seemed most natural, and to rate their selection by pressing a button on an IR button box. The set up (six buttons) allowed observers to register how convinced they were of their choice (e.g. strongly convinced, convinced, and marginally convinced). For example if an observer was strongly convinced that the top image was more natural that the bottom one, (s)he would press button 3 (see Figure 4.2), if it was marginally convinced that the bottom picture was the most natural it would press button 4 and so on. There was no time limit but observers took an average of 2.5 seconds to respond to each choice. The total experiment lasted 90 minutes approximately (divided in three sessions of 30 minutes each).

4.2.1 A new image dataset

To test the models we need a large image dataset of good quality natural scenes. From a colorimetric point of view, the obvious choice is to produce hyperspectral imagery, to reduce metameric effects. However, hyperspectral outdoor natural scenes are difficult to acquire since the exposure times needed are long and its capture implies control over small movements or changes in the scene, (not to talk of the financial cost of the equipment). There are currently good quality image databases available (such as the hyperspectral dataset built by Foster *et al.* [47] and Brelstaff *et al.* [20]), but they either contain specialised (i.e. non-general) imagery or the number of scenes is not large enough for our purposes. For this reason, and because metamerism is relatively rare in natural scenes [44], [107] we decided to acquire our own dataset of 83 images (see Figure 4.3) using a trichromatic digital colour camera (Sigma Foveon D10) calibrated to produce CIEXYZ pixel representations.

The camera was calibrated at Bristol University (UK) Experimental Psychology lab by measuring its colour sensors' spectral sensitivities using a set of 31 spectrally narrowband interference filters, a constant-current incandescent light source and a TopCon SR1 telespectroradiometer (a process similar to that by others [89], [91]). This calibration procedure was carried out by Dr. C. A. Parraga. The calibrated camera allows us to obtain a measure of the CIE XYZ values for every pixel in the image. Images were acquired around Barcelona city at

Figure 4.3: Image dataset under D65 illuminant.

Figure 4.4: Camera and grey sphere setup.

different times of the day and in three different days in July 2008. The weather was mostly sunny with a few clouds. We mounted a grey ball in front of the camera (see Figure 4.4), following the ideas of Ciurea *et al.* [26]. The ball was uniformly painted using several thin layers of spray paint (Revell RAL7012-Matt), whose reflectance was approximately constant across the camera's response spectrum and its reflective properties were nearly Lambertian (see Figure 4.5). The presence of the grey ball (originally located at the bottom-left corner of every picture and subsequently cropped out) allows us to measure and manipulate the colour of the illuminant. Images whose chromaticity distribution was not spatially uniform (as measured on the grey ball) were discarded.

4.2.2 Selected colour constancy algorithms

In this section we briefly summarise the three methods we have selected for our analysis. This explanation is a summary of that in section 2.3 for two of the methods. The third one, is just an early stage of the method explained in the previous chapter. We have chosen two well-known methods, Grey-World [21] and Shades-of-Grey [40], and a more recent method, the MaxName algorithm [115]. The Grey-World algorithm (an uncalibrated method based on a strong assumption about the scene) was selected because of its popularity in the literature. The Shades-of-Grey algorithm (another uncalibrated algorithm) was selected because it con-

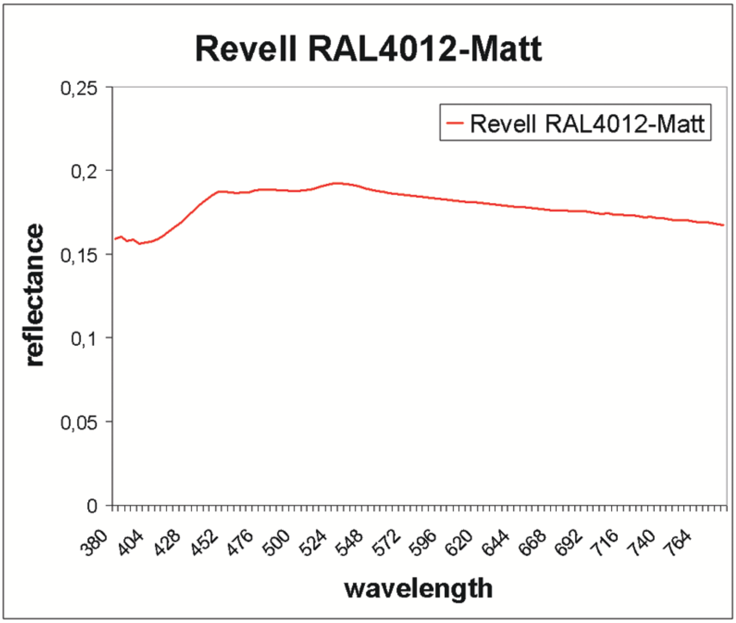


Figure 4.5: Reflectance of the paint used on the ball.

siderably improves performance with respect to Grey-World (another uncalibrated algorithm such as Grey-edge [111] could also have been used). Finally, MaxName is a previous version of the CaC methods presented in chapter 3 and it was selected because it uses high-level knowledge to correct the illuminant. We give a brief outline of these methods below.

Grey-World

It was proposed by Buchsbaum [21] and it is based on the hypothesis that mean chromaticity of the scene corresponds to grey. Given an image $I = (R, G, B)^T$ as a function of RGB values, and adopting the diagonal model of illuminant change [34], then an illuminant (α, β, γ) accomplishes the Grey-World hypothesis if

$$\frac{\int I dx}{\int dx} = k \cdot (\alpha, \beta, \gamma) \tag{4.1}$$

where k is a constant.

Shades-of-grey

It was proposed by Finlayson *et al.* [40]. This algorithm is a statistical extension of Grey-World and MaxRGB [75] algorithms. It is based on Minkowski norm of images. An illuminant (α, β, γ) is considered as the scene illuminant if it accomplishes

$$\left(\frac{\int I^p dx}{\int dx}\right)^{\frac{1}{p}} = k \cdot (\alpha, \beta, \gamma) \tag{4.2}$$

where k is a constant. Actually, this is a family of methods where $p = 1$ is Grey-World method, and $p = \infty$ is Max-RGB algorithm. In this case we have used $p = 12$, since it is the best solution for our dataset.

MaxName

This algorithm is a previous stage of the one presented in chapter 3. It is based on giving more weight to those illuminants that maximise the number of colour names in the scene. That is, MaxName builds a weighted feasible set by considering nameable colours, this is prior knowledge given by

$$\underline{\mu_k} = \int_w S(\lambda)E(\lambda)R_k(\lambda), k = R, G, B \tag{4.3}$$

where, $S(\lambda)$ are the surface reflectances having maximum probability of being labelled with a basic colour term, also called focal reflectances (from the work of Benavente [11]). In addition to the basic colour terms, we added a set of skin coloured reflectances. In Equation 4.3, $E(\lambda)$ is the power distribution of a D65 illuminant and $R_k(\lambda)$ are the CIE RGB 1955 Colour Matching Functions. We define μ as the set of all k-dimensional nameable colours obtained from Equation 4.3. The number of elements of μ depends on the number of reflectances used. Following this, we compute the Semantic Matrix, denoted as SM, which is a binary representation of the colour space as a matrix, where a point is set to 1 if it represents a nameable colour, that is, it belongs to μ, and 0 otherwise. Then, for a given input image, I, we compute all possible illuminant changes $I_{\alpha,\beta,\gamma}$. For each one, we calculate its nameability value. This is done by counting how many points of the mapped image are nameable colours in SM and can be computed by a correlation in log space:

$$Nval_{\alpha,\beta,\gamma} = log(H_{bin}(I) * \log(SM) \tag{4.4}$$

In the previous equation, $H_{bin}(I)$ is the binarised histogram of the image, $Nval$ at the position (α, β, γ) is the number of coincidences between the SM and $I_{\alpha,\beta,\gamma}$. $Nval$ is a 3-dimensional matrix, depending on all the feasible maps, (α, β, γ). From this matrix, we select the most feasible illuminant as the one that accomplishes:

$$(\alpha, \beta, \gamma) = \arg\max Nval_{\alpha,\beta,\gamma} \tag{4.5}$$

that is, the one giving the maximum number of nameable colours.

4.3 Results

The results of the experiment validate those presented by Vazquez *et al.* [115], with a different image dataset and a different set of algorithms. The main finding is that preferred solutions, namely the more natural in the psychophysical experiment, do not always coincide with solutions of minimum angular error. In fact, this agreement only happened in 43% of the

Bottom more natural than top			Top more natural than bottom		
Button 6	Button 5	Button 4	Button 3	Button 2	Button 1
Definitely	Sufficiently	Marginally	Marginally	Sufficiently	Definitely
0	0.2	0.4	0.6	0.8	1

Table 4.1: Buttons codification.

observations, independently of the degree of certainty of the observers when making the decision. Since the experimental procedure allows us to define a partition in the interval $[0, 1]$ to encode the subject selection and each observation represents a decision between two images, then for each observation we label one image as the result from Method A, and the other as the result from Method B (Method A and B are labelled as 1 and 0, respectively). The confidence of the decision is considered at three different levels (the three buttons that the subject was allowed to press (ordinal paired comparison [3])). For example, suppose that a scene processed by Method A is presented on top of the screen and a second scene processed by Method B is presented at the bottom (the physical position of the scenes was randomised in each trial, but let's consider an exemplary layout). If the subject thinks that the top picture is more natural it will press one of the top buttons in Figure 4.2 according to how much he/she is convinced. Suppose the subject presses button 3 (top-right: definitely more natural), then the response is coded as 1. If the choice is button 2 (top-center: sufficiently more natural) the response is coded as 0.8, etc. (see Table 4.1). If, on the contrary the subject thinks the bottom picture (Method B) is more natural, then he/she will press a button from the lower row. If he/she is marginally convinced, will pick button 4 (bottom-left) and the response will be coded as 0.4 according to Table 4.1. Similarly if he/she is strongly convinced, will press button 6 (bottom-right) and the response will be coded as 0. In this way we collect not only the direction of the response but its certainty. Observer's certainty was found to be correlated (corr. coef. 0.726) to a simple measure of image difference (the angular error between each image pair). This technique is similar to that used by other researchers [2], [27], [41], [50].

We have computed two different measures of observer variability. The first measure is the correlation coefficient between each individual subject and the average over all the observers (in black in Figure 4.6). Table 4.2 shows this measure. The idea behind this analysis is to detect outliers (subjects with a distribution of results significantly different to the rest of the observers, i.e. low correlation). Our second measure is the coefficient of variation (CV) [78], [79], which computes the difference between two statistical samples (see Table 4.2). Both measures were calculated for the whole 1245 observations (3 combinations of colour constancy solutions x 415 observations per combination).

From this table, and from the distribution of the plots in Figure 4.6, we decided to omit data from observer 6 (very low correlation coefficient and highest coefficient of variation) in all subsequent analysis. As a first approach to analyse our results we computed the mean of the observers' responses for each pairwise comparison. We considered that a method was selected if the mean of the encoded decisions, computed for all 9 observers, is greater than 0.5 (when the method was encoded as 1) or lower than 0.5 (when the method was encoded as 0). The performance does not vary significantly if we do not consider the cases where the average

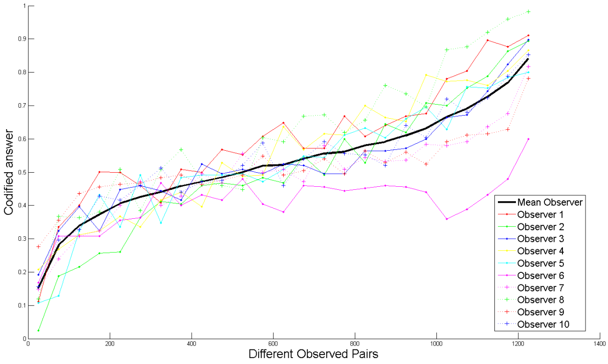


Figure 4.6: Comparison to the mean observer (black line).

Observer	1	2	3	4	5
Correlation	0.54	0.57	0.59	0.55	0.52
CV	52.49%	57.96%	37.65%	52.28%	52.69%

Observer	6	7	8	9	10
Correlation	0.23	0.48	0.63	0.61	0.55
CV	9.85%	47.12%	51.13%	25.36%	42.81%

Table 4.2: Correlation between each observer and mean observer.

value is too close to the chance rate (e.g. averages between 0.45 and 0.55). The results of these pairwise comparisons are given in Table 4.3. For each pair of methods, we show the percentage of cases where it has been selected against the others. Thus, results in Table 4.3 can be interpreted as follows: each method (in rows) is preferred a certain percentage of trials over the method in the columns. For example, Shades-of-Grey is preferred in 68.1% of the trials against Grey-world.

The percentages in Table 4.3 show that the images produced by Shades-of-Grey and Max-Name are preferred to those produced by Grey-World (68, 1% and 62, 4%). However, there is no clear preference when compared against each other (50.6% Shades-of-Grey preference vs. MaxName).

	Versus		
Selected	Shades-of-Grey	Grey-World	Max-Name
Shades-of-Grey	-	68.1%	50.6%
Grey-World	31.9%	-	37.6%
MaxName	49.4%	62.4%	-

Table 4.3: Results of the experiment in the 1-to-1 comparison.

Method	Wins
Shades-of.Grey	35.18%
Grey-World	16.33%
MaxName	39.28%
3-equally selected	8.92%

Table 4.4: Experiment results in a general comparison.

Method	Wins
Shades-of.Grey	42.65%
MaxName	36.39%
Grey-World	20.96%

Table 4.5: Results using Thurstone's Law of Comparative Judgement.

In Table 4.4 we show a global comparison of all algorithms (the percentages are computed for all 415 images). A method was considered a 'winner' for a given image if it was selected in two of the three comparisons. Methods were evaluated in the same way as we did for results in Table 4.3 (that is, a greater than a 0.5 mean value from all observers is encoded as 1). Evaluating this way, there are some cases where the three methods are equally selected (this happens in 8.92% of the images). This analysis was formulated in order to remove non-transitive comparisons (e.g. method A beats method B, method B beats method C and method C beats method A). Hence, we can conclude from these straightforward analyses that solutions from MaxName are preferred in general, but closely followed by Shades-of-Grey (39.28% and 35.18% respectively). We can also state that Grey-World solutions are the least preferred in general (with a low percentage of 16.63%). Moreover, the best angular error solution is selected in 42.96% of the cases.

We have also calculated the Thurstone's Law of Comparative Judgement [108] coefficients from our data (Table 4.5), obtained from the ordinal pairwise comparisons. Using this measure, results are not very different (Shades-of-Grey and MaxName are clearly better than Grey-World although the ranking changes) and images with minimal angular error are only selected in 45% of the cases.

