Extending Minkowski Norm Illuminant Estimation

Perla Aurora Troncoso Rey

A thesis submitted for the Degree of
Doctor of Philosophy

University of East Anglia
School of Computing Sciences

October 2012

Abstract

The ability to obtain colour images invariant to changes of illumination is called colour constancy. An algorithm for colour constancy takes sensor responses - digital images - as input, estimates the ambient light and returns a corrected image in which the illuminant influence over the colours has been removed. In this thesis we investigate the step of illuminant estimation for colour constancy and aim to extend the state of the art in this field.

We first revisit the Minkowski Family Norm framework for illuminant estimation. Because, of all the simple statistical approaches, it is the most general formulation and, crucially, delivers the best results. This thesis makes four technical contributions. First, we reformulate the Minkowski approach to provide better estimation when a constraint on illumination is employed. Second, we show how the method can (by orders of magnitude) be implemented to run much faster than previous algorithms. Third, we show how a simple edge based variant delivers improved estimation compared with the state of the art across many datasets. In contradistinction to the prior state of the art our definition of edges is fixed (a simple combination of first and second derivatives) i.e. we do not tune our algorithm to particular image datasets. This performance is further improved by incorporating a gamut constraint on surface colour -our 4th contribution.

The thesis finishes by considering our approach in the context of a recent OSA competition run to benchmark computational algorithms operating on physiologically relevant cone based input data. Here we find that Constrained Minkowski Norms oper-

ating on spectrally sharpened cone sensors (linear combinations of the cones that behave more like camera sensors) supports competition leading illuminant estimation.

Acknowledgements

This thesis would have not been possible without the guidance, advice and support of my supervisor, Prof. Graham D. Finlayson. Thank you for believing in me and in this project. Now, at the end of this adventure and while writing this final section, my cup overflows. My examiners, Dr. Maria Vanrell and Dr. Mark Fisher, have provided an interesting and gratifying opportunity to discuss this work. I would like to thank them both for their opportune comments and recommendations.

I would also like to thank my sponsors CONACYT and Fundación GCC for their sponsorship and making possible for me to pursue my PhD at UEA.

My colleagues at the Colour Lab, Dome, Roberto, Mike, Stuart and Michal, have contributed to have a happy and interesting work environment and I would like to thank you guys for that. Special thanks to Dr. Roberto Montagna for proof reading this thesis and for his helpful advice and comments.

My dear friends Wei-Wei, Alba, Mino, Gaby and Rowena, I want to thank you guys for all what we have shared, this experience would have not been the same without you all. Stephen, your support and encouragement is a pillar of strength in overcoming all difficult times. Thank you for your cherished company, you are a true inspiration in my life.

Finally, my dear family, thank you Susy, Juan, Laura, Jimmy, Emilio, dad and Irma for always being there for me. Mum, thanks for your love and advice, without you I would not be where I am today.

Contents

List of Figures

List of Tables

Publications

The following are publications by the author related to this work:

- Graham D. Finlayson, Perla A. Troncoso Rey and Elisabetta Trezzi. General ℓ^p Constrained Approach for Colour Constancy. In Proceedings of IEEE Color and Photometry in Computer Vision Workshop, International Conference of Computer Vision, ICCV, Barcelona, Spain, November 2011.

- Hamid Reza Vaezi, Mark S. Drew, Graham D. Finlayson and Perla A. Troncoso Rey. The Role of Bright Pixels in Illumination Estimation. In Proceedings of the IS& T_SID, Twentieth Color and Imaging Conference, CIC, November 2012.

Glossary

CD	Combined Derivative
CDC	Combined Derivative Constrained
DCT	Discrete Cosine Transform
GCDC	Gamut Combined Derivative Constrained
SoG	Shades of Gray
SoGC	Shades of Gray Constrained
HDR	high dynamic range

Chapter 1

Introduction

Colour is a visual feature which like shape, texture and size, is used by humans to describe objects. Of particular interest to my research is the ability of the human vision system to determine -to some extent- the colour of an object despite the ambient light. For example, take a white piece of paper viewed outside on a sunny day. The paper colour will remain white to a human observer even though it is viewed at different times of the day and the wavelength composition of the daylight changes. Formally we see illuminant independent *colour descriptors*: our visual system, by some algorithmic means, discounts the colour of the light. The ability to determine an object's colour based on its surface properties is called colour constancy. The degree to which human observers exhibits colour constancy depends in part on the viewing conditions. However, even with this caveat, the human visual system achieves reasonable colour constancy (in most cases).

In computer vision if we could "solve for" a colour descriptor (surface properties independent of the colour of the light) then this ability would be useful in a number of areas including the enhancement of digital photography, object recognition, aerial image analysis, image retrieval, and development of systems for automatic quality in-

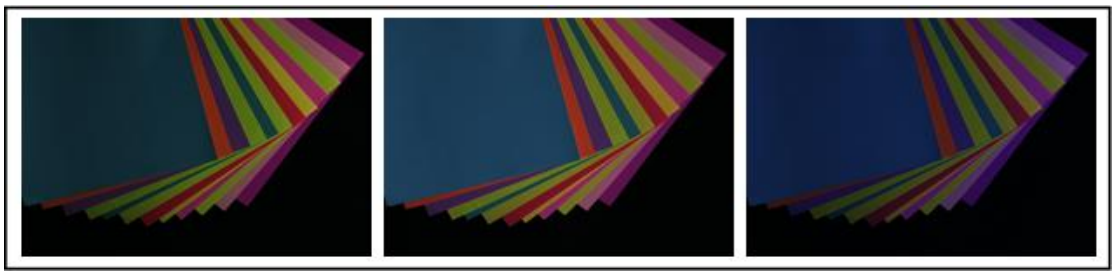

Figure 1.1: Three images showing the same scene viewed under different lights

spection [FBM98]. Yet, there is a gap between how we think about colour (a property intrinsic to the reflectance of an object) and the physics of image formation (the picture recorded by a camera). Camera images are confounded by the colour of the prevailing illuminant. For example, Figure 1.1 shows images of the same scene comprising coloured pieces of paper under different illuminants. Although the pieces of paper in the scene are the same in the three pictures, they look different due to changes in the illumination.

A number of algorithms have been proposed to solve colour constancy in computer vision (often in analogy to what is known about human visual processing). The key part of all methods in the literature is estimating (or inferring) the colour of the illuminant from an image (since, if this is known, it is simple to remove colour bias due to illumination [HHFD97]). Over and above solving the problem for its own sake, colour constancy in computer vision has mainly been considered as a preprocessing step for allowing colour to be used as an aid in tasks such as object recognition.

For example, object recognition techniques aim to automatically locate and recognise objects in digital images. Objects of interest are described by a set of features which make it possible to differentiate them from other elements in the image. It is desirable that these features are invariant to effects such as translation, rotation, scale and, of primary relevance to this thesis, to a change of illumination [HS94] [JZQ11].

Furthermore, digital photography is another practical application of computational

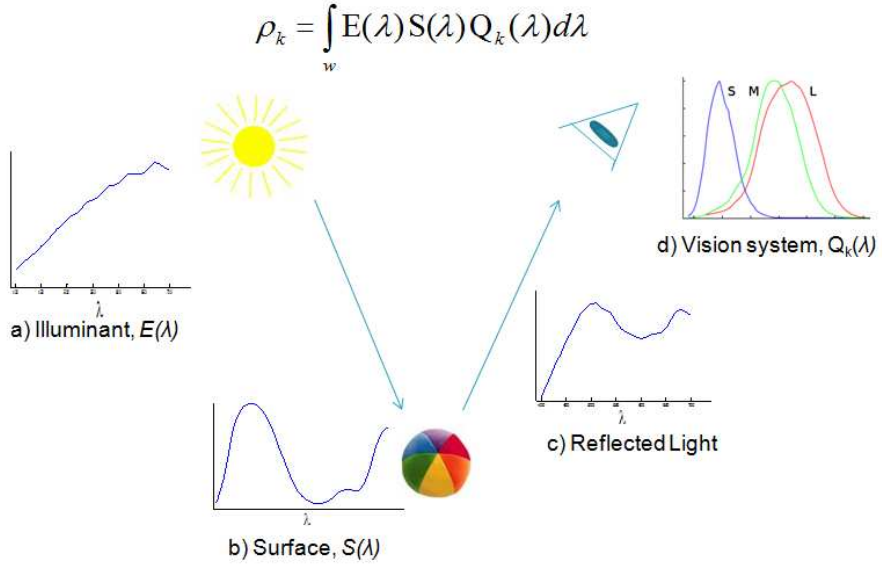

Figure 1.2: The simplified model of image formation. Figure (a) shows an ambient light and its spectral power distribution. Figure (b) shows an object and its intrinsic surface property represented by surface reflectance in Figure (c). Figure (d) represents the visual system and its sensitivity functions.

colour constancy. Specifically in tasks such as colour balancing where the ambient illumination is estimated from digital camera data and then used to "correct" the image colours according to the type of illumination.

To tackle these problems, the first step is to understand how image formation works, since it provides the basis to have a better understanding and insight to the problem. Any response from a camera (or the human vision system) to a scene is determined by three different factors: the ambient light illuminating the scene, the reflectance property of the object (or objects) in the scene, and the camera's sensor sensitivities. Now, this process is called *Image Formation* and is illustrated in Figure 1.2.

The simplified model of image formation is defined by:

$$\rho_k = \int_w E(\lambda)S(\lambda)Q_k(\lambda)d\lambda, \tag{1.1}$$

where ρ_k is the response to a surface colour over the wavelength range w (roughly 400 to 700 Nanometres). This response is produced when a surface reflectance, $S(\lambda)$, (where λ is the wavelength) is illuminated by a spectral power distribution, $E(\lambda)$. The reflected light from the surface (sometimes called the colour signal) is then perceived by the camera's sensor, Q_k, resulting in the perception of colour.

According to the simplified model of image formation, a camera's response to a given surface is defined by the product of the ambient light illuminating the object, the surface reflectance properties of the object, and the camera's sensitivities. So let us imagine, for an object with surface reflectance $S(\lambda)$, two situations where the object is illuminated by two different lights, E^1 and E^2. If we were to measure the amount of light reflected by the same object but under different lights, these reflected light will be the result of $E^1(\lambda)S(\lambda)$ and $E^2(\lambda)S(\lambda)$. The camera's response to this object under two lighting conditions will then be represented by $\boldsymbol{\rho}^1$ and $\boldsymbol{\rho}^2$. These two camera's responses are different. However, they simply represent the same object under a different light. In computer vision, the goal of colour constancy is to "see" the objects under all lights as having the same colour. If we substitute the camera sensors with the human cones the equations remain the same. Sometimes the goal of colour constancy in computer vision is to "see like we do". The colour constancy for humans is illustrated in Figure 1.3.

1.1 Discrete Image Formation

In practice we use a discrete version of the simplified model of image formation. Normally, the spectra are sampled evenly at fixed intervals, so that we can denote each of the functions $S(\lambda)$, $E(\lambda)$ and $Q_k(\lambda)$ with a n-dimensional vector:

$$\mathbf{E} = [E(\lambda_1)\ E(\lambda_2)\ ...\ E(\lambda_n)]^t \tag{1.2}$$

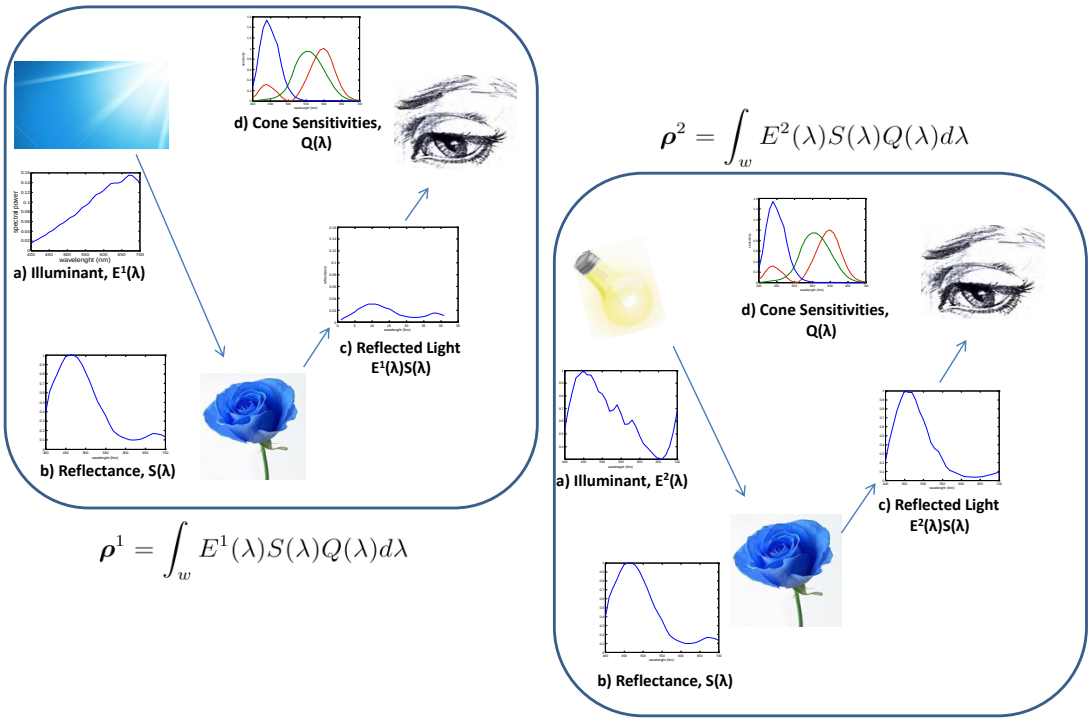

Figure 1.3: Human Colour Constancy: The human visual system is able to perceive "illuminant-invariant" colour descriptors (e.g. surface colour independent of the prevailing illumination). The figure shows the cone's response to the same object under two different lights, E^1 and E^2. The reflected light from the object is the product of $E^1(\lambda)S(\lambda)$ and $E^2(\lambda)S(\lambda)$ respectively, resulting in the cone's responses, $\boldsymbol{\rho}^1$ and $\boldsymbol{\rho}^2$. Even though $\boldsymbol{\rho}^1$ and $\boldsymbol{\rho}^2$ are different, they represent the same object under two different lights.

$$\mathbf{S} = [S(\lambda_1)\ S(\lambda_2)\ ...\ S(\lambda_n)]^t \tag{1.3}$$

$$\mathbf{Q}_k = [Q_k(\lambda_1)\ Q_k(\lambda_2)\ ...\ Q_k(\lambda_n)]^t \tag{1.4}$$

By adopting this discrete representation, Equation 1.1 can be written as the following summation:

$$\rho_k = \sum_{i=1}^{n} E(\lambda_i)S(\lambda_i)Q_k(\lambda_i)\triangle\lambda \tag{1.5}$$

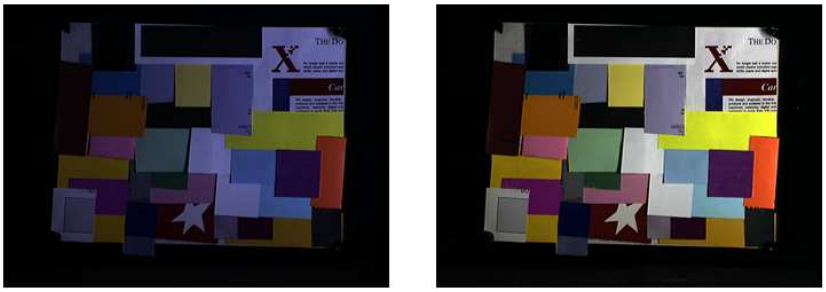

Figure 1.4: Two images viewed under two different lighting conditions

If we further assume that the $\triangle\lambda$ is incorporated into the sensor response functions, the simplified image formation equation, Equation 1.1, can be rewritten as:

$$\rho_k = \sum_{i=1}^{n} E(\lambda_i)S(\lambda_i)Q_k(\lambda_i) \tag{1.6}$$

Typically, we sample these functions with increments of 1, 2, 4, 5 or 10 Nanometres. The n used in Equation 1.5 and Equation 1.6 is between 31 and 301 (the visible range is 300 Nanometres wide). However, it has been found [SSS92] that with increments of 10 Nanometres almost all colorimetric integrals can be calculated according to Equation 1.6 without loss of accuracy.

1.2 The Computational Colour Constancy Problem

In computer vision, the colour constancy problem can be defined as the ability to estimate, from a digital image, colour descriptors that depend only on the surface reflectance, $S(\lambda)$, and not biased by the illuminant, $E(\lambda)$. In this context, the only available information for doing so is provided by the sensor response, $\boldsymbol{\rho}$.

Typical attempts to achieve computational colour constancy divide the task as a two

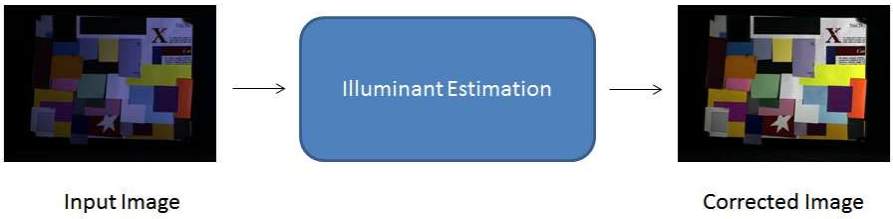

Figure 1.5: Colour constancy typically works in two stages. First, from a given image, it estimates the illuminant of the scene. Second, it removes the illuminant influence to obtain a corrected illuminant-invariant image.

step problem. First, we might attempt to calculate $E(\lambda)$ given only RGB pixel values in an image or, as an easier alternative, to estimate the 3-dimensional vector $e = [\hat{R}\ \hat{G}\ \hat{B}]^t$ corresponding to the RGB values of the prevailing light. This first step is known as *Illuminant Estimation*. Second, if the illuminant is estimated correctly then the colour cast due to the illuminant is readily removed from images [HHFD97]. This can be achieved simply by mapping each RGB value in the image by the estimate of the light (typically using a 3×3 matrix or a 3×3 diagonal matrix). When a diagonal matrix is used the diagonal terms are often the reciprocal of the R, G and B of the illuminant. Clearly dividing the light by itself gives $(1, 1, 1)$, i.e. a coloured illuminant maps to white. Remarkably "dividing out" the illuminant colour tends to work well for all surface colours for almost all typical cameras that have fairly narrow-band sensitivities and can be made to work even for systems that have broader response functions [FDF94b]. Figure 1.5 illustrates "typical" computational colour constancy.

While in the early years of colour constancy research the methods for removing illuminant colour bias achieved close attention, generally, if the illuminant is known, image colours can be made to "look correct". Thus, the primary focus of this thesis is to investigate and extend the state of the art in illuminant estimation.