Finally, we have computed two overall analyses (considering all scenes as one) in order to extract a global ranking for our colour constancy methods: the Thurston's Law of Comparative Judgement [108] and the Bradley-Terry [18] analysis. Table 4.6 shows the results of the Bradley and Terry's cumulative logit model for pairwise evaluations extended to ordinal comparisons [3]. These results are shown on the 'Estimate' column where the estimate reference has been set to 0 for the smallest value (Grey-World model). The standard error of this ranking measure shows that the two best models (Shades-of-Grey and MaxName) are better than Grey-World and arguably close to each other. Table 4.7 shows a similar analysis using Thurstone's Law of Comparative Judgement [108] and considering all scenes as one.

As we mentioned above, our experiment shows that images having minimum angular error with respect to the canonical solution are selected in fewer than half of the observations (when we ask people for the most natural image, the response, does not always correspond to the optimal physical solution). Moreover, this result is maintained even if we discard

Parameter	DF	Estimate	Std. Er	Wald 95% Conf. limits	Chi-Sq	Pr> >ChiSq
Shades-of-Grey	1	1.609	1.2231	-0.7882 4.0063	1.73	0.1883
MaxName	1	1.0256	0.8435	-0.6278 2.6789	1.48	0.2241
Grey-World	0	0	0	0 0	.	.

Table 4.6: Results using Bradley-Terry ordinal pairwise comparison analysis.

Parameter	DF	Estimate	Std. Er	Wald 95% Conf. limits	Chi-Sq	Pr> >ChiSq
Shades-of-Grey	1	0.196	0.0031	0.19 0.0021	4040.2	<0.0001
MaxName	1	0.1283	0.0031	0.1223 0.1343	1743.22	<0.0001
Grey-World	0	0	0	0 0	.	.

Table 4.7: Results using Thurstone's law of comparative judgement binary pairwise comparison analysis.

responses with low levels of certainty. In order to quantify this fact, in the next section we will introduce a new measure to complement the current performance evaluation of colour constancy algorithms.

4.4 Perceptual performance evaluation

Assuming the ill-posed nature of the problem, the difficulty of finding an optimal solution and the results of the present experiment, we propose an approach to colour constancy algorithms that involves human colour constancy by trying to match computational solutions to perceived solutions. Hence, we propose a new evaluation measurement, the Perceptual Angular Error, which is based on perceptual judgements of adequacy of a solution instead of the physical solution. The approach that we propose in this work does not try to give an alternative line of research to the current trends which focus on classifying scene contents to efficiently combine different methods: here we try to complement these efforts from a different point of view that we could consider as more 'top-down', instead of the 'bottom-up' nature of the usual research. As mentioned before, the most common performance evaluation for colour constancy algorithms consists in measuring how close their proposed solution is to the physical solution, independently of the other concerns. This has been computed as

$$e_{ang} = acos\left(\frac{\underline{p}_w \cdot \widehat{\underline{p}}_w}{\|\underline{p}_w\| \|\widehat{\underline{p}}_w\|}\right) \tag{4.6}$$

which represents the angle between the actual white point of the scene illuminant,$\underline{p}_w$, and the estimation of this point given by the colour constancy method, $\widehat{\underline{p}}_w$, which can be understood as a chromaticity distance between the physical solution and the estimate. The current consensus is that none of the current algorithms present a good performance on all the images [48], and a combination of different algorithms offers a promising option for further research. Our proposal here is to introduce a new measure, the perceptual angular error,

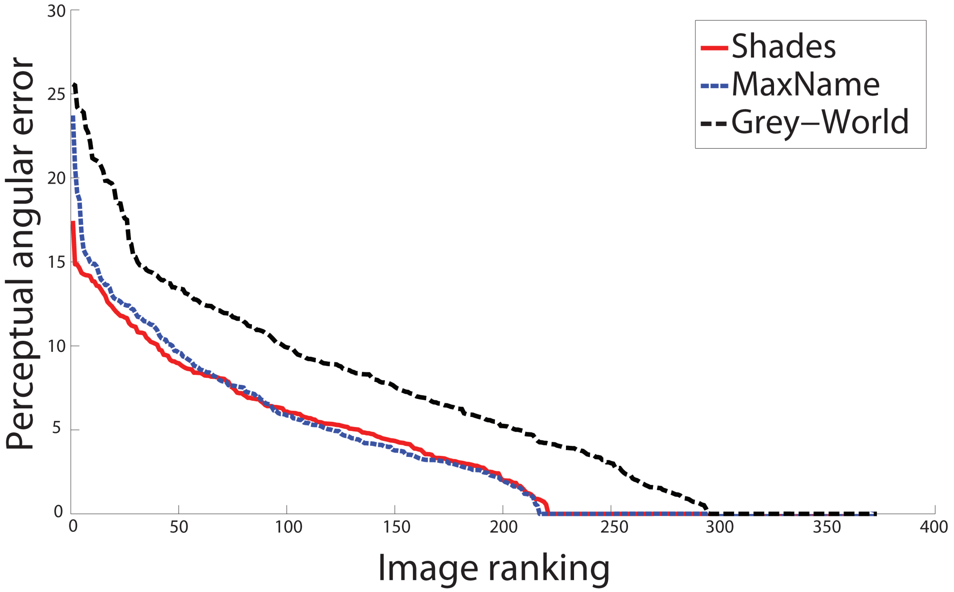


Figure 4.7: Estimated perceptual angular error (between method estimations and preferred illuminants).

e^p_{ang} , that would be computed in a similar way:

$$e_{ang} = acos\left(\frac{\underline{p}^p_w \cdot \widehat{\underline{p}}_w}{\|\underline{p}^p_w\|\|\widehat{\underline{p}}_w\|}\right) \tag{4.7}$$

where $\underline{p}^p_w$ is the perceived white point of the scene (which should be measured psychophysically) and $\hat{p}_w$ is an estimation of this point, that is the result of any colour constancy method, as in Equation 4.6. The difficulty of this new measurement arises from the complexity of building a large image dataset, where , the perceived white point of the images has been measured. In this work we propose a simple estimation of this perceived white point by considering the images preferred in the previous experiment. Hence, the perceived white point is given by the images coming from the colour constancy solutions that have been preferred by the observers. The preferred solutions, that is, the most natural solutions, can give us an approximation to the perceived image white point. Making the above consideration, in Figure 4.7 we can see how the estimation of the perceptual angular error works for the three tested algorithms. In the abscissa we plot a ranking of the observations in order to get the perceptual errors in descending order. In the ordinate we show the estimated perceptual angular error for each created image (that is, 415 different inputs to the algorithms). A numerical estimation of the perceptual angular error could be the area under the curves plotted in Figure 4.7. In the figure we can see that both Shades-of-Grey and MaxName work quite similarly, while Grey-World presents the highest perceptual error. This new measurement agrees with the conclusion we summarised in the previous section and provides a complementary measure to evaluate colour constancy algorithms. In Figure 4.8 we show a similar plot for the usual angular error.

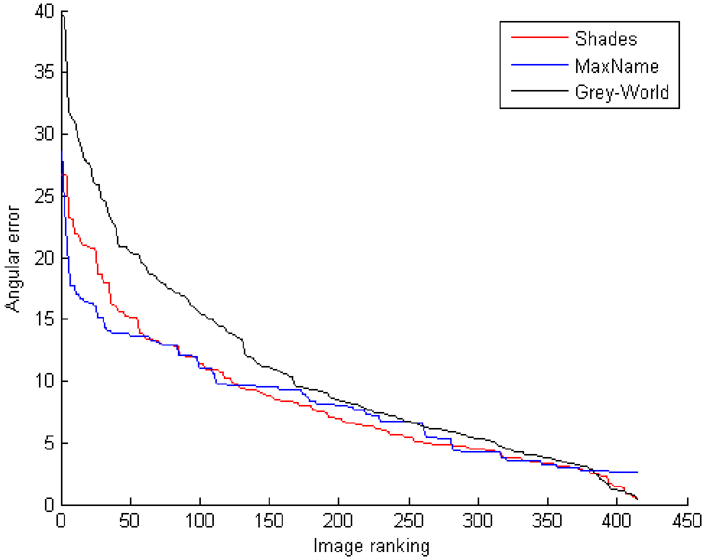


Figure 4.8: Angular error plot for the different methods on 415 images of the dataset.

	Mean	RMS	Median
Max-Name	7.64	8.84	6.78
Shades-of-Grey	7.84	9.70	5.95
Grey-World	10.05	12.70	7.75

Table 4.8: Angular error for the different methods on 415 images of the dataset.

In Tables 4.8 and 4.9 we show the different statistics on the computed angular errors. In Table 4.8, the angular error between the estimated illuminant and the canonical illuminant are shown. In this case, MaxName and Shades-of-Grey present better results than Grey-World. In Table 4.9 equal statistics are computed for the estimated perceptual angular error. The results on this table confirm the conclusions we obtained from Figure 4.7.

4.5 Conclusion

This chapter explores a new research line, the psychophysical evaluation of colour constancy algorithms. Previous research points out the need to further explore the behaviour of high-level constraints that are necessary for the selection of a feasible solution (to avoid the dependency of current evaluations on the statistics of the image dataset). With this aim in mind, we have performed a psychophysical experiment in order to compare three computational colour constancy algorithms: Shades-of-Grey, Grey-World and MaxName. The

	Mean	RMS	Median
MaxName	3.86	6.02	2.61
Shades-of-Grey	3.79	5.66	2.86
Grey-World	6.70	9.01	5.85

Table 4.9: Estimated perceptual angular error for the different methods on 415 images of the dataset.

results of the experiment show that the Shades-of-grey and MaxName methods have quite similar performance which is better than that obtained by the Grey-World method and that in almost half of the judgements; subjects have preferred solutions that are not the closest ones to the optimal solutions.

Considering that subjects do not prefer the optimal solutions in a large percentage of judgements; we have introduced a new measure, based on the perceptual solutions to complement current evaluations: the Perceptual Angular Error. It tries to measure the proximity of the computational solutions versus the human colour constancy solutions. The current experiment allows the computation of an estimate of the perceptual angular error for the three explored algorithms. However, our main conclusion is that further work should be done in the line of building a large dataset of images linked to the perceptually preferred judgements.

To this end a new, more complex experiment, perhaps related to the one proposed in [93], must be done in order to obtain the perceptual solution of the images, independently of the algorithms being judged.

Part II

Part 2: Sensor sharpening

Chapter 5

A new sensor basis to predict colour names and unique hues

In this chapter we define a new set of sensors which fit psychophysical data. These new sensors are defined in order to predict specific tasks performed by the human visual system. To this end, we base our work in a previous work done by Philipona and O'Regan. In this work surface colour is defined by the relationship between the visually accessible information under a set of different lights and the visually accessible information about the surface viewed under the same set of lights. This relationship is a 3 × 3 matrix transform, and from its eigenvalues, three reflection properties are derived. These reflection properties are related to the World Colour Survey (WCS) data and the unique hues via a singularity index.

In this chapter we make two contributions. First, we prove that the Philipona and O'Regan mathematical formulation is strongly related to spectral sharpening methods. Then, inserting the spectral sharpening idea in the Philipona and O'Regan model allows us to define a new set of sharp sensors capable of predicting psychophysical data. We also develop a new mathematical definition of singularity which has two additional properties to the one of Philipona and O'Regan: it is compact and it is related to a well-known colour measure.

5.1 Introduction

Opponent space has been defined as a confrontation of nonmixable colours. That is, it is impossible to perceive a reddish green or a yellowish blue. Red, green, yellow and blue are considered 'cardinal' colours and their hues are considered unique. In opponency theory [61] it is proposed that colour is encoded in our brain using three axes. The first axis represents the intensity, namely, black, grey and white. For convenience we can represent this channel as an interval [0,1] where 0 and 1 are black and white respectively. The second axis called red-green encodes values between these two colours: red might be coded by -1, green by +1 and 0 indicates an equilibrium point, that is, a colour that is neither reddish or greenish. Finally, the third yellow-blue channel measures the relative proportions of yellow or blue in a stimulus (where again we can use the numbers in [-1,1] to describe yellowness through the equilibrium point to blueness). In opponent theory yellow is perceived when the yellow-blue channel has a strong yellow response and the red-green channel is in equilibrium. The process to perceive red, green and blue can be analogously defined. However, there is not a widely accepted theory explaining the uniqueness of these four colours, since the opponent theory does not adequately predict hues perceived as unique [73], [110].

Whether the opponent theory is able to explain this fact or not, what is widely assumed is the asymmetry in human perception of different colour surfaces. Specific colour properties, such as red, green, yellow and blue, or possibly purple, orange and pink in other cultures, hold a special status in perception. Explanations for this fact could be essentially found in the neuronal representation of colour in the human visual system [91], or could be due to cultural or linguistic reasons [69], but it is an open issue.

How this asymmetric perception can be achieved in the human visual system has been studied by Philipona and O'Regan (from now on PO) [92]. In their work they explore the hypothesis of a representation that copes with the reflection properties of surfaces independently of the lighting conditions of the observation. They build a linear biological model by finding a linear constraint between the trichromatic representation of the illuminant and the trichromatic representation of the reflected light. This is a biological approach towards what physicists define as reflectance: the relationship between the spectrum of light illuminating a surface and the spectrum of light reflected by the surface. Practically, this is equivalent to the relation between the RGB values of a surface under different lights and an achromatic surface viewed under the same light set. For each surface, this linear model finds a matrix containing the reflectance properties which are illuminant invariant. They propose the eigenvalues of this matrix as a triplet representing the inherent reflectance properties of the surface. These coefficients are denoted as (r_1^s, r_2^s, r_3^s) where s represents the surface that is being represented with these three coefficients.

These reflectance coefficients are used to compute a singularity index that will quantify the degree of asymmetry of the corresponding surface. This index is built in such a way that it allows predicting the psychophysical data of the unique hues or the colour names of the WCS [1].

A major strength of PO formulation is that the reflection properties are independent of the sensor basis chosen. It will work equally well given any combination of the original human cones. A disadvantage is that the meaning of the reflection properties is not clear (they do not simply correlate with redness, greenness and blueness for example). Indeed, the three reflection parameters are without any explicit order.

In this chapter we generalise the PO biological model by defining a 3×3 matrix able to substitute the eigenvectors matrix in their model for all the surfaces. Then, this matrix represents a new set of sensors. In order to define these new sensors, we use the spectral sharpening technique proving that its formulation is related to the PO biological model.