1.2.1 Illuminant Estimation

Illuminant estimation algorithms are based on different assumptions about the expected state of the world. Depending on the kind of assumptions these algorithms adopt, they can, broadly, be divided into three categories: physics-based, statistical and those using a combination of the first two.

Physics-based algorithms work by assuming a known physical process is presented, e.g. a specular highlight [Lee92] [TW96] or a mutual interreflection [FDH91] [Tom96], and then apply this information to the physics of image formation to arrive at an illuminant estimation.

Statistical algorithms for illuminant estimation look at digital images as a distribution of measurements and use statistical tools to estimate the illuminant. The most well known algorithms in this category are based on the principles of mean and maximum value in a distribution, the **Grey World** [Buc80] and the **MaxRGB** [Lan77] respectively. An illuminant estimation using the **Grey World** algorithm is calculated by:

$$\mathbf{e} = [\mathrm{mean}(\mathbf{R})\ \ \mathrm{mean}(\mathbf{G})\ \ \mathrm{mean}(\mathbf{B})]^t, \tag{1.7}$$

where $\mathbf{e}$ represents the 3-dimensional vector for the illuminant estimation, and $\mathbf{R}$, $\mathbf{G}$ and $\mathbf{B}$, are the three colour channels in the image.

MaxRGB instead assumes that there is at least one white patch, or three surfaces maximally reflecting in each R, G, B channel, reflecting all the light that is illuminating it. An illuminant estimation, e, given by **MaxRGB** is calculated:

$$\mathbf{e} = [\max(\mathbf{R})\ \ \max(\mathbf{G})\ \ \max(\mathbf{B})]^t \tag{1.8}$$

Later, in 2004, the **Shades of Gray** algorithm was proposed by Finlayson and Trezzi [FT04], which is based on the Minkowski family norms. They showed that a summary of the statistics in the image taking the form of a Minkowski norm can be used to

estimate the illuminant.

A Minkowski norm is defined in a N dimensional real space. Let $\mathbf{X} = [X_1 ... X_N]^t$ be a vector in $\mathbb{R}^N$. For every $p \geq 1$ the quantity

$$\|\mathbf{X}\|_p = \left\{ \sum_{i=1}^{N} |X_i|^p \right\}^{1/p} \tag{1.9}$$

defines a norm in $\mathbb{R}^N$ [KF75] and X_i represents the ith component of $\mathbf{X}$.

As $p \to \infty$ the pth norm approaches the maximum component of $\mathbf{X}$ [FT04]. Re-writing this equation to normalise it by the number of points in the distribution we have:

$$\mu_p(\mathbf{X}) = \frac{\|\mathbf{X}\|_p}{N^{(1/p)}}. \tag{1.10}$$

This p norm can be then used as an illuminant estimator, which results in the formalism of the **Shades of Gray** algorithm, where an estimation of the illuminant is calculated by:

$$\mathbf{e} = [\mu_p(\mathbf{R})\, \mu_p(\mathbf{G})\, \mu_p(\mathbf{B})]^t, \tag{1.11}$$

Furthermore, the **Grey World** and the **MaxRGB** algorithms are represented by the extreme norms of a Minkowski Family [FT04]. Typically (and remarkably) a norm between four and six was empirically shown to delivers good performance for different image data sets [FT04].

It has been shown that a p-norm summary based on image differences (derivatives of different order) can deliver improved performance [JGG07]. This is the principle of the **Grey Edge** algorithm, where an estimation of the illuminant, e, is given by:

$$e_k^{n,p,\sigma} = \left(\int \left\| \frac{\partial^n I_k^\sigma(\boldsymbol{x})}{\partial \boldsymbol{x}^n} \right\|^p dx \right)^{1/p} \quad k \in R, G, B \tag{1.12}$$

where k denotes R, G or B, and n is the order of the derivative ($n = 0$ means no derivative is taken, $n = 1$ indicates first derivative and $n = 2$ second derivative). The term $I_k^\sigma(x)$ is the result of a convolution of the image I with a Gaussian filter, G with σ pixel standard deviation. This convolution is given by:

$$I_k^\sigma(x) = I_k \otimes G^\sigma \tag{1.13}$$

where G represents a Gaussian filter.

An extension to this work allows the use of the discrete cosine decomposition of 8×8 windows in an image [CHZ08] [CHZ12], i.e. the DCT is used to carry out the spatial filtering (instead of edges).

Intriguingly, and worryingly, the performance of van de Weijer et al. [JGG07] derivative approach appears highly database dependent. Thus, while by using derivatives of different order and scale (i.e. the standard deviation of the smoothing Gaussian) we can get good performance, the best derivatives and the best scale change with the images under investigation. One of the goals of this thesis is to adopt this *edge-based* idea (it really looks like this ideas has potential since it delivers good, albeit tuned results) but to try and derive a method which works across data sets without tuning.

Simple statistical algorithms can also be improved by incorporating some additional information. Forsyth [For90] observed that under a reference light, the observed colours in an image are limited and fall inside a subspace of the RGB space. Thus when estimating an scene's illuminant only lights that are consistent with the set of all plausible colours viewed under a reference light need to be considered. Furthermore, in [Fin96] a constraint, a priori, is also placed on the set of typical lights, i.e., purple lights do not occur in nature so these and other improbable lights are removed from the set of answers to illuminant estimation. Recently, a constraint in the form of a feasible illuminant set

was also incorporated by Finlayson and Trezzi [FT05] into a Minkowski Family norm formulation. Constraining the set of plausible lights helps, significantly, our ability to estimate the light. We will also adopt this constraint in the work presented in this thesis.

1.2.2 New Contributions

In this thesis we present a new algorithm for illuminant estimation incorporating the idea of colour edges with a weak constraint in illumination. A specific goal of this work is to formulate an algorithm which is not tuned to datasets. The building block in our approach is a simple x- and y- derivative calculation together with the "edge" image calculated using a Laplacian operator. These "combined" derivatives are used in all our experiments.

The foundation of our new algorithm is a fusion of four principles that we believe are complementary in nature: our *combined derivative*, *CD*, approach, the Minkowski p-norm estimator, a gamut constrained (which determines which lights are feasible), and an error optimisation to select the best illuminant for a given test image.

To evaluate our new approach we use five benchmark datasets for colour constancy algorithms: the Simon Fraser University (SFU) datasets [BMFC02] [Bar02], the Funt and Ciurea's Grey Ball dataset [CF03b], the Gehler's Macbeth Colour Checker set [GRB$^+$08] (and the linear version provided by Funt and Shi [SF10]), and the Funt and Shi's HDR dataset [FS10]. Results show that our new approach delivers good illuminant estimation, most of the time outperforming more complex art algorithms. Furthermore, our algorithm's performance can be described as stable across the different datasets, which looks promising since one of the objectives of colour constancy algorithms is to arrive at a *general* and optimum solution (non image-data-dependent).

A further contribution of this work is related to the efficiency in terms of computational cost and processing time when computing illuminant estimation. Our new *constrained* approach keeps the simplicity of statistical algorithms without giving up good performance.

A final contribution of this thesis is the experiments we realised in the context of human vision. One of the assumptions underlying our *constrained* framework for illuminant estimation is the diagonal model for illumination change. However, it is known that this model works better for "sharp" [FFB95] [Süs05] [FS00a] [FS00b] sensors in the context of human vision. Thus we adapt the principles of our *constrained* approach to illuminant spectrum recovery and also incorporate spectral sharpening to the task. We report our findings and results, and show that the performance of our algorithm is good. Furthermore, our algorithm outperforms most of the other approaches presented in the 2011 *Illuminant Spectrum Recovery Competition* organised by the Optical Society of America, OSA.

1.3 Outline of the thesis

The outline of this thesis is described below.

In Chapter 2 we present a review of the state of the art for illuminant estimation algorithms.

Our investigations begin with a short, but important, contribution set forth in Chapter 3. There, we revisit the constrained Minkowski norm formulation presented previously in the literature. We propose to adopt a modified formalism. Not only will we argue its improved suitability, it is also shown to deliver much improved results. Second, we show how the basic formalism can be implemented efficiently.

In Chapter 4 we present a new approach for colour constancy based on a combined derivative approach, a constraint in illumination, and the use of the Minkowski Family

Norm (the new approach is called CDC). Not only will we show that our new algorithm delivers excellent (and leading) estimation results, it is (unlike the previous leading methods) not tuned to the data sets. The same basic algorithm work across datasets.

In Chapter 5 we propose to incorporate the principles of Gamut Mapping as criterion to select feasible illuminant for a given image. This criterion is incorporated into our combined derivative approach. We will present experiments of this new variant and show that estimation is improved.

In Chapter 6 we apply a variant of our *constrained* approach to an experiment proposed by OSA (Optical Society of America) on Autumn 2011 for illuminant spectrum recovery. *CDC* is based on the assumption of a diagonal model of illumination change. Since the diagonal model works best for a "sharp" linear combination of cone response, we also incorporate spectral sharpening and demonstrate with our experiments and results that we can achieve good results. We compare our results with those of the competitors of the experiment last year, Autumn 2011, and we show that a simple Minkowski norm approach, using sharp sensors, outperforms all the algorithms that took part in that competition (over 30 approaches).

Finally, in Chapter 7 we present our general conclusions and future work.

Chapter 2

Illuminant Estimation

A number of algorithms for illuminant estimation have been proposed over the years and much progress has been made in the field. Nowadays, illuminant estimation is achieved to some extent, often under specific circumstances (e.g. significant colour casts on digital photographs are rare). However, a general solution (or near perfect one) to illuminant estimation has yet to be found.

To estimate scene illuminant, some sort of assumptions need to be made about the expected state of the world. Researchers in the area have closely investigated different hypotheses and so different algorithms for illuminant estimation have been proposed. While some approaches are more popular than others (for example, statistical approaches are thought as sometimes too simple), it cannot be declared that one approach is right or wrong when compared with a second one. Instead, the good (or bad) performance of one approach depends on how often their principles are met (or not) in the world [FS10]. All algorithms work poorly for some classes of image.

In this Chapter we present a review of the state of the art for illuminant estimation algorithms. Typically, these approaches are divided according to the sort of assumptions they are based on: statistical, learning approaches, gamut mapping algorithms, and physics-based approaches.

2.1 Statistical Algorithms

Statistical algorithms for illuminant estimation assume that simple statistics from an image are sufficient to infer the scene illuminant. In a first subcategory, statistics were taken from raw RGB pixel values. Recently, considering the output of linear, "derivative type", filters instead of the pixel values has been shown to improve estimation. In this section we describe the most well known statistical algorithms for these two subcategories.

2.1.1 Using Pixel Values

The Grey World Algorithm

One of the best known assumptions used for illuminant estimation is the Grey World hypothesis, which assumes that the average colour of a scene is grey [Buc80]. The Grey World algorithm estimates the scene illuminant by computing the mean of each RGB element in an image (an illuminant estimation given by Grey World is defined by Equation 2.1). The Grey World algorithm sometimes produces good results (despite its simplicity) when a sufficient range of colours are present in the image and if the image is captured under a single uniform illuminant. However, the Grey World estimate is biased by large coloured areas in the image, e.g., if a large blue area is present in the image, the estimated illuminant will be bluer than it actually is. An illuminant estimation by The Grey World algorithm is given by:

$$\mathbf{e} = [\text{mean}(\mathbf{R}) \;\; \text{mean}(\mathbf{G}) \;\; \text{mean}(\mathbf{B})]^t, \tag{2.1}$$

where $\mathbf{e}$ represents the 3-dimensional vector for the illuminant estimation, and $\mathbf{R}$, $\mathbf{G}$ and $\mathbf{B}$, are the three colour channels in the image.

Further work has focused on adapting the Grey World algorithm to achieve im-

proved estimation. For example, Barnard et al., characterise a wide range of surfaces whose mean is then used to estimate the illuminant [BCF02]. That is if, statistically, the mean of the world, under the reference light with colour $[1\,1\,1\,]$, is the vector p and the image mean for an unknown light is the RGB q then the map p/q corrects for the illuminant colour (the estimated light is the reciprocal vector q/p). Which is a complex way of saying: "what should we do if the average colour in the image is achromatic". Another strategy that has improved the Grey World estimation is to segment the image and calculate the average of all regions instead of individual pixels [BMCF02] [GJT87]. Incorporating this simple pre-processing step mitigates, to some extent, the sensitivity of the algorithm to the presence of large coloured areas in the image.

The MaxRGB Algorithm

MaxRGB is generally attributed to be part of Land's Retinex theory [Lan77]. This algorithm assumes that there is a white patch in the scene that reflects the maximum possible light of the scene illuminant. In practice, this is the maximum pixel value for each channel in the image. The estimate of the illuminant given by the MaxRGB algorithm is defined in Equation 2.2. MaxRGB is an example, similar to the case of Grey World, of statistical algorithm leading to often good illuminant estimation. However, by definition, the MaxRGB operator finds a single value in an image and use it as illuminant estimation, which is not a "robust"solution since it can be sensitive to noise (for example, when pixels having maximum R, G or B values due to error in the imaging process, such as clipping).

$$\mathbf{e} = [\max(\mathbf{R})\ \ \max(\mathbf{G})\ \ \max(\mathbf{B})]^t \tag{2.2}$$

Recently, Funt and Shi [FS10] carry out an exhaustive analysis of MaxRGB and conclude that its apparent "poor"performance may in part be due to the lack of dynamic range in standard 8-bit image data. This results in an inability to capture the full colour

range presented in the original scene. Thus, Funt and Shi proposed a new strategy to improve the MaxRGB performance by using simple pre-processing operations prior to the estimation. This pre-processing includes image resizing and smoothing operation (median filter). Other works have also showed that a smoothing operation alleviates the influence of noisy pixels [GG07] [Ebn09].

Shades of Gray

Furthermore, Finlayson and Trezzi [FT04] presented the Shades of Gray algorithm, an illuminant estimation algorithm based on ℓ^p norms, which is defined by:

$$\mathbf{e} = [\mu_p(\mathbf{R})\, \mu_p(\mathbf{G})\, \mu_p(\mathbf{B})]^t, \tag{2.3}$$

where the quantity μ_p is given by the normalised Minkowski Norm:

$$\mu_p(\mathbf{X}) = \frac{\|\mathbf{X}\|_p}{N^{(1/p)}}. \tag{2.4}$$

Shades of Gray generalises the Grey World and MaxRGB algorithms in a Minkowski Norm framework for illuminant estimation. That is, using a norm $p = 1$ (Equation 2.3), we are indeed calculating the mean operation (i.e. The Grey World estimator). When, instead, as $p \to \infty$, the Minkowski norm estimator approaches the maximum component in the image, equivalent to the MaxRGB estimation. Since for every $p > 1$, the ℓ^p norm gives an estimation somewhere in between the extreme norms, one might wonder if whether an norm in between these extremes might provide a reasonable estimation, based on a different assumption of the scene average. Thus, Shades of Gray proposes to estimate the scene illumination as a weighted colour average in a scene: the greater p is, the more important the weights for colours that have at least a high value in one of the RGB colour channels.

Significantly, across data sets the authors found that a p norm of around 4 or 6 con-

sistently delivered better performance than either Grey World or MaxRGB. Effectively, a p norm of 5 or 6 weights brighter pixels more yet is a more robust operator than the maximum.

2.1.2 Using Linear filters

In a second class of statistical algorithms there are those using outputs of linear filters instead of raw pixel values.

The Grey-Edge Hypothesis

The Grey-Edge algorithm was proposed in [JGG07] and it is a Minkoski-Norm-based algorithm on first and second derivatives in images (edges). The idea is to take a derivative (of a given order and scale) of an image and then calculate a Minkowski Norm (in the same way as Finlayson and Trezzi [FT04]). The formal definition of Grey Edge is defined by:

$$e_k^{n,p,\sigma} = \left(\int \left\| \frac{\partial^n I_k^\sigma(\boldsymbol{x})}{\partial \boldsymbol{x}^n} \right\|^p dx \right)^{1/p} \quad k \in R, G, B \tag{2.5}$$

where $I_k^\sigma(x)$ is the convolution of the image I_k with a Gaussian filter G, with σ pixel standard deviation, and n is the order of the derivative ($n = 0$ means no derivative is taken, $n = 1$ indicates first derivative and $n = 2$ second derivative). Finally, $\boldsymbol{e}$ is the 3-dimensional RGB vector of the estimated light source,

$$\boldsymbol{e}^{n,p,\sigma} = \alpha[e_R^{n,p,\sigma} \;\; e_G^{n,p,\sigma} \;\; e_B^{n,p,\sigma}]^t \tag{2.6}$$

where α is a scalar such that $||\boldsymbol{e}^{n,p,\sigma}|| = 1$, i.e., we seek only to recover the RGB of the illuminant up to an unknown multiplicative scalar.

Furthermore, the Grey-Edge algorithm was extended to a weighted Grey-Edge algorithm [GGJ09] to account for different types of edges in an image (materials, high-

light, geometry and shadows). It is expected that for the Grey-Edge algorithm, different edges can provide a different impact on the estimation. Gijsenij et al., conclude that specular edges (specular edges are those caused by highlights) provide more information about the scene illuminant and so they are given a higher weight. Even though this weighted scheme has shown to improve illuminant estimation, it requires more complex implementation and higher computational cost.

Although the Grey-Edge approach is a more rich computational formalism (and so can deliver improved illuminant estimation) its generality is also problematic. The best results for a given data set require a unique set of parameters (order of derivative, p norm and smoothing parameter). For this reason, the method seems unnaturally "tuned" per data set. Thus, the simplicity of calculating a Minkowski norm is lost. Remedying this weakness is a central theme of this thesis.

Using spatial correlations

A further example of an illuminant estimation algorithm incorporates the use of the Discrete Cosine Decomposition, DCT, of 8×8 windows in an image [CHZ08]. Chakrabarti et al. propose to model the spatial dependencies between pixels as an alternative approach to illuminant estimation (instead of considering individual pixels or derivatives). The accuracy of the estimation using this approach relies again on an adequate training phase that includes sufficient (extensive) examples of images taken under a reference (known) illumination.

Using colour categories

Finally, an algorithm based on colour category correlation is presented in [VCVBT12]. The authors use the eleven basic term categories defined by Berlin and Kay [BK69], which are defined as the universal basic colour terms defined in most evolved languages. Then, they define a *naming hypothesis*, which is an illuminant constraint used

to determine how likely a feasible illuminant is, according to its ability to transform image colours to colour categories. The above is determined for every element in a set of feasible illuminants. Then two criteria to select the best illuminant are proposed: selecting the illuminant with the maximum probability or using a combining criterion where estimations provided by other methods (for example Grey World or MaxRGB) are taken into account.