Moreover, we also develop a new mathematical definition of singularity overcoming the drawbacks of the PO index. In particular, our measure will be compact, independent of the order of the reflectance properties and related to achromaticness.

This chapter is organised as follow. In section 5.2 we explain the biological model of PO. After this, in section 5.3 we formulate the singularity index defined by the same authors. In section 5.4 we define our model based on a set of sharp sensors. This is followed by section 5.5 where we formulate our new compact singularity index. In section 5.6 we summarise the different results obtained.

5.2 Philipona and O'Regan biological model

The linear biological model introduced in [92] is built on the assumption that the human vision system is able to extract the reflection properties of the world surfaces independently of the lighting conditions of the observation. In other words, it delivers a canonical representation of the reflectance.

This model is based on the computation of the LMS (or the CIE R,G,B) coordinates. These coordinates represent physical properties of the light reflected by a surface reaching the observer eye which lose part of the colour information due to the photopigments absorption. The authors refer to this as the accessible information [92].

The PO model finds a matrix containing the reflectance properties for each surface. This matrix is defined from the relation between the accessible information of the incident illuminant and the accessible information of the reflected light. From this matrix the authors are able to extract a colour triplet (reflectance) that is the colour of the surface independent from the illuminant.

To build the data they select a wide number of illuminants and reflectances and a set of photopigments sensitivities. For the photopigments sensitivities they used mainly the 10-deg Stiles and Burch colour matching functions (CMFs) [104], but, they observed that by using the Stockman and Sharpe [105] cone fundamentals no noticeable change is produced. That is, they showed that both sensitivities were equivalent. For the set of illuminants (from now on set E) they used the 99 daylight spectra from Romero *et al.* [99], a Gaussian sample of 200 spectra constructed from the basis functions $S0$, $S1$ and $S2$ derived by Judd *et al.* [68], and the 239 daylight spectra from Chiao *et al.* [24]. Finally, the reflectances used are the set of 1600 Munsell glossy chips from the Spectral Database of the University of Joensuu Colour Group [90].

The authors define what they name the visually accessible information for a single surface and a single light. Denoted as $\underline{v}^s_{\underline{e}}$, the accessible information about the reflected light for a given surface s is defined as

$$\underline{v}^s_{\underline{e}} = \int_\omega R_i(\lambda)E(\lambda)S(\lambda)d\lambda,\ i = 1,2,3 \tag{5.1}$$

where $\underline{e}$ is related to the illuminant of the scene, λ is defining wavelengths, $E(\lambda)$ is the spectral power distribution of the light in each wavelength, $R_i(\lambda)$ the absorption of photopigments present in L,M and S photoreceptors respectively and $S(\lambda)$ the reflectance of a surface. Second, they define $\underline{u}_{\underline{e}}$ as the accessible information about the incident illuminant

$$\underline{u}_{\underline{e}} = \int_\omega R_i(\lambda)E(\lambda)d\lambda,\ i = 1,2,3 \tag{5.2}$$

Authors repeat this for the N different lights defined. They arrange the N response vectors for the surface in a $3 \times N$ matrix, v^s, and for the lights in a $3 \times N$ matrix, u. These two matrices are related by finding the best 3×3 matrix transform A^s such that

$$v^s \simeq A^s \cdot u \tag{5.3}$$

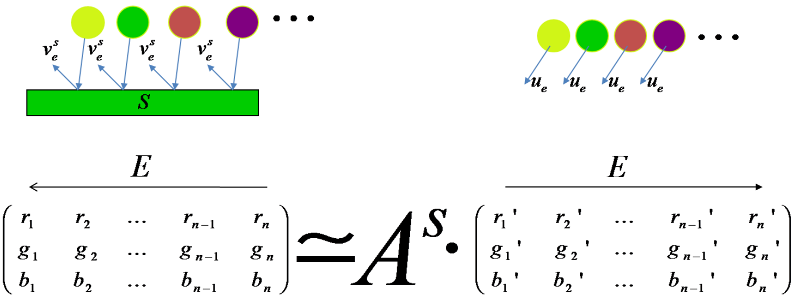


Figure 5.1: General scheme of the PO biological model, where u and v_s are respectively the accessible information about the incident light and the accessible information from the reflected light and A^s is the matrix that relates them

where the superscript s denotes dependence on a surface. Then, they solve for the matrix A^s by linear regression.

$$A^s = v^s[u]^+ \tag{5.4}$$

where $[]^+$ denotes the pseudoinverse. In Figure 5.1 we show a general scheme of the model.

Once the matrix A^s is found, they compute the eigenvalue/eigenvector decomposition of matrix A^s, this is

$$A^s = U^s R^s (U^s)^{-1} \tag{5.5}$$

where R^s is a diagonal matrix containing the eigenvalues of A^s and U^s is a 3×3 matrix containing the respective eigenvectors. Then, the model can be rewritten as follows

$$v^s \simeq U^s R^s (U^s)^{-1} \cdot u \tag{5.6}$$

PO show in their paper that U^s and R^s form a basis, and therefore, these eigenvalues are a colour triplet relating the surface reflectance and a white reflectance. From now on, we will denote these eigenvalues as a vector, $\underline{r}^s$, where the vector is composed by the values in the diagonal of R^s. This triplet allows the definition of a formulation that explains the location of WCS colour names and the unique hues as we will explain in section 5.5.

5.3 Philipona and O'Regan singularity index

PO propose a singularity index that correlates the reflection triplet found by their method with the red, green, yellow and blue colour names found in the WCS data and with the unique hues. They formulate this singularity index as follows. Let $\underline{r}^s = (r_1^s, r_2^s, r_3^s)$ denote the reflection properties of a surface. Without loss of generality it is possible to assume that they are in decreasing order ($r_1^s \geq r_2^s \geq r_3^s$).

PO define β_1^s and β_2^s as:

$$\beta_1^s = \frac{r_1^s}{r_2^s} \tag{5.7}$$

$$\beta_2^s = \frac{r_2^S}{r_3^s} \tag{5.8}$$

Finally, the singularity index is given by maximising:

$$SI = \max\left(\frac{\beta_1^s}{\max(\beta_1^s)}, \frac{\beta_2^s}{\max(\beta_2^s)}\right) \tag{5.9}$$

In summary, PO propose that colour names occur when this SI is high, that means that the reflection properties are, in some sense, singular. A triplet of reflection values are said to be singular when one or two of the values are close to 0 (and there is at least one value much larger than 0).

This singularity index has two main drawbacks. First, the formulation is cumbersome, and it requires the knowledge of the reflectance values for all the surfaces in order to be able to normalise the output. Second, it is not related to any known property of colour. For these reasons, in section 5.5 we define a new singularity index overcoming these problems.

5.4 Spectral sharpening to generalise the biological model

In section 5.2 we have defined the PO biological model. We have shown that the model can be explained by an equation of the form given in the equation 5.6

$$(U^s)^{-1} \cdot v^s \simeq R^s (U^s)^{-1} \cdot u \tag{5.10}$$

It is important to notice that in the PO model the basis of eigenvectors (U^s) is different for each surface. The question that arises here is if it is possible to find a matrix T that can substitute the different $(U^s)^{-1}$ for all s. If such a matrix exists, it will be defining a new set of sensors since, by linearity of the integral and using equation 5.1 and 5.2, we will have the following relations

$$\begin{gathered} T \cdot \int_\omega R_i(\lambda)E(\lambda)S(\lambda)d\lambda = R^s \cdot T \cdot \int_\omega R_i(\lambda)E(\lambda)d\lambda \Rightarrow \\ \Rightarrow \int_\omega (T \cdot R_i(\lambda))E(\lambda)S(\lambda)d\lambda = R^s \cdot \int_\omega (T \cdot R_i(\lambda))E(\lambda)d\lambda \end{gathered} \tag{5.11}$$

We notice that equation 5.10 is related to the one of the transformed diagonal mapping (equation 2.4). Then, we can find a sharpening matrix T for this problem applying the data-based sharpening method (section 2.1.3) to equation 5.10

$$\Delta^s = T \cdot v^s [T \cdot u]^+ \tag{5.12}$$

Rearranging this last equation in order to assure that Δ^s is diagonal we find

$$T^{-1} \Delta^s T = v^s [u]^+ \tag{5.13}$$

Then, T is related to the eigenvector decomposition of

$$v^s [u]^+ = U \Delta^s U^{-1} \tag{5.14}$$

Therefore, $T = U^{-1}$.

Applying this method to the different surfaces will give us a set of s different matrices $\{T_s\}$ very similar to each other. In order to select the best matrix over the set we search for the sharpening matrix T_σ that can substitute all the different U^s by minimising the error

$$v^s = T_\sigma R^s (T_\sigma)^{-1} u \;\forall s \tag{5.15}$$

The selection of T_σ is based on two criteria. Firstly, we discard the different matrices T_s that have complex values, and secondly, we select the matrix that obtains a singularity index for Δ^s values closest to the singularity index obtained by PO. Mathematically, let us define, $\{\Phi_l\}_{l=1,\cdots,L} = \{T \in T_s | \forall t_{ij} \in T, t_{ij} \in \Re\}$. Then,

$$T_\sigma = \arg\max_l \operatorname{corr}(SI([r_1^s, r_2^s, r_3^s]), SI(\Phi_l \cdot v^s [\Phi_l \cdot u]^+)) \tag{5.16}$$

where corr represents the lineal correlation coefficient.

The matrix T_σ defines a new set of sharp sensors predicting psychophysical data. The sharp sensors obtained by applying T_σ are very similar to those shown in Figure 5.2.

The results in Figure 5.2 are the improved version of the sensors defined by T_σ by applying spherical sampling (section 2.1.5). Using spherical sampling we can minimise the error committed by T_σ. To this end, we define a new set of sharpening matrices $\{\Gamma_j\}_{j=1,\cdots,J}$ around the original matrix T_σ.

Again, we search the matrix whose reflection properties are close to those of PO. To this end we first select matrices where (r_1^s, r_2^s, r_3^s) and Δ^s (ordering the values) are sufficiently close. This means, we reduce the set $\{\Gamma_j\}$ into one called $\{\Upsilon_k\}_{k=1,\cdots,K}$

$$\{\Upsilon_k\} = \left\{ \quad T \in \Gamma_j | \frac{\sum_{s=1}^{\#S} d([r_1^s, r_2^s, r_3^s], T \cdot v^s [T \cdot u]^+)}{\#S} \leq \alpha \right\} \tag{5.17}$$

where S is the set of reflectances and $\#S$ is the cardinality of the set.

Then, we apply the selection based on the correlation as before:

$$T_B = \arg\max_k \operatorname{corr}(SI([r_1^s, r_2^s, r_3^s]), SI(\Upsilon_k \cdot v^s [\Upsilon_k \cdot u]^+)) \tag{5.18}$$

By using this procedure we find the best sharpening matrix T_B

$$T_B = \begin{pmatrix} 1.2566 & -1.1480 & 0.1235 \\ -0.3057 & 1.6502 & -0.3276 \\ -0.0169 & 0.0219 & 1.0659 \end{pmatrix} \tag{5.19}$$

we will call this matrix the biological sharpening matrix. This sharpening matrix defines the new set of sensors that subsumes the PO biological model, which are shown in Figure 5.2

It is worth remarking here that by using our sensors we obtain the reflectance information for each sensor. That is, the reflection properties are related to specific colour sensations defined by a unique basis that can be understood as a linear combination of the cone sensitivities. This allow us to avoid one of the biggest problems in the PO formulation, where reflectance properties did not carry any information about which colour is representing each value, since they do not refer to a unique basis.

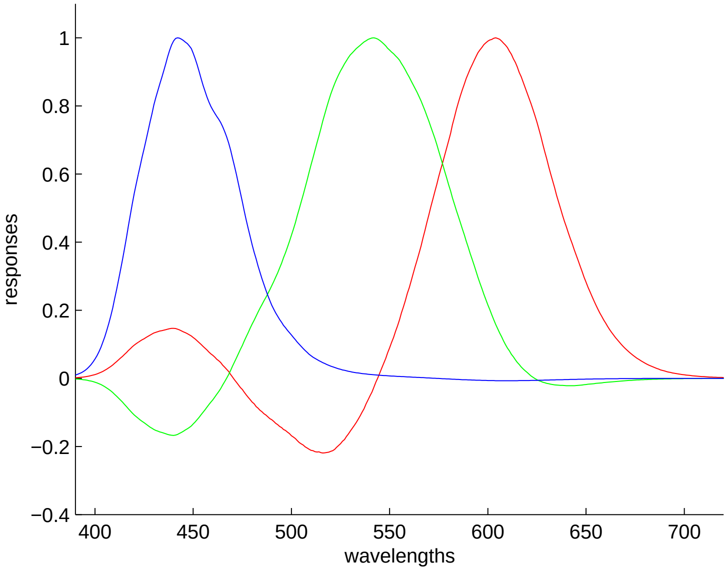


Figure 5.2: Sharp sensors predicting psychophysical data.

5.5 A new compact singularity index

In this section we propose a new singularity index that pursuits a simpler and more compact formulation than the one proposed by PO (section 5.3) with some specific properties. The first property we want to fulfil is that our new function should be independent of the order of the values.The second property we want to fulfil is to normalise it independently of which is the maximum value of the components. Our proposal is to increase the importance of a particular coefficient over the other two by means of a mathematical function. To this end, we propose to use a cubic function normalised by the product of the components, that is, we compute the terms:

$$I_1 = \frac{(r_1)^3}{r_1 \cdot r_2 \cdot r_3} \tag{5.20}$$

$$I_2 = \frac{(r_2)^3}{r_1 \cdot r_2 \cdot r_3} \tag{5.21}$$

$$I_3 = \frac{(r_3)^3}{r_1 \cdot r_2 \cdot r_3} \tag{5.22}$$

Let us note that we we will no longer use the superscript s to relate reflection properties with a particular surface.