2.2 Learning Approaches

First attempts in this category are those using machine learning techniques. For example, in [CFB02] an approach for illuminant estimation using a neural network is presented. The idea behind this approach is to use a neural network to determine the chromaticity of an scene illuminant. For this, the neural network is first trained with several images under different illuminants (in [CFB02] the chromaticity of the lights are used). Once the neural network has "learnt" (or defined a series of rules), it will have a "model" to determine the chromaticity of the illuminant from an image. Approaches in the same category include those using support vector regression [FX04] [XF06] [WXL09], linear regression techniques [AGKA06] [AGA07][VAA09], and thin-plate spline interpolation [XSF+07].

2.2.1 Bayesian Methods

Other approaches, instead, are based on the Bayesian formulation [DIS95] [BF97] [Sap99] [TCRK01] [RML03] [GRB+08]. These approaches first try to model the relation among surfaces, illuminants, and sensor responses. Then, a priori distribution is calculated to define the probability that a particular light and surface exists in the world. Finally, they estimate the illuminant from the posterior distribution calculated on the image data. Even when these algorithms can achieve good results, they are still

influenced by changes in the collection of surfaces in the scene [BF97]. In an attempt to find a more appropriate model, Rosenberg et al., proposed a Bayesian approach that uses a non-Gaussian probabilistic model (in contrast with the previous model based on Gaussian distribution [DIS95] [BF97]). This approach combines the principle of the Bayesian framework combined with non-parametric image models, based on the assumption of correlation among nearby pixels [RML03].

2.2.2 Combining algorithms

It has been shown that algorithms can achieve good (or reasonable good) illuminant estimation when the image content meets the algorithm assumptions. However, no illuminant estimation algorithm has been proved to be a general and unique solution for all image data. Inspired by this observation, a final group of learning algorithms aims to combine approaches to improve estimation. Combining the output of various algorithms is proposed in [CF99] [SHF05] [BGS08]. In [CF99], three illuminant estimation algorithms are considered: Grey World, MaxRGB and an approach based on neural networks. Cardei and Funt showed that using an optimised (in a least-squares sense) weighted average of the three algorithm estimations, results in more accurate estimation than that provided by the algorithms individually. Another example [BGS08] combines the results of five (based on rather different assumptions) illuminant estimation algorithms: Grey World, MaxRGB, Grey-Edge, Color-by-Correlation [FHH01a] and Colour in Perspective [Fin96]. While combined approaches can deliver improved estimation performance, the benefit is modest. Like the general Minkowski-edge formalism, it is possible that the increased performance is due to more and more parameters being available to "tune" the algorithm.

Finally, in [GG07] [GG11], high-order statistics (of the spatial distribution of edges in an image) are used to determine what illuminant estimation algorithm to use for a specific scene content (using predefine categories of images). More specifically the

edge distribution is parameterised by two numbers which control the shape of a Weibull distribution. The Weibull parameters then index a particular constancy algorithm (a variant of the Grey Edge approach).

2.2.3 Using semantic information

Further approaches consider some sort of semantic information to account for the best colour constancy algorithms according to image categories. For example, Bianco et al. proposed an indoor-outdoor classifier for image content [BCCS08a] based on decision forests.

In [LGG+09] [RLG09], to use a stage classifier that distinguishes medium-level semantic classes is proposed [NSRG10]. Also, likelihood of semantic content have proven to be useful [JSV07]. Similar approaches applied what they called "memory color" (colours that are associated with specific object categories) to refine the estimate of the illuminant [RNK+09].

2.3 Gamut Mapping Algorithms

Let us imagine we can define a set of all feasible surfaces in the world. If these surfaces are illuminated by a particular light, their observed colours will occupy a subregion in the RGB colour space. If we then use a different light to illuminate these surfaces, the subregion representing the observed colours will change, reflecting the illumination change. This fact was first observed by Forsyth [For90] when looking at real images and it inspired the Gamut Mapping Framework for colour constancy. For example, the bluest blue RGB cannot occur under the reddest light.

Gamut Mapping algorithms work by finding a mapping (diagonal matrix) that map observed image colours inside the gamut of colours under a reference illuminant. When many maps are possible, one member must be chosen from the map set (effectively

using a criterion akin to the ideas of Grey World and MaxRGB). However, unlike the simple statistical estimator approach, the set of feasible maps places a strong constraint on the plausible light (i.e. the Grey World solution is often not possible).

Gamut Mapping algorithms have proved to be among the most promising for illumination estimation. Forsyth's original Gamut Mapping assumptions about the world were very restrictive (conditions relating to a Mondrian World, e.g., no presence of specularities, uniform illumination). More recent versions of gamut mapping deal with more realistic assumptions about the sort of surfaces that can be found in the real world and with non-uniform illumination [Fin96] [BF99] [Bar99] [Kob00] [FHH01a] [TEW01] [FX03] [FHX05] [FHT06] [GGJ07] [GGJ10]. Specifically, algorithms by Barnard and Funt [BF99], and Tominaga et al., [TEW01] can work when specularities [BF99], fluorescent surfaces [Bar99], or coloured metals [BF99] are presented in the images. Over the years different approaches have been proposed to alleviate the limitations of Forsyth's gamut mapping. In the following sections we discuss the key aspects of gamut mapping algorithms, where an optimal general solution is still yet to be found.

2.3.1 Model for illumination change.

A diagonal model for illumination change was first used in Forsyth's original gamut mapping. However, it has been much debated that the diagonal model does not hold in practice and can be only an approximation. Therefore if using this model, it is necessary to consider the cases where the model of illumination change is not appropriate.

2.3.2 Representation of the canonical gamut.

Ideally the canonical gamut represents all physically realisable surfaces in the world under a reference light. In practice it is very difficult to have the complete full set of all possible surfaces and so just a subset of these surfaces is commonly used.

Another consideration is how to represent the canonical gamut. Originally, gamut mapping algorithms represented the canonical gamut by its convex hull, enclosing all observable colours. Due to the properties of convex hulls it is sufficient to work only with the points representing the hull and not all the inner points. The issue with convex hulls is that they have a very complex form, with many faces represented by triangles (in a 3-dimensional space) and its computation is not trivial. A consequence of this convexity is that computing the set of maps requires intersecting several convex sets. Yet for a single illuminant we do not have to compute a mapping set at all. Indeed, it was shown that by incorporating the estimation into the problem formulation, Gamut Mapping Colour Constancy could be found using convex programming. In convex programming gamuts are represented using "half-spaces". Another simplified canonical gamut representation is given by a cube in the RGB space.This cube is found by taking the maximum and minimum RGB values of the gamut [MF10].Since the cube is a much simpler shape than the convex hull, gamut mapping implementation is simpler and less computationally expensive. Monsy and Funt [MF10] present experiments using a cube to represent the canonical gamut and report better results in terms of median illuminant estimation error. Furthermore, Monsy and Funt propose how to tune this simplified version of gamut mapping to minimise the maximum error.

2.3.3 2- or 3-dimensional space.

Forsyth's gamut mapping was first proposed in a RGB 3-dimensional space but due to its computationally cost in 3-dimensional space, Finlayson [Fin96] proposed the use of a 2-dimensional chormaticity space. It was proven that the two versions are equivalent in terms of recovery light approximation up to an unknown intensity.

2.3.4 Criteria for mapping selection.

Different criteria have been proposed to select a single mapping. From the set of feasible mappings Forsyth selects the maximum volume mapping [For90] [Fin96] [BF99] [Bar99], while the mean is used in [Bar95] [BF99] [Bar00] [Bar99], median in [FH00], a cone-based mean is used in [FH98], 3-d non-perspective mapping selection is used in [FH98] and a hybrid approach is proposed in [Kob00]. In [FX03] [FHX05] [MF10] an optimisation using linear programming is proposed, and recently the use of image derivatives is presented in [GGJ07] [GGJ10].

2.3.5 Illumination.

Even when gamut mapping can find a mapping that transforms an image gamut to fit inside the canonical gamut, this mapping may not represent a plausible illuminant. It is also important to ensure that the solution to gamut mapping results in a real light.

Some researchers have proposed to include a constraint in illumination to ensure that the resulting mapping corresponds to a real light [Fin96] [Bar95] [FH98] [BF99] [Bar99] [FH00] [FHH01a] [FHT06]. For example, lights in the real world are yellowish, whitish and bluish, but purple lights do not occur.This constraint in illumination can be included by considering the gamut of possible illuminants, i.e., the algorithm recovers a feasible set of mappings representing the possible unknown illuminants which are feasible. Also, the problem of illuminant estimation can be cast as one of illuminant classification: Colour-by-correlation [FHH01a] proposes a probabilistic approach for illuminant classification; [TEW01] uses an illuminant classification approach; and GCIE [FHT06] provides three different criteria to select an illuminant consistent with an image gamut from a set of possible illuminants.

2.4 Non-Statistical Algorithms

Most of the algorithms for illuminant estimation are based on the simplified Lambertian model of image formation (presented in Equation 1.6). Instead, Physics-based algorithms work by assuming that a known physical process is present (such as specularities or mutual illumination) and then they use this knowledge to derive an estimate of the illuminant [Tom96] [FDH91] [Lee92] [GB94]. The algorithms that exploit specular highlights assume the world contains reflectances which can be modelled using the dichromatic reflectance model. This dichromatic reflection model is given by:

$$\rho_k(x) = m_b(x)S(\lambda)E(\lambda)Q_k(\lambda) + m_s(x)E(\lambda_b)Q_k(\lambda_b) \tag{2.7}$$

where the factors m_b and m_s determine the relative amount of surface and specular reflectance that contribute to the overall light reflected at the position x.