These last functions allow as to normalise and boost the different coefficients. Then, these functions can be simply combined by a sum. In this case, if the surface has a singularity it will be reflected in at least one of these three components, and it will eventually appear in the sum, hence our Compact Singularity Index (CSI) is given by

$$CSI = I_1 + I_2 + I_3 = \frac{(r_1)^3 + (r_2)^3 + (r_3)^3}{r_1 \cdot r_2 \cdot r_3} \tag{5.23}$$

Let us now continue explaining different properties that can be derived. Let us note that achromatic surfaces will have three equal reflection coefficients and will fall on the diagonal axis of the space defined by these coefficients. Then, in this space, a chromatic measure can be computed as the determinant of the following matrix

$$M = \begin{pmatrix} r_1 & r_2 & r_3 \\ r_2 & r_3 & r_1 \\ r_3 & r_1 & r_2 \end{pmatrix} \tag{5.24}$$

that is given by

$$det(M) = (r_1)^3 + (r_2)^3 + (r_3)^3 - 3 \cdot r_1 \cdot r_2 \cdot r_3 \tag{5.25}$$

whose normalisation brings to the compact singularity function

$$\begin{aligned} \frac{det(M)}{r_1 \cdot r_2 \cdot r_3} &= \frac{(r_1)^3 + (r_2)^3 + (r_3)^3 - 3 \cdot r_1 \cdot r_2 \cdot r_3}{r_1 \cdot r_2 \cdot r_3} && (5.26) \\ &= \frac{(r_1)^3 + (r_2)^3 + (r_3)^3}{r_1 \cdot r_2 \cdot r_3} - 3 && (5.27) \\ &\propto \frac{(r_1)^3 + (r_2)^3 + (r_3)^3}{r_1 \cdot r_2 \cdot r_3} && (5.28) \\ &= CSI && (5.29) \\ & && (5.30) \end{aligned}$$

Another interesting property is the independence of our CSI from intensity if it is considered as a colour representation, that is

$$\begin{aligned} &\frac{(s \cdot r_1)^3 + (s \cdot r_2)^3 + (s \cdot r_3)^3}{(s \cdot r_1) \cdot (s \cdot r_2) \cdot (s \cdot r_3)} = \\ = &\frac{s^3((r_1)^3 + (r_2)^3 + (r_3)^3)}{s^3 \cdot (r_1 \cdot r_2 \cdot r_3)} = \frac{(r_1)^3 + (r_2)^3 + (r_3)^3}{r_1 \cdot r_2 \cdot r_3} \end{aligned} \tag{5.31}$$

Finally, we introduce another interesting property of this formulation, since it can be seen as an approximation of the perceptual space given by

$$r_1 = (\rho_1)^{\frac{1}{3}}, r_2 = (\rho_2)^{\frac{1}{3}}, r_3 = (\rho_3)^{\frac{1}{3}} \tag{5.32}$$

Hence, by replacing equation 5.32 in equation 5.23 we found

$$CSI = \frac{(r_1)^3 + (r_2)^3 + (r_3)^3}{r_1 \cdot r_2 \cdot r_3} = \frac{\rho_1 + \rho_2 + \rho_3}{(\rho_1)^{\frac{1}{3}} (\rho_2)^{\frac{1}{3}} (\rho_3)^{\frac{1}{3}}} =$$
$$= \frac{\rho_1 + \rho_2 + \rho_3}{(\rho_1 \cdot \rho_2 \cdot \rho_3)^{\frac{1}{3}}} \propto \frac{\text{aritmean}}{\text{geomean}} \tag{5.33}$$

where aritmean refers to the arithmetic mean and geomean refers to the geometric mean in a perceptual space.

From our first index, another interesting analysis can be made by combining in a different way the values of I_1, I_2 , I_3. In particular we will define two other indices, and we will denote them as the normalised compact singularity index and the simplified compact singularity index

$$CSI_{norm} = \frac{I_1}{\max(I_1)} + \frac{I_2}{\max(I_2)} + \frac{I_3}{\max(I_3)} \tag{5.34}$$

$$CSI_{simp} = \frac{\max((r_1)^3, (r_2)^3, (r_3)^3)}{r_1 \cdot r_2 \cdot r_3} = \max(I_1, I_2, I_3) \tag{5.35}$$

5.6 Results

PO showed that their model and singularity index are able to predict colour naming, unique hues and hue cancellation by using three experiments. In this chapter we are going to use the two first experiments of their paper.

Their first experiment was related to the fitting of the WCS data. PO computed their singularity index in the 40 Munsell chips that are typically used in colour naming experiments. They compared the results of their index with the data obtained for these particular chips in the WCS data.

Their second experiment was related to the location of the unique hues. Unique hues are defined as the wavelengths that human perceive without any trace of other tints. These hues are usually found in experiments where observers face *aperture colours*. The stimuli used in these experiments are generated by sending lights of controlled spectral composition directly into the eye. PO found the unique hues estimated by their method as follows. They computed the singularity index for all the surfaces in the Munsell book and those from some other databases. They then related these singularities with the x,y 1931 values for the respective surfaces. Finally they related the peaks of their index in the x, y space with the wavelength associated to that point. In Table 5.1 we can see the different studies about unique hues and the PO prediction (this table has been extracted from the PO paper).

5.6.1 Comparison of PO's model vs SS model

This section tests whether our newly derived sensors account for the colour names and unique hues in a similar manner as the PO biological model. To this end, we will use the singularity index defined by PO for both models.

Let us start by fitting the WCS data. In Figure 5.3 we can observe the comparison between the surface created by the singularity index using the PO model on the WCS grid

		Unique	yellow	Unique	green
Dataset	Subjects	Mean(nm)	Range(nm)	Mean(nm)	Range(nm)
Schefrin	50	577	568-589	509	488-536
Jordan-Mollon	97	–	–	512	487-557
Volbrecht	100	–	–	522	498-555
Webster (a)	51	576	572-580	544	491-565
Webster (b)	175	580	575-583	540	497-566
Webster (c)	105	576	571-581	539	493-567
PO Prediction	–	575	570-580	540	510-560

		Unique	blue	Unique	red
Dataset	Subjects	Mean(nm)	Range(nm)	Mean(nm)	Range(nm)
Schefrin	50	480	465-495	–	–
Jordan-Mollon	97	–	–	–	–
Volbrecht	100	–	–	–	–
Webster (a)	51	477	467-485	EOS	–
Webster (b)	175	479	474-485	605	596-700
Webster (c)	105	472	431-486	EOS	–
PO Prediction	–	465	450-480	625	590-EOS

Table 5.1: Unique hues found in the different experiments and the PO prediction. EOS means End Of Spectrum.

(Figure 5.3.(a)) and the results of our sensors (Figure 5.3.(b)). We can observe that they are almost identical. In Figure 5.4 the contour plots of the previous figures are shown. In this way we can compare the location of the different maxima in the grid.

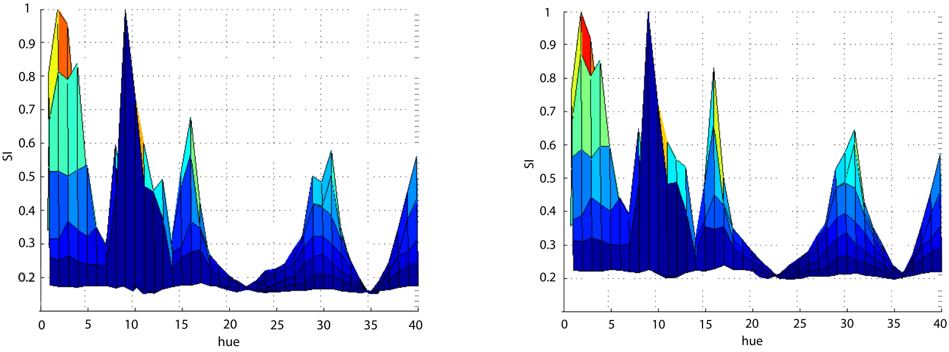
Let us also check the second experiment performed by PO in order to locate the unique hues. To this end, as explained before, we are going to locate the singularity indices computed using both models in the 1931 x, y space. In Figure 5.5 we show the singularity index for the different x,y values for both models, and in Figure 5.6 we can observe a contour plot of this singularity index.

Finally, with our new approach we obtain the spectral information of the sensors, therefore, we can apply the singularity index directly to these. In Figure 5.10.(a) we show the spectral result for the singularity index applied directly to the sensors. We can observe how the four unique hues pop out, and that they are very similar to the values in Table 5.1.

5.6.2 Comparison PO's SI vs our CSI

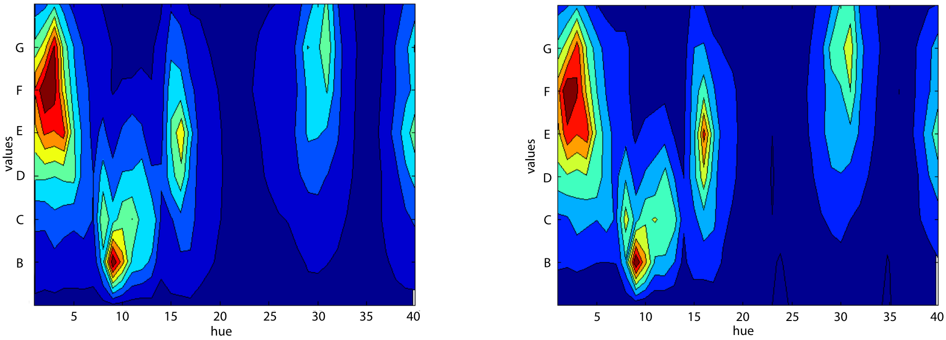
In this section we show that our new compact singularity index fits the data as precisely as the PO singularity index with the additional benefit that our index gives extra colour information.

We begin with the comparison of how well the PO singularity index and our new index fit the WCS data. In Figure 5.7 we can see the coloured surface representing the results of the WCS experiment together with the contour lines for both indices. On Figure 5.7.(a) we show the contour lines for the SI index proposed by PO and on Figure 5.7.(b) the contour



(a) (b)

Figure 5.3: Prediction of the PO singularity index on the WCS data through the hue axis (a) PO model results (b) sharp sensors.



(a) (b)

Figure 5.4: Prediction of the PO singularity index on the WCS data through the hue-value axis (a) PO model results (b) sharp sensors results.

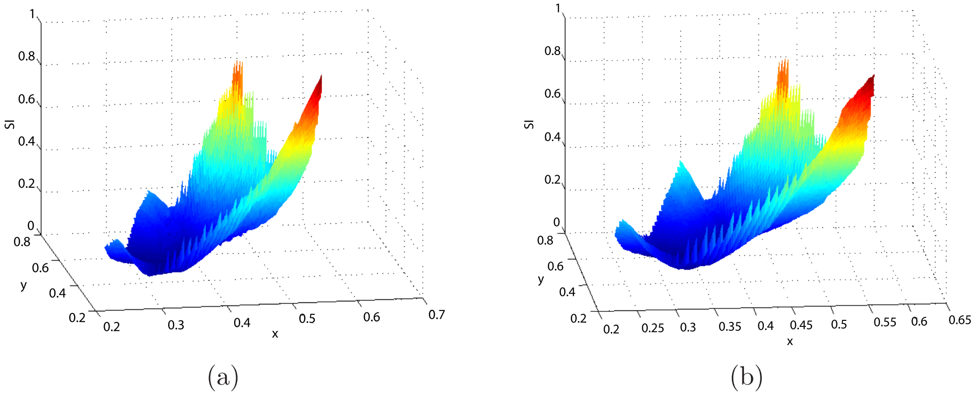


Figure 5.5: Prediction of the PO singularity index on the unique hues data on the x,y chromaticity plane (a) PO model results (b) sharp sensors results.

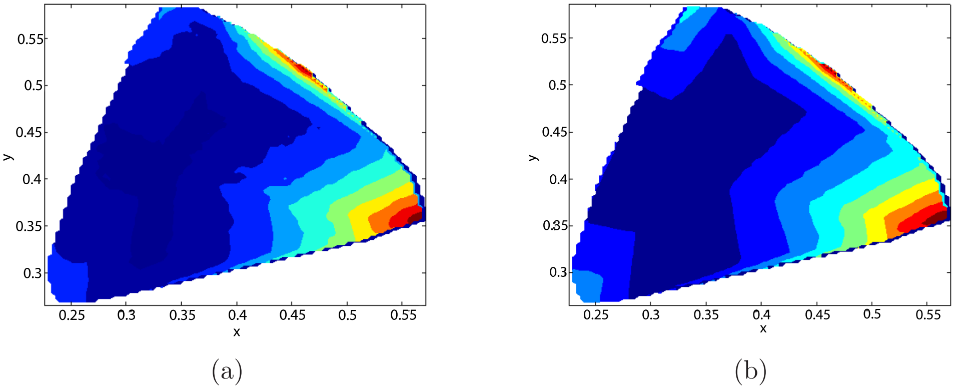
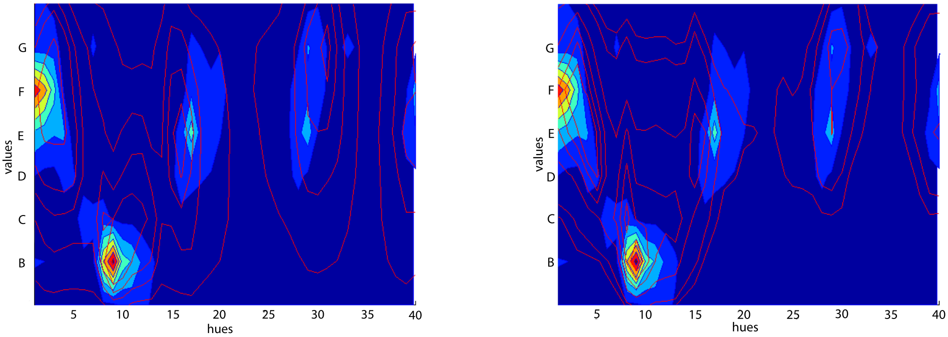


Figure 5.6: Prediction of the PO singularity index on the unique hues data on the x,y plane plane (a) PO model results (b) sharp sensors results.



(a) (b)

Figure 5.7: Comparison of the WCS data (background) versus: (a) the PO's index, (b) our index (right).

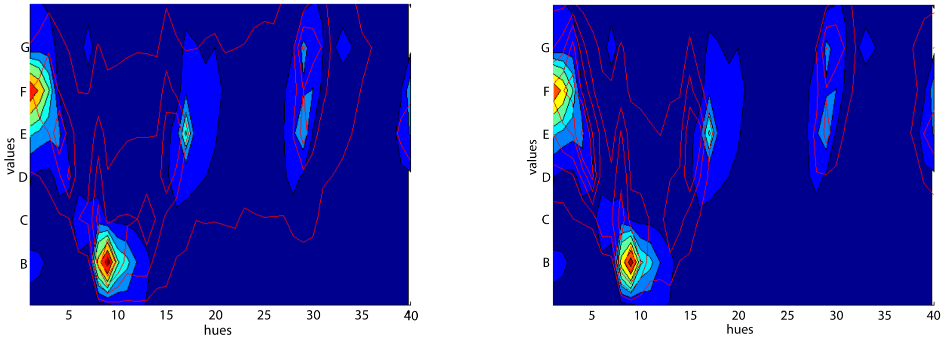
lines for our CSI index. We can see how our new index outperforms the previous one for the blue colour, while the fitting of the rest is quite similar. Moreover in Figure 5.8.(a) we can see the behaviour of our normalised index, where again blue is giving a better fitting than in the PO index, but red in this case is fitted worse. Finally, in the simplified version (Figure 5.8.(b)) red is not detected correctly while blue is perfectly located.

We also tested the ability of our new compact singularity index to locate unique hues. In this case, we have applied two procedures. The first procedure is the same used by PO (related to the x, y space). Results for this procedure with the CSI index are shown in Figure 5.9. Plots in this figure should be compared with Figures 5.5 and 5.6. The second procedure consists of directly applying the different indices to the sharp sensors in order to extract a spectral function locating the unique hues. In this case we applied the three different versions of our index and the results are shown in Figure 5.10.(b),5.10.(c),5.10.(d). The wavelength values where the unique hues arise for these indices are reported in Table 5.2.

5.6.3 Invariant reflectance properties on natural images

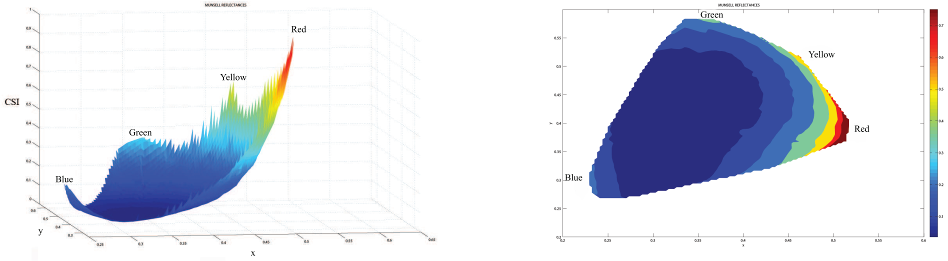
One of the main advantages of our formulation is the possibility of working on hyperspectral images directly. This can be done by applying equation 5.11 to a multispectral image $S(\lambda)$ knowing that when using our new sensors the reflection properties can be found just by dividing v_s and u component by component.

We have computed the reflectance properties on some images of the multispectral dataset acquired by Foster and Nascimento [47], [84]. In Figure 5.11 we show different reflection properties for some images (left column) and their corresponding singularity index (right column). We can appreciate how the singularity index provides the information of unique hues and the focal colours from the scenes.



(a) (b)

Figure 5.8: WCS Data (background) versus (a)our normalised index and (b) our simplified index.



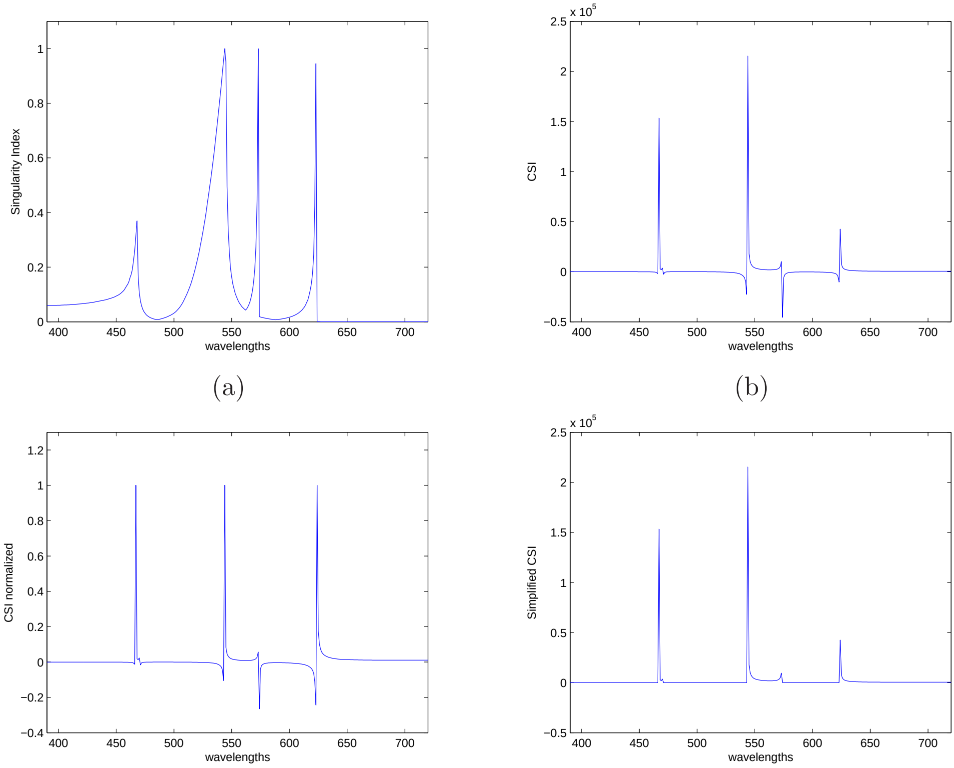
(a) (b)

Figure 5.9: CSI index in the x, y space: (a) surface plot, (b) contour plot.

	Unique	yellow	Unique	green
Index	Mean(nm)	Range(nm)	Mean(nm)	Range(nm)
CSI	574	567-577	544	538-554
Normalised CSI	574	567-577	544	538-554
Simplified CSI	573	568-580	544	537-550
PO	573	570-574	544	515-555

	Unique	blue	Unique	red
Index	Mean(nm)	Range(nm)	Mean(nm)	Range(nm)
CSI	467	452-475	624	610-EOS
Normalised CSI	467	450-476	624	615-EOS
Simplified CSI	467	460-480	624	620 -650
PO	468	456-470	623	610-625

Table 5.2: Unique hues located by our indices applied directly to the sensors versus the Philipona index applied to the sensors. EOS means End Of Spectrum.



(a) (b)

(c) (d)

Figure 5.10: Unique hues found by (the PO index) (b) our CSI index, (c) our normalised index and (d) our simplified index (right).

Figure 5.11: Reflection properties of some images of the Foster and Nascimento dataset (left) and their singularity index (right.)

5.7 Conclusions

The main novelty of this chapter is the definition of a new set of sensors using the spectral sharpening technique. These new sharp sensors allow us to predict unique hues and WCS data in the same manner as the PO model did.

An advantage of this application relies on the fact that we can obtain a unique basis to represent all reflection properties of any surface, more independently than how they were obtained in PO's model, where reflection properties where defined in a different basis for each surface.

In the PO model the appearance of a specific linguistic colour name was linked to the notion of a singularity index. The singularity index was a function of the mapping taking the response of a surface measured under all lights to corresponding responses for a white surface. The mapping was a 3 by 3 matrix and the singularity function was large when the eigenvalues of this matrix had at least one eigenvalue significantly smaller than the other two. In our work the rather complex and indirect method of PO was replaced by finding a single sharpen sensor basis with respect to which unique colours correspond to RGB values which have a similar (yet much simpler) singularity property: at least one of the R, G and B responses is relatively small. Our compact singularity index can be seen as an achromatic measure, that is, colours with a high chromatic value in these reflection components are the ones considered as unique hues and basic colour terms.

Chapter 6

Experiments using spherical sampling

In this chapter we define optimal combinations of sensors that better deal with some measurements. We show that these new sensors are sharp and that they can be found by the use of the spherical sampling technique (Section 2.1.5). This technique allows us to obtain all the sharpening matrices close to a given one providing a refinement. In particular, we use it to minimise the CIE Delta E measure and maximise colour ratio stability. Lastly, we extend the spherical sampling paradigm to the multispectral case, where the objective is to model the interaction of light and surface without using all the spectra.

6.1 Introduction

In the previous chapters we have shown that the spectral sharpening procedure has been effective in computational colour constancy for making the diagonal model work better. In that case, the error measure used for selecting the best approximation has been usually based on a least squares criteria (see section 2.1.7). In this chapter we apply spherical sampling to other different computational colour problems where the solution is also a set of sharp sensors. In particular the problems treated in this chapter are i) to find the sharpening matrix that best minimises the CIE Delta E error measure ii) to find the sharpening matrix that best maximises colour ratio stability and iii) to find the best sharpening matrix that allows us to compute multispectra processing without using the whole spectra.

The first problem is related to finding the sharpening matrix that minimises the error perceived by humans when using a diagonal model. The error measure usually defined takes into account a minimisation criteria useful for computations. But, colour constancy is a human property, and then, it is interesting to study which are the sharp sensors that better correlate to human perception.

The second experiment is related to relational colour constancy. Relational colour constancy is defined as the invariance of perceived ratios between the colours of surfaces under different illuminants. In other words, the ratio between colours is stable under different illuminants. This is a different formulation from the approach followed in colour constancy (see section 1.3). For this reason, this experiment complements the thesis by studying the sharp sensors from a different colour constancy perspective.

Finally, the third experiment is useful in different computer graphics problems such as rendering [30]. In these problems different calculations, like multiplications, over the entire spectrum are carried out. Sharp sensors can be used in this problem in order to compute a basis to represent the different spectral functions. Then, the different computations needed can be done in this new basis reducing the complexity of the problem. Later on, the result can be expressed back in a multispectral form.

In this chapter we show that the spherical sampling technique outperforms state-of-the art in these three different problems.

6.2 Experiment 1: Minimising CIE Delta E

This experiment searches for the set of sharp sensors that best correlates with human perception when applying the diagonal model. To this end, this experiment considers a minimisation related to a perceptually relevant space, in particular the CIELab space, instead of the usual RGB space.

The experiment works as follows. We first calculate the colour descriptor for a reflectance $S(\lambda)$ under the $D65$ illuminant and the colour matching functions R_k following equation 1.9. We denote this value as $\underline{b}$. It is then converted to its CIELab representation, we denote it as $\underline{a}$

$$\underline{a} = Lab(\underline{b}) = Lab\left(\int_\omega S(\lambda)D65(\lambda)R_k(\lambda)d\lambda\right) \tag{6.1}$$

Alongside, we compute the value for the same reflectance $S(\lambda)$ under another illuminant $E(\lambda)$. We denote it as $\underline{c}$.

$$\underline{c} = \int_\omega S(\lambda)E(\lambda)R_k(\lambda)d\lambda \tag{6.2}$$

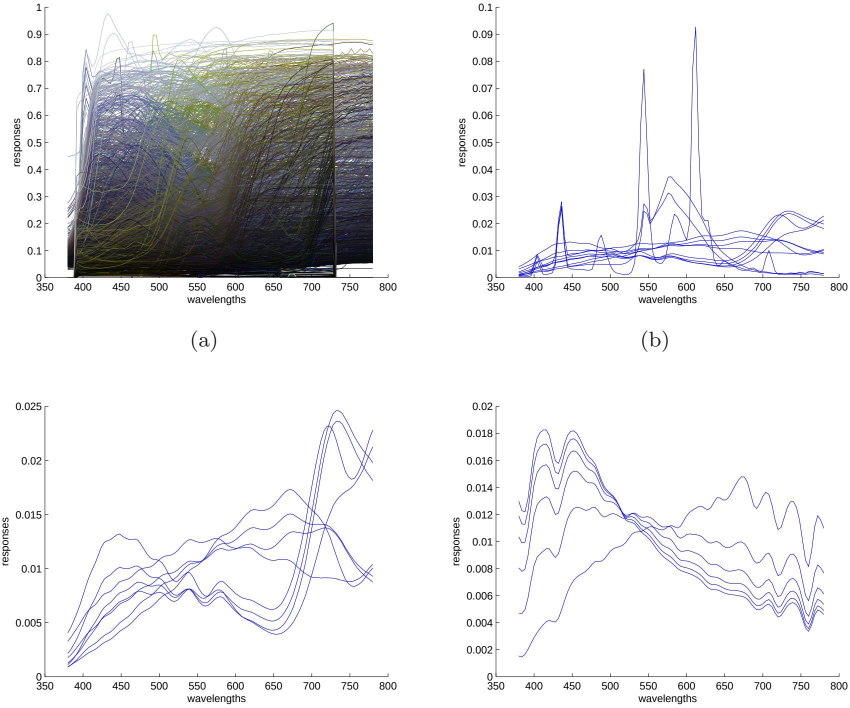
By using equation 2.4 we find a transformation matrix T and a diagonal matrix D from the colour Matching Functions that relate illuminant E with illuminant $D65$ in a diagonal form. Thus, we can compute an approximation of $\underline{b}$ from the value $\underline{c}$ by converting this value to the $D65$ illuminant. We denote this approximation as $\underline{b}^E$

$$\underline{b}^E = T^{-1} \cdot D \cdot T \cdot \underline{c} \tag{6.3}$$

Now, we also find an approximation to $\underline{a}$ by converting $\underline{b}^E$ to CIELab. We call this new approximation $\underline{a}^E$. The error between $\underline{a}$ and $\underline{a}^E$ is defined by Δ_ϵ, that is, the Euclidean distance between them, which is a perceptual measure. Formally,

$$\Delta_\epsilon = \|\underline{a} - \underline{a}^E\| = \|Lab(\underline{b}) - Lab(\underline{b}^E)\| = \|Lab(\underline{b}) - Lab(T^{-1} \cdot D \cdot T \cdot \underline{c})\| \tag{6.4}$$

The experiment has been performed with three different datasets. All of them share the same set of reflectances. These reflectances are the 1995 reflectances defined by Barnard *et*



(a) (b)

(c) (d)

Figure 6.1: (a) Reflectances in the three datasets, (b) illuminants in dataset 1, (c) illuminants in dataset 2, (d) illuminants in dataset r.

al. [9]. The set of reflectances is composed by the 24 Macbeth color checker patches, 1269 Munsell chips, 120 Dupont paint chips,170 natural objects [117], the 350 surfaces in Krinov Dataset [72], and 57 additional surfaces.

The difference between the three datasets are the illuminant sets used. In the first dataset we have used the 11 illuminants from Barnard *et al.* [9]. Illuminants in the second dataset are a subset of the previous ones. They are the 8 illuminants which represent daylight emulations. Finally, for dataset 3 we have defined a set of 6 planckian illuminants (4000K, 6000K, 8000K, 10000K, 12000K and 14000K).

In Figure 6.1(a) we can see an overlapped plot of the spectral composition of the reflectances, and the spectral composition of the illuminants for dataset 1, dataset 2 and dataset 3 in Figures 6.1(b), 6.1(c),6.1(d) respectively.

The transformation matrices we have compared against our spherical sampling approach are: Bradford, HPE, Fairchild and CAT02 as experimentally learned and the measurement tensor method as one that is theoretically found. The initial colour basis matrix for the measurement tensor method has been set to the identity matrix. In Table 6.1 we can see how the spherical sharpening outperforms all the previous methods. Results are calculated by computing the mean of Δ_ϵ for all the surfaces.