Some algorithms have proposed to use specularities present in the image to estimate the scene illuminant [Tom96] [TW96] [Sha85] [Lee92] [LJBS86] [Hea91] [KSK] [GJT86] [RNI04].

The basic idea behind these algorithms is to find pixels where the body reflectance factor m_b in Equation 2.7 is zero (or close to zero) since they reflect most of the scene's light (pure specularities according to Equation 2.7 have the same colour as the illumination). Estimation will then be derived using these pixels, since their colours are similar or identical to the colour of the illuminant.

However, it is not easy to find specular highlights in images. For example, the brightest pixels in an image may or may not be specular, and, even if they are, they may not be "pure". While techniques exist for deriving the "pure" specular colour by integrating information from two or more surfaces [Sha85] these algorithms tend to work in laboratory conditions but not in the real world.

In other work on physics-based illuminant estimation, a number of authors have

instead used mutual illumination [Tom96] [FDH91] between two surfaces to provide further information of to the illuminant colour. Mutual reflection occurs when light reflected from one surface is incident upon a second [FDH91]. They analysed how sensor responses are modified in the mutual reflection region and showed that a good approximation of the surface reflectance function can be recovered by using the extra information available. However, again it has proved difficult to engineer these approaches to work outside laboratory conditions: it is not simple to identify image regions where interreflection occurs.

2.5 Evaluating Illuminant Estimation Algorithms

It is common practice to evaluate an algorithm's accuracy by determining how close the estimated light colour is to the ground truth light colour. This distance can be measured using angular error. Given the measured light $\mathbf{e} = [R\,G\,B]^t$ and an estimation $\hat{\mathbf{e}} = [R_e\,G_e\,B_e]^t$, the angular error between them is defined as:

$$\theta = \text{angle}(\mathbf{e}, \hat{\mathbf{e}}) = \cos^{-1}\left(\frac{\mathbf{e}\cdot\hat{\mathbf{e}}}{\|\mathbf{e}\|\|\hat{\mathbf{e}}\|}\right) \tag{2.8}$$

where $\mathbf{e}\cdot\hat{\mathbf{e}}$ represents the dot product of $\mathbf{e}$ and $\hat{\mathbf{e}}$. Angular error is independent of pixel intensity.

When defining an algorithm's performance or comparing several algorithms, care must be taken to select a sensible statistical figure or merit [Hor06] [HF06]. Indeed, Hordley and Finlayson showed that the error distribution tends to be strongly skewed and hence using a mean angular error (sometimes proposed in the literature) is not always an appropriate measure to compare algorithm relative performance. Instead, it was argued that the median angular error was a more appropriate summary statistic [HF06].

Angular error (and its statistical summary) is normally used to define how successful an algorithm is in recovering the scene illumination. From a mathematical and computational point of view, angular error is a good measurement to use when evaluating an algorithm's accuracy at estimating the illuminant. However, how can we determine if this 'mathematical' accuracy is relevant to a human observer (from a perceptual perspective)? Recently, Gijsenij et al. [AG09] analysed the correlation between an algorithm's performance in terms of angular error and the quality of the output image re-rendered to discard the illuminant estimated by the algorithm. From their experiments they conclude that there is in fact a high correlation: angular error correlates with perceptual judgements. In other words, subjective and objective evaluations of colour constancy algorithms are highly correlated [AG09].

When comparing an algorithm's performance, how can one determine that an improvement in the illuminant estimation is relevant? In terms of angular error, it has been shown that a deviation of 1° from the ground truth is not noticeable, and 3° is noticeable but acceptable [FHM05] [FF08].

Hordley concluded that a 2° angular error can be considered good enough for colour constancy for object recognition [Hor06]. Regarding the perceptual difference between two algorithms, Gijseni et al., show that improvements of up to 6% in terms of angular error are not noticeable to human observers [AG09]. Conversely, reducing angular error from 5 degrees to 4.5 (9%) is significant. Moreover, the same significance is, relativistically, observed in improving 3 degrees to 2.7 (it is the % step that matters).

2.6 Image data sets used to evaluate computational colour constancy algorithms

To investigate the performance colour constancy algorithms the following five benchmark image data sets are often used: the Simon Fraser calibrated data sets A and B [BMFC02] [Bar02], the HDR image data set [FS10], the Grey Ball data set [CF03b] and Gehler's Cambridge data set [GRB$^+$08] [SF10].

2.6.1 Simon Fraser University

The Simon Fraser University, SFU, datasets, contain images created under laboratory conditions [BMFC02] and it is available in [Bar02]. The images consist of thirty one scenes of coloured objects captured under eleven different lights. Images are provided in a standard 8-bit TIFF format, but also a 16-bit TIFF format is available to allow tests using extra dynamic range.

For this dataset, two groups of images are commonly used: a 321 image set (group A) and a 220 image set (group B). Figure 2.1 shows examples of images in this data set.

2.6.2 The Ciurea and Funt Grey Ball Set

The grey ball data set [CF03b] contains more than 11,000 images captured under different lighting conditions. The images were extracted from a 2 hours of video recorded in a variety of indoor and outdoor conditions. A small grey sphere is attached to the camera which appears in all scenes. This grey sphere was then used to measure the RGB of the light for every scene. Some images in this data set are shown in Figure 2.2. The images are available in [CF03a].

The Grey Ball is the largest dataset for illuminant estimation (that we know of). However, the images are extracted from a video, and hence the main disadvantage of

this data set its the high correlation of the images in the same sequence. It has been proposed the use of a representative subset instead of the complete set, for example, in [JGG07] a subset of 150 (10 images per scene) are considered representative, enough, of the whole set of images. Instead, Bianco [BCCS08b] have proposed a subset of 1135 images.

2.6.3 Gehler's MacBeth Colour Checker Set

Gehler's data set [GRB$^+$08] consists of 568 images of scenes under different lighting conditions, with a large variety of content: people, landscapes, buildings, objects indoors, outdoor images. Each scene's illuminant is obtained using a Gretag MacBeth Colour Checker Chart that was placed in the scenes (through analysis of the achromatic patches). The main advantage of this database is that contains more realistic (and with high quality) images. Figure 2.3 shows examples of images in this data set.

Since Gehler's image data set was not linear, Shi and Funt [SF10] presented a linear version of Gehler's data set, which is available on-line.

2.6.4 Shi and Funt's HDR Set

Funt and Shi [FS10] compiled a new dataset of HDR images. The images were created using multiple exposures of the same scene. To measure the scene light they take an extra image with reference colour checkers placed in the scene (similar to Gehler's data set). In Figure 2.4 examples of images in this set are presented.

2.6.5 Other datasets

There are in fact many other data sets which, while we metion here, are not use in this thesis. Further examples of image data sets available to date include the hyperspectral images of natural scenes [NFF02] [FNA04], the SFU hyperspectral data [BMFC02],

and the Barcelona Calibrated Image Dataset [VPVB09] [PBV10] [PVCV09].

2.6.6 Conclusions

In this Chapter we presented the state of the art for illuminant estimation algorithms. We described the most relevant algorithms in the field and presented the cases where they achieve accurate estimation and the cases where they fail for this task. We discussed the state of the art approaches for illuminant estimation, we describe how these computational algorithms are evaluated and the measurements that have been proposed to compare with each other. Finally, we presented the image data sets more widely used to test illuminant estimation algorithms and how they have been used. This review of the state of the art provides the basis to present a new statistical algorithm for illuminant estimation, which will be described in the following chapters of this thesis.