Method	Dataset 1	Dataset 2	Dataset 3
Spherical sampling	**2.20**	**0.93**	**0.64**
Measurement Tensor	2.46	0.98	0.69
Bradford	2.59	1.14	0.84
Fairchild	2.32	1.12	0.85
Cat02	2.38	1.26	0.94
HPE	3.62	1.86	1.39

Table 6.1: Results of the first experiment in the three different datasets (measure: mean Δ_ϵ).

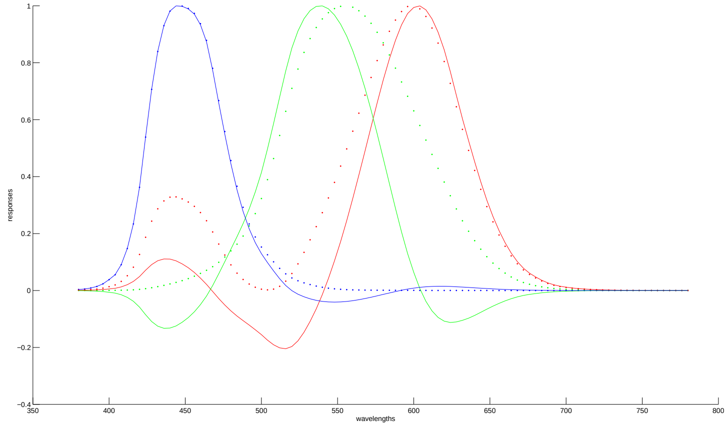


Figure 6.2: Original sensors -dotted lines- compared to spherical sampling sensors -solid lines- for the second dataset in the first experiment.

In order to have a quantified measure about how different are the new sensors from the previous ones, we have computed the angular distance between the spherical sampling sensors and each one of the others over the sphere (equation 2.24). Results can be seen in Table 6.2. Finally, in Figure 6.2 we can see the original sensors -dotted lines- versus the sensors resulted from the spherical sampling method -solid lines-.

6.3 Experiment 2: Maximising colour ratio stability

Relational colour constancy assumes that the relations between the colours of the surfaces are invariant [46]. This is quite a different idea compared with conventional colour constancy. Relational colour constancy is related to Retinex. Different studies have corroborated the importance of cone ratios in human colour constancy [45], [85], [16]. In order to find the sensors that best achieve relational colour constancy we have designed an experiment maximising colour ratio stability.

The experiment has been individually performed for each sensor, under the assumption

Method		Dataset 1			Dataset 2			Dataset 3	
	R	G	B	R	G	B	R	G	B
Meas. Tensor	5.47°	6.64°	0.43°	1.87°	1.25°	0.24°	0.19°	3.23°	1.34°
Bradford	21.45°	10.08°	1.79°	25.50°	7.70°	1.42°	24.05°	11.22°	1.41°
Fairchild	23.92°	8.48°	1.13°	27.94°	6.74°	2.00°	26.40°	10.51°	2.00°
CAT02	27.42°	11.22°	0.78°	31.46°	9.18°	2.74°	29.97°	12.85°	3.74°
HPE	38.41°	24.40°	0.13°	42.45°	21.82°	3.25°	41.06°	24.87°	3.25°

Table 6.2: Angular distance between the sensors mentioned and the sensors found by spherical sampling in the first experiment.

Method		Dataset 1			Dataset 2			Dataset 3	
	R	G	B	R	G	B	R	G	B
Spherical s.	**0.0569**	**0.1060**	**0.0595**	**0.0397**	**0.0606**	**0.0331**	**0.0585**	**0.0549**	**0.0395**
Meas. Tensor	0.1174	0.3529	0.0779	0.1498	0.2209	0.2529	0.1818	0.2609	0.2775
Bradford	0.0838	0.1252	0.1598	0.0844	0.0688	0.1428	0.1022	0.0689	0.1342
Fairchild	0.0733	0.2767	0.2074	0.0676	0.2071	0.2132	0.0647	0.1944	0.1623
CAT02	0.0829	0.1139	0.1576	0.0792	0.0750	0.1817	0.0710	0.1259	0.1482
HPE	0.1171	0.1299	0.0660	0.1193	0.1233	0.0588	0.1162	0.1264	0.0588

Table 6.3: Results of the second experiment in the three different datasets.

that the ratio stability for one sensor is independent of the other two sensors.

Let us now define, a vector $\underline{b}$ $(m \times 1)$ containing the colours for a set of m reflectances under the $D65$ illuminant with a given sensor (following equation 1.9). This means $b = [\rho_1, \cdots, \rho_m]$. Then, we define the vector of colour ratios $\underline{a}$, with dimension $(m-1)^2$ as follows

$$\underline{a} = \left[\frac{\rho_i}{\rho_j}; \frac{\rho_i}{\rho_{j+1}}; \cdots\right]; \quad \rho_i, \rho_j \in \underline{b}, \quad \rho_i \neq \rho_j. \tag{6.5}$$

If we have a second vector of ratios for the same reflectances under a different illuminant $\underline{a}^e$ the total ratio error is defined by

$$\epsilon = \frac{1}{n}\sum_{e=1}^{n} \frac{\|\underline{a} - \underline{a}^e\|}{\|\underline{a}\|}. \tag{6.6}$$

We have performed this experiment with the same datasets as in experiment 1 and we have used the same methods. The results showed in Table 6.3 are the mean of the ϵ measure. In the table we can observe that spherical sampling outperforms all the other methods.

We have also computed the angular distance between the spherical sampling sensors and the rest of the sensors. Results are reported in Table 6.4. Finally, we can also see in Figure 6.3 the original sensors compared to the sensors found by spherical sampling.

Method		Dataset 1			Dataset 2			Dataset 3	
	R	G	B	R	G	B	R	G	B
Meas. Tensor	14.22°	8.38°	4.29°	11.21°	7.81°	3.85°	25.24°	7.93°	3.85°
Bradford	11.44°	1.26°	1.77°	14.37°	0.57°	1.32°	2.01°	1.07°	1.32°
Fairchild	13.68°	3.83°	1.14°	16.71°	2.92°	0.74°	2.68°	3.78°	0.73°
CAT02	17.26°	2.10°	0.83°	20.26°	1.59°	1.25°	6.30°	2.47°	1.25°
HPE	28.49°	13.10°	0.16°	31.40°	13.88°	0.61°	17.72°	13.47°	0.61°

Table 6.4: Angular distance between the sensors mentioned and the sensors found by spherical sampling in the second experiment.

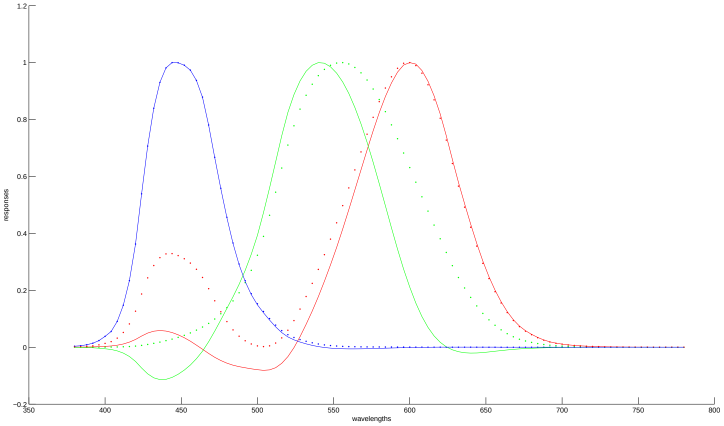


Figure 6.3: Original sensors -dotted lines- compared to spherical sampling sensors -solid lines- for the second dataset in the second experiment.

6.4 Experiment 3: Multispectral case

Sharp sensors have also been proved to be useful in modelling the interaction of light and surface without using all the spectra. In the work of Drew and Finlayson [30] they show that given a set of basis functions extracted from the spectra, by sharpening this basis we can model the interaction of light and surface with less complexity and also recover the spectra.

Let us suppose we have a set of reflectances $\{S_i\}_{i=1,\cdots,I}$, a set of illuminants $\{E_j\}_{j=1,\cdots,J}$, and the CIE XYZ functions that we denote as R. Then,

$$\underline{\rho}_{i,j} = \int_{\omega} S_i(\lambda)E_j(\lambda)R(\lambda)d\lambda \tag{6.7}$$

is the associated colour descriptor. Let us sample all the $\{E_j\}$ and $\{S_i\}$ at n wavelengths. Now we can construct a vector $\underline{w}$ of dimensions $(n \times 1)$

$$\underline{w} = diag(E_j) \cdot S_i \tag{6.8}$$

We compute this vector for every possible combination of lights and surfaces. We create the matrix W of dimension $n \times (I \cdot J)$.

$$W = [\underline{w}_1, \cdots, \underline{w}_{I \cdot J}] \tag{6.9}$$

Then, we perform a singular value decomposition over this matrix, $W = U\Sigma V^t$, where U is a $n \times n$ matrix, Σ a $n \times (I \cdot J)$ and Σ a $(I \cdot J) \times (I \cdot J)$.

From this decomposition we can extract a basis for the values on W. This basis will be formed by the first p columns of the matrix U. Then, $Q = [U_1, \cdots, U_p]$ has dimension $n \times p$. In our case, we use $p = 6$. Once we have these sensor basis Q, we compute the sharp sensor basis $\hat{Q}$. Here is where the different sharpening methods come into effect.

The goal is to check if the new sharp basis is good enough to recover the RGB values $\underline{\rho}_{i,j}$.

Let us now define an illuminant $\underline{E}$ of dimension $n \times 1$ and a spectral function $\underline{S}$ of the same dimension. They will be represented in the new basis as $\underline{s}$ and $\underline{e}$, which will have dimension 6×1

$$\underline{s} = \hat{Q}^t \underline{S} \tag{6.10}$$

and

$$\underline{e} = \hat{Q}^t \underline{E} \tag{6.11}$$

Moreover, as $\hat{Q}$ is not orthogonal, we need to use the pseudoinverse

$$\underline{\hat{E}} = \hat{Q}^+ \underline{e} \tag{6.12}$$

and

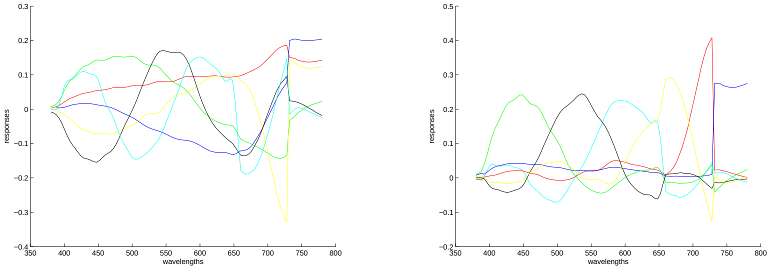
$$\underline{\hat{S}} = \hat{Q}^+ \underline{s} \tag{6.13}$$

Then, in this basis we can compute the vector v of dimension 6×1

$$\underline{v} = diag(\underline{e})\underline{s} \tag{6.14}$$

We can relate this vector v with an spectral composition V (dimension $n \times 1$)

$$\underline{\hat{V}} = \hat{Q}^+ \underline{v} \tag{6.15}$$



(a) (b)

Figure 6.4: (a) Basis functions for dataset 2 (b) Sharp functions for dataset 2.

Now, by considering the same sensitivities R used to compute equation 6.7, the RGB values can be computed by

$$\underline{b} = R^t \underline{V} = (R^t \hat{Q}^+)\underline{v} \tag{6.16}$$

Then, we can find a matrix $B = [\underline{b}_1, \cdots, \underline{b}_{I \cdot J}]$ of recovered values. Then for a particular $\underline{\rho}$ and its correspondent approximated solution $\underline{b}$, the error measure is

$$\Delta_\epsilon = \mathrm{d}(Lab(\underline{\rho}), Lab(\underline{b})) \tag{6.17}$$

where d is the euclidean distance.

We have compared in this experiment the spherical sampling versus two different methods (one of them with two initialisations). These are sensor-based spectral sharpening $L_2 - L_1$ constrained [29] and measurement tensor extension to nD. This last method have two different initialisations (I) with the identity as initial matrix and (II) with the solution of spectral sharpening as initial matrix.

Particularly, for spherical sampling we have calculated 100.000.000 points over the $6D$ sphere and we have done five iterations searching sensors that were in approximately 3° from the previous solutions.

Results are expressed in the form of the mean of Δ_ϵ and are shown in Table 6.5. We have used the same datasets as in the previous experiments. Again, spherical sampling outperforms the other methods. Furthermore, in Table 6.6 we show the minimal and the maximal angular distance between the six sensors found by spherical sampling and the ones for the other methods. Finally,in Figure 6.4 we can see (a) the original basis (Q) and (b) the sharpen basis for the dataset 2.

6.5 Conclusions

Sensor sharpening methods have been proven of critical importance for handling problems related to sensor calculations. Their most common application is the one related to colour constancy and the Von Kries law which allows the reduction of the complexity of the illumination change problem from a 9-dimension problem to a 3-dimension one. For this reason, most of the research in sensor sharpening has been focused in finding sharp sensors to tackle

Method	Dataset 1	Dataset 2	Dataset 3
Spherical sampling	**11.51**	**2.14**	**2.09**
Database sharpening L2-L1	16.18	12.37	7.27
Measurement tensor (I)	12.28	2.32	2.40
Measurement tensor (II)	12.28	2.28	159.20

Table 6.5: Results of the third experiment in the three different datasets.

Method	Dataset 1		Dataset 2		Dataset 3	
	Min	Max	Min	Max	Min	Max
Database sharpening	44.57°	114.02°	45.84°	66.25°	46.78°	65.50°
Measurement tensor (I)	2.87°	13.40°	1.97°	14.18°	5.20°	10.61°
Measurement tensor(II)	2.87°	13.40°	1.95°	14.18°	47.04°	169.13°

Table 6.6: Angular distance between the sensors mentioned and the sensors found by spherical sampling in the third experiment.

this particular problem. However, here we explore other applications that can be dealt by using sharp sensors.

In this chapter we have firstly shown that, when the measure for evaluating these sensors is not defined as a least squares minimisation, usual sharp sensors do not give the desired solution. In this case, the use of spherical sampling becomes crucial in order to deal with these new measures, i.e., CIE Delta E, or colour ratio stability. We have shown how spherical sampling overcomes the state of the art in this problems.

Finally, we have extended the spherical sampling method to more than 3 dimensions. This has allowed us to show that this method is also useful for modelling the interaction of light and surface without considering the whole spectra.