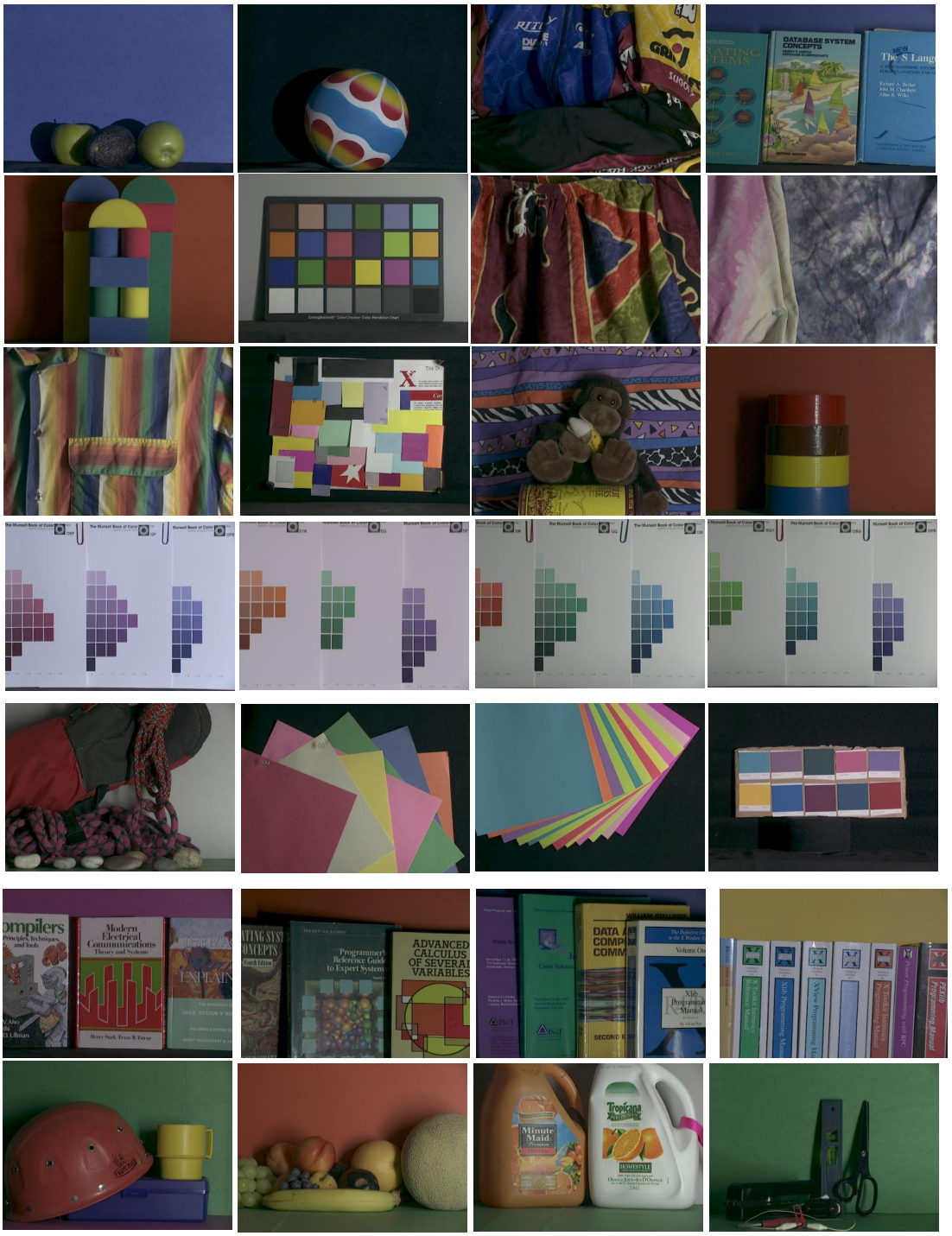

Figure 2.1: Examples of images in Simon Fraser University, SFU, data set.

Figure 2.2: Examples of images in Ciurea and Funt's Grey Ball dataset

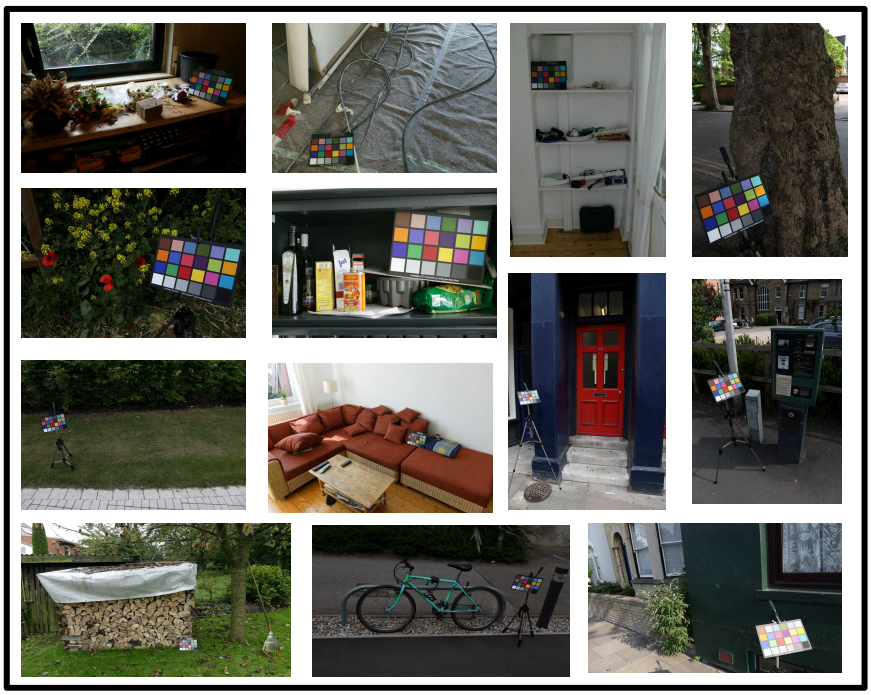

Figure 2.3: Examples of images in Gehler dataset

Figure 2.4: Examples of images in Shi and Funt's HDR dataset

Chapter 3

Revisiting the Constrained ℓ^p Minkowski Norm

In this chapter we carry out preliminary work essential for the rest of the thesis. We make two small but important contributions. First, we observe that the baseline constrained Minkowski formalism (Section 3.1) is not, in a technical sense, well formulated. We adopt a modified formulation in Section 3.2. Also, when we reimplemented Minkowski norm illuminant estimation, we found that a naive implementation was extremely slow. Indeed, the computational complexity was proportional to the number of pixels in the image. The optimisation necessarily implemented as a search albeit one which is guaranteed to find the global optimal was very slow indeed. In Section 3.3 we present a minor modification where we quantise pixel brightness and derive a formalism where the complexity of finding the illuminant is proportional to the number of quantisation levels. Our new formalism is much faster. So, while the contributions of this chapter are modest, they are essential to using constrained Minkowski norm illuminant estimation.

3.1 Constrained ℓ^p Minkowski Norm for Colour Constancy

It is well known that the mean of a distribution is the single point that is closest to all other points in the set in the least-squares sense. But, what happens if the mean colour in our image is clearly not a plausible colour? For example, if the majority of a scene is a field of purple lavender then the average will be purple, but, no purple light exists in nature. So, for a *Constrained* $p = 2$ Minkowski norm algorithm we need to calculate a given illuminant that summarises a set of points (this is the sum of the squared differences from a plausible illuminant to the image points). More generally, for the p-norm formulation, which light amongst the set of all plausible lights is the illuminant that minimises the sum of p-norm deviations?

Before presenting the details of how we incorporate this illuminant constraint, we draw attention to an important technical issue. Suppose that the best illuminant, $\boldsymbol{e}$, in a constraint set summarizes the colours in an image with an error u ($u > 0$). Furthermore, let us suppose that this light is the actual correct answer to the illuminant estimation problem. Let us now multiply all the red responses in the image by (say) 1.5 (we make the illuminant colour redder). The correct light for this second image is the light vector $\boldsymbol{e}$ but where we scale the R component by 1.5. Let this light also be in the constraint set. When we calculate the error w for this light we find that $w \neq u$. That is, the correct answer in both circumstances induces a different error measure. We found this issue to be (slightly) problematic. Thus, at the outset we change our measure of error. If (r, g, b) denotes the colour of the light (one light in a constraint set) and (R_i, G_i, B_i) is one pixel value then the error for this pixel value is $(R_i/r - 1)^2 + (G_i/g - 1)^2 + (B_i/b - 1)^2$. That is we divide by the illuminant. Now, the error with respect to the normalised RGBs is the deviation from (1,1,1). Calculating the error in this way results in the same error as the illuminant colour changes.

So, let us assume that we have a representative set W of RGBs measured for different lights. This illuminant constraint set could be obtained by measuring the RGB values of a white surface under different lights or by just demarcating part of the RGB space as being candidates for light colour. An example of such an illuminant set is presented by Barnard et al. [BMFC02]. They measure a set of plausible common illuminant sources which includes artificial and natural illuminants, i.e., the artificial light sources used to capture the images of the SFU data set [Bar02] and spectra of natural light sources measured in and around the SFU campus in Canada at various times of the day, in a variety of weather conditions. This is an example of illuminant set W. Let $\boldsymbol{w}^j$ denote the RGB of the jth illuminant in W. Let I be the $n \times 3$ matrix of image RGBs. For an illuminant $\boldsymbol{w}^j = [r, g, b]^t$ let us define the normalised image $L^j : \mathbf{L}^j_R = [R_1/r...R_n/r]^t, \mathbf{L}^j_G = [G_1/g...G_n/g]^t, \mathbf{L}^j_B = [B_1/b...B_n/b]^t$.

Given a finite set of lights W it could well be that the correct answer for an image is in W but that its magnitude is wrong. Indeed if the mean of an image is μ and we multiply the image by 10 the mean must also scale $\mu \rightarrow 10\mu$ (but we cannot expect all lights of all intensities to be modelled by W). We must somehow formulate this unknown *intensity* scale into our problem formulation.

In [FT05] the error calculation for a single light j is defined as:

$$\min_{\alpha} \left\{ \sum_{k \in R,G,B} ||\alpha \mathbf{1} - \mathbf{L}^j_k||^p \right\} \; k \in R, G, B \tag{3.1}$$

where α represents the scaling vector, $\mathbf{L}^j_k$ is the normalised image by the light j, $\mathbf{1}$ is the corresponding vector of 1s and $||.||^p$ denotes the sum of the absolute value of the arguments raised to the power p.

3.2 Modifying the constrained Minkowski formalism

Now the same scene viewed under two lights will return the same error measure. However, we found that there is an inherent bias in choosing a $\boldsymbol{w}^j$ whose elements are large. Indeed, regardless of whether it is the right answer, a large $\boldsymbol{w}^j$ will reduce the magnitude of L^j and so reduce the magnitude of the whole error expression. To mitigate this problem, in [FT05] the illuminant set contained unit length RGB vectors. But, even when this is done, there is a significant bias against more chromatic lights being selected. Chromatic lights will tend to have at least one of $\mathbf{R}$, $\mathbf{G}$ or $\mathbf{B}$ being relatively small and so, in dividing out the light will still tend to increase the magnitude of L^j and hence the error. That is if one of our lights was defined as $[100,\ 200,\ 100]$ and another by $[0.1,\ 0.2,\ 0.1]$ the former will always induce a smaller error irrespective of whether it was, in fact, the correct colour of light.