Chapter 7

Conclusions and further work

The main goal of this thesis was to introduce colour naming known results in order to deal with computational colour constancy for natural images. In the process of this thesis, we have tackled the problem from two different points of view. Firstly, we have worked on the selection of the illuminant of the scene. Secondly, we have worked in the selection of a set of sharp sensors that adapts to human colour naming categorisation. Then, our conclusions will be also split in these two points of view. In this chapter, we also show future research directions, that vary from the combination of the different results of this thesis, to the extension of each of the contributions separately.

7.1 Conclusions

Computational colour constancy is an ill-posed problem. In this work we have advanced the field by inserting psychophysically learned colour categories as an important information for colour constancy. In the first part of the thesis, the one dealing with the illuminant selection, we have made the following contributions:

Definition of a framework to introduce colour categorisation for computational colour constancy
In particular, we have experimented with basic colour terms defined in [12]. Our approach performs similarly to state-of-the-art methods, showing that basic colour names are useful for solving colour constancy.

Definition of a new calibrated image dataset
Although the creation of the dataset was not a goal of this thesis itself, it has been developed as an integral part of one of the projects of this thesis. This new dataset has three main advantages over previous ones: i) It uses a calibrated camera, ii) the illuminant of the scene is recorded in a grey-ball (although this was already done by Ciurea and Funt [26]) , iii) the quality of images is higher than in previous image datasets.

Definition of a psychophysical experiment in order to check if physical constancy error measures adapt to the human perception
This experiment is based on a pairwise image comparison between the solutions of different computational methods. The results have shown that the image with the best physical error is only preferred by the observers in less than 50% of the cases.

Definition of a new perceptual constancy measure
Based on the previous result, we have defined a new colour constancy measure: the perceptual angular error. This measure copes with the preference of humans, and gives smaller error to those solutions closest to the solutions chosen by the observers.

On the second part of the thesis, dealing with the use of sensor sharpening techniques, the main contributions are:

Definition of a new basis of sharp sensors that predicts unique hues and colour names
We have shown that by sensor sharpening we are able to generalise the recovery of the reflection properties defined by the PO's biological model.

Definition of a new compact singularity index relating colour names and unique hues with achromaticness
Related to the previous contribution, we have defined a new compact singularity index to predict unique hues and colour names from our sharp sensors. This new compact singularity index has two main advantages: i) it is a compact index and ii) it is related to a well-known colour property, achromatricness.

Definition of some experiments to find the sharp sensors that best cope with some perceptual measures
We have performed different experiments in order to find: a) the best sensors minimising the CIE Delta E error and b) the best sensors minimising colour ratio stability.

Extension of the spherical sampling technique to n-D
We have extended the spherical sampling procedure to n dimensions in order to deal with the problem of multispectral data recovery without spectra .

7.2 Further work

There is still a considerable body of work to be done in order to allow computer vision applications to mirror HVS. Our work has opened several new research directions. The thesis result with the most immediate impact is the insertion of the narrow band sensors (defined in the second part of the thesis) into the category hypothesis. It connects both parts of the thesis strengthening the importance of both results. We expect further improvements in the CaC methods due to an improved performance of the diagonal model.

Another extension of this work can be derived from the insertion of spatial and contextual information to improve the category hypothesis. Such additional data, that has not been used yet, complements the colour information, since psychophysical results have shown that context is very important in colour perception [102].

Yet another way to continue with this research relies on the definition of individual sensors that can be adjusted to the colours of a scene. The idea underlying this research line is to find the best sensors for each image with the goal of better representing the most important information. This research will be related to the one already started in [114].

Appendix A

Appendix: Lovisolo and Da Silva sphere sampling

The method proposed by Lovisolo and Da Silva [77] is based on the next assumption: when we have a big number of points K, there is one set of vectors uniformly distributed on an n-dimensional hyper-sphere that defines a tiling on the hyper-sphere by identical hypercubes of dimension N-1.

To start with, remember that a point p over the n-sphere in spherical coordinates will be represented as $p = [1\ \omega_1 \cdots \omega_n]$. The equivalence to Cartesian coordinates is trivial. Then, defining the length of each hypercube has δ, it can be defined as an addition of small variation of this spherical coordinates $\Delta_{\omega_1}, \cdots, \Delta_{\omega_n}$ as follows

$$\begin{aligned} \delta &= \Delta_{\omega_1} \\ \delta &= \Delta_{\omega_j} \sum_{i=1}^{j-1} \sin \omega_j,\ j = 2 : n-1 \end{aligned} \tag{A.1}$$

It should be noted, that in this point given δ and a point $\omega_1, \cdots, \omega_n$ we can compute the different values for $\Delta_{\omega_1}, \cdots, \Delta_{\omega_n}$. On the other side, it is possible to compute the length δ assuming that the sum of the area for all the hypercubes shall be equivalent to the area of the sphere. From this we have

$$A_n = K\delta^{n-1} \tag{A.2}$$

where K is the number of hypercubes or points we want to sample, and A_n is the area of the sphere which can be computed as

$$\begin{aligned} A_n &= \frac{n\pi^{n/2}}{(n/2)!} \quad \text{for n even} \\ A_n &= \frac{n 2^n \pi^{(n-1)/2} \left(\frac{n-1}{2}\right)!}{n!} \quad \text{for n odd} \end{aligned} \tag{A.3}$$

From all this information, Lovisolo and Da Silva proposed the next algorithm

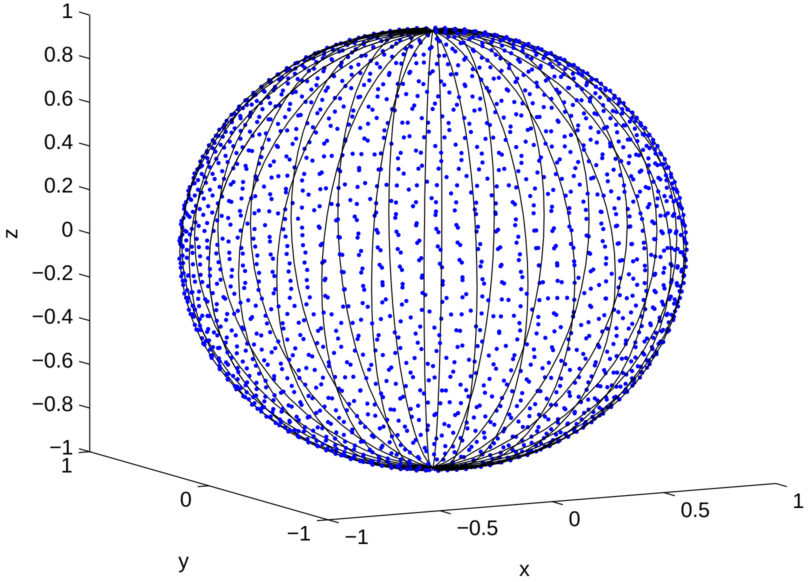


Figure A.1: Points sampled with the algorithm

```
compute $\Delta_{\omega_1}$ from equation A.1
for $\omega_1 = \frac{\Delta_{\omega_1}}{2}$ to $\pi$ in increments of $\Delta_{\omega_1}$ do
    compute $\Delta_{\omega_2}$ from equation A.1
    for $\omega_2 = \frac{\Delta_{\omega_2}}{2}$ to $\pi$ in increments of $\Delta_{\omega_2}$ do
        ...
        compute $\Delta_{\omega_{n-2}}$ from equation A.1
        for $\omega_{n-2} = \frac{\Delta_{\omega_{n-2}}}{2}$ to $\pi$ in increments of $\Delta_{\omega_{n-2}}$ do
            compute $\Delta_{\omega_{n-1}}$ from equation A.1
            for $\omega_{n-1} = \frac{\Delta_{\omega_{n-1}}}{2}$ to $2\pi$ in increments of $\Delta_{\omega_{n-1}}$ do
                for $i = 1 ... n-1$ do
                    compute $x_i = \sin \omega_{w_1} \cdots \sin \omega_{w_{i-1}} \cos \omega_{w_i}$
                end for
                compute $x_n = \sin \omega_{w_1} \cdots \sin \omega_{w_{n-2}} \sin \omega_{w_{n-1}}$
            end for
        end for
    end for
end for
```

An example of points sampled with this algorithm: 2000 points in $3D$ can be seen in figure A.

Bibliography

[1] http://www.icsi.berkeley.edu/wcs/data.html.

[2] A. Agresti. Analysis of ordinal paired comparison data. *Journal of the Royal Statistics Society, Serie C, Applied Statistics*, (41):287, 1992.

[3] A. Agresti. *An Introduction to Categorical Data Analysis.* Wiley, Hoboken, NJ, 2 edition, 2007.

[4] J. L. Barbur and K. Spang. Colour constancy and conscious perception of changes of illuminant. *Neuropsychologia*, 46(3):853 – 863, 2008. Consciousness and Perception: Insights and Hindsights - A Festschrift in Honour of Larry Weiskrantz.

[5] K. Barnard, V. C. Cardei, and B. V. Funt. A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data. *IEEE Transactions on Image Processing*, 11(9):972–984, 2002.

[6] K. Barnard, F. Ciurea, and B. Funt. Sensor sharpening for computational color constancy. *Journal of the Optical Society of America A*, 18(11):2728–2743, 2001.

[7] K. Barnard, L. Martin, A. Coath, and B. V. Funt. A comparison of computational color constancy algorithms. II. Experiments with image data. *IEEE Transactions on Image Processing*, 11(9):985–996, 2002.

[8] K. Barnard, L. Martin, and B. Funt. Colour by correlation in a three dimensional colour space. In *European Conference on Computer Vision*, pages 275–289, 2000.

[9] K. Barnard, L. Martin, B. Funt, and A. Coath. A data set for colour research. *Color Research and Application*, 27:147–151, 2002.

[10] R. Benavente, M Vanrell, and R. Baldrich. A data set for fuzzy colour naming. *Color Research and Application*, 31(1):48–56, Feb 2006.

[11] R. Benavente, M. Vanrell, and R. Baldrich. Parametric fuzzy sets for automatic colour naming. *Journal of the Optical Society of America A*, 25(10):2582–2593, 2008.

[12] B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution*. University of California Press, Berkeley, CA, 1969.

[13] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini. Improving color constancy using indoor-outdoor image classification. *IEEE Transactions on Image Processing*, 17(12):2381–2392, December 2008.

[14] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini. Automatic color constancy algorithm selection and combination. *Pattern Recognition*, 43:695–705, March 2010.

[15] S. Bianco, F. Gasparini, and R. Schettini. A consensus based framework for illuminant chromaticity estimation. *Journal of Electronic Imaging*, 17:023013–1–9, 2008.

[16] K. T. Blackwell and G. Buchsbaum. Quantitative studies of color constancy. *Journal of the Optical Society of America A*, (5):1772–1780.

[17] R. M. Boynton and C. X. Olson. Locating basic colors in the osa space. *Color Research and Application*, 12(2):94–105, 1987.

[18] R. A. Bradley and M. E. Terry. The rank analysis of incomplete block designs — I. The method of paired comparisons. *Biometrika*, 39:324–345, 1952.

[19] D. H. Brainard and W. T Freeman. Bayesian color constancy. *Journal of the Optical Society of America A*, 14:1393–1411, 1997.

[20] G. Brelstaff, C.A. Parraga, T. Troscianko, and D. Carr. Hyperspectral camera system: acquisition and analysis. In *In Proc. of SPIE Conference on Geographic Information Systems, Photogrammetry, and Geological/Geophysical Remote Sensing*, pages 150–159, 1995.

[21] G. Buchsbaum. A spatial processor model for object colour perception. *Journal Franklin Institute*, 310:1–26, 1980.

[22] V. C. Cardei, B. Funt, and K. Barnard. Estimating the scene illumination chromaticity by using a neural network. *Journal of the Optical Society of America A*, 19(12):2374–2386, 2002.

[23] V. C. Cardei and B. V. Funt. Committee-based color constancy. In *Color Imaging Conference*, pages 311–313, 1999.

[24] C. Chiao, T. W. Cronin, and D. Osorio. Color signals in natural scenes: characteristics of reflectance spectra and effects of natural illuminants. *Journal of the Optical Society of America A*, 17(2):218–224, 2000.

[25] H. Chong, S. Gortler, and T. Zickler. The von kries hypothesis and a basis for color constancy. In *International Conference on Computer Vision*, 2007.

[26] F. Ciurea and B. Funt. A large image database for color constancy research. In *Color Imaging Conference*, pages 160–164, 2003.

[27] P. Courcoux and M. Semenou. Preference data analysis using a paired comparison model. *Food Quality and Preference*, 8(5-6):353 – 358, 1997. Third Sensometrics Meeting.

[28] A. M. Derrington, J. Krauskopf, and P. Lennie. Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, 357:241–265, 1984.

[29] M. S. Drew and G. D. Finlayson. Spectral sharpening with positivity. *Journal of the Optical Society of America A*, 17(8):1361–1370, 2000.

[30] M. S. Drew and G. D. Finlayson. Multispectral processing without spectra. *Journal of the Optical Society of America A*, 20(7):1181–1193, 2003.

[31] M. D. Fairchild. *Color Appearance Models*. Wiley-IST, 2nd edition, 2005.

[32] G. D. Finlayson. Color in perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):1034–1038, 1996.

[33] G. D. Finlayson, M. S. Drew, and B. V. Funt. Color constancy: Enhancing von kries adaptation via sensor transformation. In *Human Vision, Visual Processing, and Digital Display IV*, pages 473–484, 1993.

[34] G. D. Finlayson, M. S. Drew, and B. V. Funt. Color constancy: generalized diagonal transforms suffice. *Journal of the Optical Society of America A*, 11(11):3011–3019, 1994.

[35] G. D. Finlayson, M. S. Drew, and B. V. Funt. Spectral sharpening: sensor transformations for improved color constancy. *Journal of the Optical Society of America A*, 11(5):1553–1563, 1994.

[36] G .D. Finlayson, S. D. Hordley, and P. M. Hubel. Color by correlation: A simple, unifying framework for color constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, November 2001.

[37] G. D. Finlayson, S. D. Hordley, and I. Tastl. Gamut constrained illuminant estimation. *Internation Journal of Computer Vision*, 67(1):93–109, 2006.

[38] G. D. Finlayson, S. D. Hordley, and R. Xu. Convex programming colour constancy with a diagonal-offset model. In *Internation Conference on Image Processing*, pages 948–951, 2005.

[39] G. D. Finlayson and S. Susstrunk. Spherical Sampling and Color Transformations. In *Proc. IS&T/SID 9th Color Imaging Conference*, volume 9, pages 321–325, 2001.

[40] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *Color Imaging Conference*, pages 37–41, 2004.

[41] J. Fleckenstein, R.A. Freund, and J.E. Jackson. A paired comparison test of typewriter carbon papers. *Tappi Journal*, (41):128, 1958.

[42] D. A. Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–35, 1990.

[43] D. H. Foster. Does colour constancy exist? *Trends in Cognitive Science*, 7(10), 2003.

[44] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster. Frequency of metamerism in natural scenes. *Journal of the Optical Society of America A*, 23(10):2359–2372, 2006.

[45] D. H. Foster and S. M. C. Nascimento. Relational colour constancy from invariant cone-excitation ratios. *Proceedings: Biological Sciences*, 257(1346):115–121, 1994.

[46] D. H. Foster, S. M. C. Nascimento, B. J. Craven, K. J. Linnell, F. W. Cornelissen, and E. Bremer. Four issues concerning color constancy and relational color constancy. *Vision Research*, (37):1341–1345.

[47] D. H. Foster, S M.C. Nascimento, and K. Amano. Information limits on neural identification of colored surfaces in natural scenes. *Visual Neuroscience*, 21(03):331–336, 2004.

[48] B. Funt, K. Barnard, and L. Martin. Is machine colour constancy good enough. In *In Proceedings of the 5th European Conference on Computer Vision*, pages 445–459. Springer, 1998.

[49] B. V. Funt, M. S. Drew, and J. Ho. Color constancy from mutual reflection. *International Journal of Computer Vision*, 6(1):5–24, 1991.

[50] G. Gabrielsen. Paired comparisons and designed experiments. *Food Quality and Preference*, 11(1-2):55 – 61, 2000.

[51] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp. Bayesian color constancy revisited. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1–8, 06 2008.

[52] A. Gijsenij and T. Gevers. Color constancy using natural image statistics and scene semantics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (in press), 2011.

[53] A. Gijsenij, Th. Gevers, and J. van de Weijer. Generalized gamut mapping using image derivative structures for color constancy. *International Journal of Computer Vision*, 86(2–3):127–139, 2010.

[54] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins studies in the Mathematical Sciences, 3rd edition, 1996.

[55] H. Grassman. Zur theorie der farbenmischung. *Annuary of Physics and Chemistry*, 89:69–84, 1853.

[56] J. Guild. The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London. Series A*, 230:149–187, 1931.

[57] T. Hansen, M. Olkkonen, S. Walter, and K. R. Gegenfurtner. Memory modulates color appearance. *Nature Neuroscience*, 9(11):1367–1368, October 2006.

[58] T. Hansen, S. Walter, and K. R. Gegenfurtner. Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision*, 7(4), 2007.

[59] R. A. Harshman. Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multi-modal factor analysis. *UCLA Working Papers in Phonetics*, 16(1):84, 1970.

[60] M. Hedrich, M. Bloj, and A. I. Ruppertsberg. Color constancy improves for real 3D objects. *Journal of Vision*, 9(4), 2009.

[61] E. Hering. *Zur Lehre vom Lichtsinne*. Sechs Mittheilungen an die kaiserliche Akademie der Wissenschaft in Wien, 1891.

[62] T. Hofmann. Probabilistic latent semantic indexing. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '99, pages 50–57, New York, NY, USA, 1999. ACM.

[63] S. D. Hordley and G. D. Finlayson. Re-evaluating colour constancy algorithms. In *International Conference on Pattern Recognition*, pages 76–79, 2004.

[64] A. Hurlbert. Perceptual constancy. *Edited by Walsh V., Kulikowski J, Cambridge UK: Cambridge University Press.*, pages 283–322, 1998.

[65] A. Hurlbert. Colour vision: Is colour constancy real? *Current Biology*, 9(15):558–561(4), 1999.

[66] A. Hurlbert. Colour constancy. *Current Biology*, 21(17):906–907, 2007.

[67] A. Hurlbert and K. Wolf. Color contrast: a contributory mechanism to color constancy. *Progress on Brain Research*, 144, 2004.

[68] D. B. Judd, D. L. Macadam, G. Wyszecki, H. W. Budde, H. R. Condit, S. T. Henderson, and J. L. Simonds. Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America A*, 54(8):1031–1036, 1964.

[69] P. Kay. Color categories are not arbitrary. *Cross Cultural Research*, 39(1):39–55, 2005.

[70] G. Klinker, S. Shafer, and T. Kanade. A physical approach to color image understanding. *International Journal of Computer Vision*, 4(1):7–38, January 1990.

[71] J. M. Kraft and D. H. Brainard. Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences USA*, 96:307–312, 1999.

[72] E. L. Krinov. Spectral reflectance properties of natural formations. *National Research Council of Canada*, 1947.

[73] R. G Kuehni. Variability in unique hue selection: A surprising phenomenon. *Color Research and Application*, 29:158–162, 2004.

[74] E. H. Land. The retinex. *American Scientist*, 52:247–264, 1964.

[75] E. H. Land and J. J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America A*, 61(1):1–11, 1971.

[76] H. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A*, 3:694–1699, 1986.

[77] L. Lovisolo and E.A.B. da Silva. Uniform distribution of points on a hyper-sphere with applications to vector bit-plane encoding. *IEE Proceedings - Vision, Image, and Signal Processing*, 148(3):187–193, 2001.

[78] M. Luo, A. Clarke, P. Rhodes, A. Schappo, S. Scrivener, and C. Tait. Quantifying color appearance I. LUTCHI color appearance data. *Color Research and Application*, (16):166, 1991.

[79] M. Luo, A. Clarke, P. Rhodes, A. Schappo, S. Scrivener, and C. Tait. Quantifying color appearance II. Testing color models performance using LUTCHI color appearance data. *Color Research and Application*, (16):181, 1991.

[80] D. H. Marimont and B. A. Wandell. Linear models of surface and illuminant spectra. *Journal of the Optical Society of America A*, 9(11):1905–1913, 1992.

[81] J. J. McCann. Some additional comments on CIECAM97. *Color Research and Application*, 24:215–217, 1999.

[82] N. Moroney. Thousands of on-line observers is just the beginning. In *Human Vision and Electronic Imaging*, volume 7249 of *Proceedings of the SPIE*, pages 724005–724005(8), 2009.

[83] D. Mylonas and L. MacDonald. Online colour naming experiment using munsell colour samples. In *European Conference on Colour in Graphics, Imaging and Vision (CGIV'10)*, pages 27–32, Joensuu(Finland)), Jun 2010. IS&T.

[84] S. M. C. Nascimento, F. Ferreira, and D. H. Foster. Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A*, 19:1484–1490, 2002.

[85] S. M .C. Nascimento and D. H. Foster. Relational color constancy in achromatic and isoluminant surfaces. *Journal of the Optical Society of America A*, (17):225–231.

[86] I. Newton. *Opticks*. 1704.

[87] M. Olkkonen, T. Hansen, and K. R. Gegenfurtner. Color appearance of familiar objects: Effects of object shape, texture, and illumination changes. *Journal of Vision*, 8(5), 2008.

[88] M. Olkkonen, T. Hansen, and K. R. Gegenfurtner. Categorical color constancy for simulated surfaces. *Journal of Vision*, 9(12):1–18, 11 2009.

[89] A. Olmos and F.A.A. Kingdom. A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, (33):1463, 2004.

[90] J. P. S. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of munsell colors. *Journal of the Optical Society of America A*, 6(2):318–322, 1989.

[91] C. A. Parraga, T. Troscianko, and D. J. Tolhurst. Spatiochromatic properties of natural images and human vision. *Current Biology*, (12):483, 2002.

[92] D. L. Philipona and J. K. O'Regan. Color naming, unique hues and hue cancellation predicted from singularities in reflection properties. *Visual Neuroscience*, 3-4(23):331–339, 2006.

[93] P. D. Pinto, J. M. M.l Linhares, and S. M. C. Nascimento. Correlated color temperature preferred by observers for illumination of artistic paintings. *Journal of the Optical Society of America A*, 25(3):623–630, 2008.

[94] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou. Minimal discrete energy on the sphere. *Mathematical Research Letters*, (1):647–662, 1994.

[95] P Rakic. Specification of cerebral cortical areas. *Science*, 241(4862):170–176, 1988.

[96] D. C. Rich. Publication cie 159: A colour appearance model for colour management systems: CIECAM02. *Color Research and Application*, 31:158, 2006.

[97] C. Riess, E. Eibenberger, and E. Angelopoulou. Illuminant Estimation by Voting. Technical report, 2009.

[98] A. R. Robertson. The CIE 1976 color-difference formulae. *Color Research and Application*, 2:7–11, 1977.

[99] J. Romero, A. García-Beltrán, and J. Hernández-Andrés. Linear bases for representation of natural and artificial illuminants. *Journal of the Optical Society of America A*, 14(5):1007–1014, 1997.

[100] G. Sapiro. Color and illuminant voting. *IEEE Transactions on Image Processing*, 21(11):1210–1215, November 1999.

[101] S. A Shafer. Using color to separate reflection components. *Color Research and Application*, 10(4):210–218, 1985.

[102] S. K. Shevell and F. A. A Kingdom. Color in complex scenes. *Annual Review of Psychology*, 59:143–146, 2008.

[103] H. E. Smithson. Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions: Biological Sciences*, 360(1458), 2005.

[104] W. S. Stiles and J.M. Burch. NPL colour-matching investigation: Final report. *Optica Acta*, (6):1–26, 1959.

[105] A. Stockman and L.T. Sharpe. Spectral sensitivities of the middleand long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40:1711–1737, 2000.

[106] J Sturges and T. W. A Whitfield. Locating basic colours in the munsell space. *Color Research and Application*, 20:364–376, 1995.

[107] M.G.A. Thomson, S. Westland, and J. Shaw. Spatial resolution and metamerism in coloured natural scenes. *Perception*, (29):123, 2000.

[108] L. L. Thurstone. A law of comparative judgement. *Psychological Review*, 34:278–286, 1927.

[109] F. Tous. *Computational framework for the white point interpretation based on colour matching*. PhD thesis, Computer Vision Center, 2006.

[110] A. Valberg. Unique hues: An old problem for a new generation. *Vision Research*, 41:1645–1657, 2001.

[111] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *IEEE Transactions on Image Processing*, 16(9):2207–2214, 2007.

[112] J. van de Weijer, C. Schmid, and J. Verbeek. Using high-level visual information for color constancy. In *International Conference on Computer Vision*, oct 2007.

[113] M. T. Vanleeuwen, C. Joselevitch, I. Fahrenfort, and M. Kamermans. The contribution of the outer retina to color constancy: a general model for color constancy synthesized from primate and fish data. *Visual Neuroscience*, 24(3).

[114] J. Vazquez-Corral and M. Vanrell. Content-based colour space. Technical report, Computer Vision Center, 2007.

[115] J. Vazquez-Corral, M. Vanrell, R. Baldrich, and C. A. Párraga. Towards a psychophysical evaluation of colour constancy algorithms, Jun 2008.

[116] H. von Helmholtz. *Handbuch der Physiologischen Optik*. Voss, 1866.

[117] R. Vrhel, M.J. anf Gershon and L.S. Iwan. Measurement and analysis of object reflectance spectra. *Color Research and Application*, 19:4–9, 1994.

[118] A. Werner. The influence of depth segmentation on colour constancy. *Perception*, 35(9):1171–84, 2006.

[119] A. Werner. Color constancy improves, when an object moves: High-level motion influences color perception. *Journal of Vision*, 7(14), 2007.

[120] J. A. Worthey and M. H. Brill. Heuristic analysis of von kries color constancy. *Journal of the Optical Society of America A*, 3(10):1708–1712, 1986.

[121] W. D. Wright. A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30(4):141, 1929.

[122] G. Wyszecki and W.S. Stiles. *Color science: concepts and methods, quantitative data and formulae*. John Wiley & Sons, 2nd edition, 1982.

[123] S. N. Yendrikhovskij. Computing color categories from statistics of natural images. *Journal of Imaging Science and Technology*, 45:409–417, 2001.

Final acknowledgement

This work has been partially supported by projects TIN2007-64577, TIN2010-21771-C02-1 and Consolider-Ingenio 2010 CSD2007-00018 of Spanish MEC (Ministery of Science)

List of Publications

This dissertation has led to the following communications:

Journal Papers

- Javier Vazquez-Corral, Carlos A. Párraga , Maria Vanrell, Ramon Baldrich. Colour Constancy Algorithms: psychophysical evaluation on a new dataset. *Journal of Imaging Science and Technology*, 53(3), 031105, May-June, 2006

Published abstracts

- Javier Vazquez-Corral, Carlos A. Párraga , Maria Vanrell. Ordinal pairwise method for natural images comparison, *Perception (Supplements)*, 2009
- Carlos A. Párraga, Javier Vazquez-Corral, Maria Vanrell. A new cone activation-based natural images dataset. *Perception (Supplements)*, 2009

Conference Contributions

- Javier Vazquez-Corral, Graham D. Finlayson, Maria Vanrell. A new compact similarity measure to predict World Colour Survey data and unique hues. In *Proceedings of the 5th European Conference on Colour in Graphics, Imaging and Vision (CGIV'10)*, 2010.
- Javier Vazquez-Corral, Maria Vanrell, Ramon Baldrich, Carlos A. Párraga. Towards a psychophysical evaluation of Colour Constancy Algorithms. In *Proceedings of the 4th European Conference on Colour in Graphics, Imaging and Vision (CGIV'08)*, 2008.
- Eduard Vazquez, Ramon Baldrich, Javier Vazquez-Corral, Maria Vanrell. Topological histogram reduction towards colour segmentation. In *Proceedings of the Third Iberian Conference on Pattern Recognition and Image Analysis, IBPRIA07*, 2007
- Javier Vazquez-Corral, Maria Vanrell, Anna Salvatella, Eduard Vazquez. A colour space based on the image content. In *Proceedings of the Tenth Catalan Conference on Artificial Intelligence , CCIA07*, 2007

Internal Workshops & Technical Reports

- Javier Vazquez-Corral, Maria Vanrell. Color: Representation, Constancy and Psychophysics. In *Proceedings of the Fourth CVC Internal Workshop on Computer Vision:*

Advances in Research and Development, CVCRD09, pages 64–69, 2009.

- Javier Vazquez-Corral, Maria Vanrell. A Colour space that adapts to the coloured blob detection. In *Proceedings of the Second CVC Workshop, Computer Vision: Advances in Research & Development, CVCRD07*, pages 7–12, 2007.
- Javier Vazquez-Corral, Maria Vanrell. Content-based Colour Space. *CVC Technical Report* Num. 94, 2007.