Here we modify this minimisation problem by switching the position of the scalar α to avoid this problem. We then arrive at the following *"percentage error"* formalism of illuminant estimation:

$$\min_{j,\alpha}\left\{\sum_{k\in R,G,B}||\mathbf{1}-\mathbf{L}_k^j\alpha||^p\right\} \tag{3.2}$$

where the magnitude of $\boldsymbol{w}^j$ (or the individual components thereof) is not an issue.

Let us now stack the normalized red, green and blue values as a single long vector denoted $\mathbf{L}$ which is of size $m \times 1$ ($m = 3n$ where n is the number of pixels). We write this general ℓ^p constrained illuminant estimation as:

$$\min_{j,\alpha}\left\{\sum||\mathbf{1}-\mathbf{L}^j\alpha||^p\right\} \tag{3.3}$$

It is our contention that Equation 3.3, the third constrained illuminant estimation formalism will simultaneously deliver the same error as the illuminant changes. In

addition, it is not biased by the magnitude of the numbers describing an illuminant RGB. Clearly, we want to investigate if this "percentage error" can indeed improve the illuminant estimation (or rather, the illuminant selection).

First, an important aspect for our algorithm is the selection of the illuminant set. For experiments we use two illuminant sets: the SFU 11 and the SFU 87 illuminant set. The chromaticities of the lights for each set are presented in Figure 3.1 and were measured by Brainard et al. [BMFC02]. The first illuminant set is represented by red stars and contains 87 general lights which includes day lights measured at different times of the day. The second illuminant set consists of the 11 lights that were used to create the images of the actual scenes and is represented by the blue circles in the figure. The 87 illuminant set represents the case where more general information about illuminants is available, while the 11 illuminant set represents a case where just limited information is at hand. For this experiment we will use the SFU 11 illuminant set, the SFU 87 illuminant set will be used in the following chapters.

Now that the illuminant set has been defined, we tested our new formalism using the 321 Simon Fraser University set of calibrated images [BMFC02] and the SFU 11 illuminant set. In this experiment, we used an estimator using the optimisation expression defined in Equation 3.1, for norms in the $[1, 16]$ range. We do the same for the "percentage error" formalism defined in Equation 3.3. The results for this experiment are presented in Figure 3.2, where the mean and angular error for SoG using Equation 3.1 are represented by blue square line and the red circle respectively. The results using Equation 3.3 are represented by the green triangle line and the star magenta line, mean and median angular error respectively.

We can observe from the figure that using our "percentage error" formalism helps to select the best illuminant for a test image. And, for this particular data set the performance increment is highly significant. Let us take, for example, the best result for the original SoG formalism corresponding to norm four, which achieves a mean angular

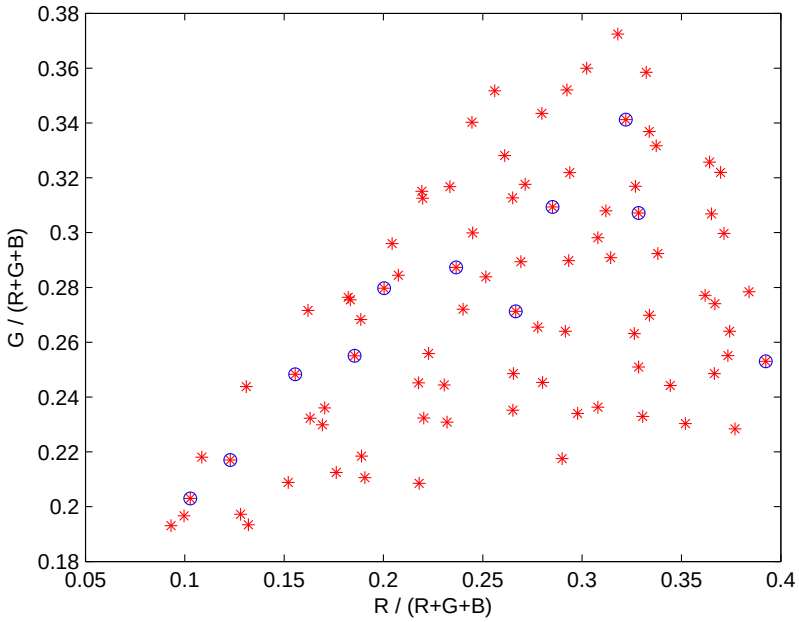

Figure 3.1: Chromaticities of the two illuminant sets used for SFU (A and B) [Bar02]. The red star points represent the chromaticities for the 87 light set, and the blue circle points represent the 11 light set.

error of 6.7 degrees and a median of 5.1. For the new "percentage formalism", the mean and median angular errors for the same norm four are 5.7 and 3.6 degrees respectively. This improvement corresponds to 1 and 1.5 degrees in terms of mean and median angular errors. This is equivalent to a 14.9% and 29.4% reduction in mean and median angular errors. This improvement is encouraging. Furthermore, we can observe from the figure that for norms larger than six, the angular error reduction achieved by using the "percentage formalism", is more significant. For example, for norm sixteen, our new illuminant estimator reduces the median angular error by 3.5 degrees, corresponding to a 60.3 % reduction.

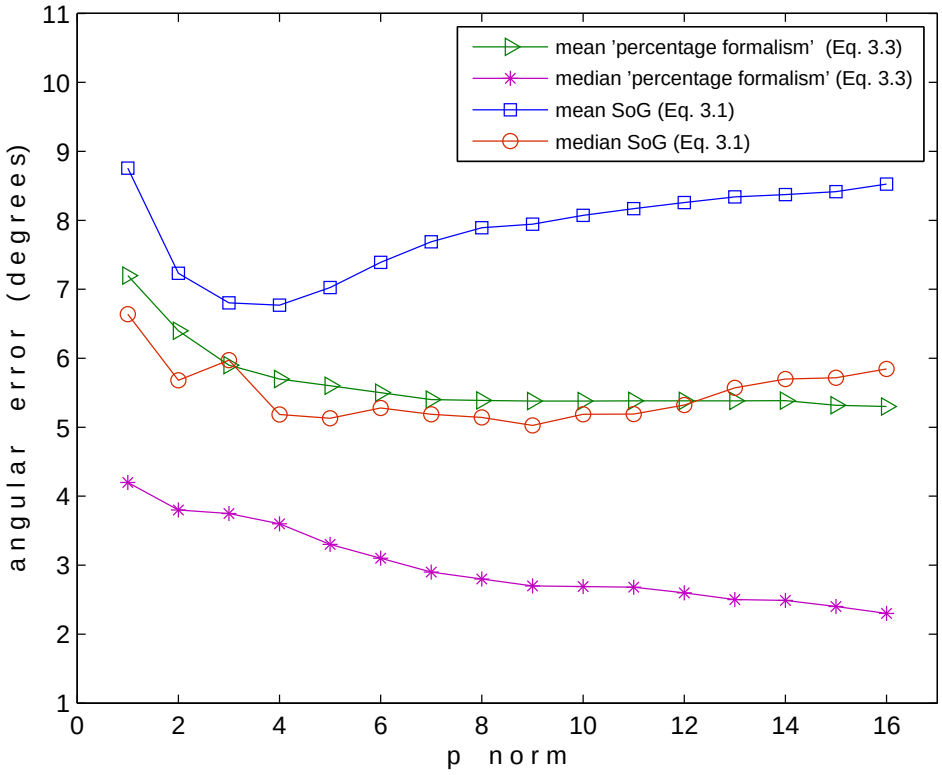

Figure 3.2: Performance for SFU A using two different optimisation criteria (see text for more details). Results for a constrained estimator using Equation 3.1 are represented by the blue square line and the red circle line (mean and median angular error respectively). For a constrained estimator using Equation 3.3, the mean and median angular error are represented by the green triangle line and the star magenta line respectively.

3.3 Improving Computational Performance

In previous sections we show how the Minkowski-based framework for illuminant estimation is extended to include extra information regarding feasible lights (in the form of a constraint in illumination) [FT05]. To select the best illuminant from a predefined set, an error is minimised for a test image based on 'distances' which are calculated for every pixel in an image. Even when this results in an improved estimation, it comes (not surprisingly) at an increase in the computational time to derive the estimation. Furthermore, this cost is large. Carrying out the minimisation might take up to several minutes

for a single image.

In this thesis we want to investigate if, for a constrained Minkowski estimator, this calculation time can be reduced. To do so, we propose simply the use of histograms to describe the image colour distribution. So that Equation 3.3 becomes:

$$\min_{j,\alpha} \left\{ \sum_{k \in \{R,G,B\}} ||\mathbf{1} - \mathbf{q}^{j,k}\alpha||^p \mathbf{H}^{j,k} \right\} \tag{3.4}$$

where $\mathbf{L}^{\mathbf{j}}$ in Equation 3.3 is replaced with the quantised sensor values in $\mathbf{q}$ and $\mathbf{H}$ is the vector with the counts for the values in $\mathbf{q}$. This equation looks similar as before. However, the vector $\mathbf{q}$ is a small number of quantised sensor values. The p norm of the quantised value vector minus the $\mathbf{1}$'s vector is then weighted by the frequency (number of times) that the value occurs in the image (for each channel). This small modification to the formalism is **essential** to making the constrained Minkowski norm algorithm practical for illuminant estimation.

So, almost by definition, we expect to see a large increase in performance, but let us measure the time taken to calculate estimation errors for commonly used datasets. We compare the performance of our constrained approach using pixels and histograms in terms of mean, median, minimum and maximum values, as well as first and third quantiles (Q_1 and Q_3 respectively) of the processing distribution.

Here we run the constrained Minkowski norm formalism for a single p-norm (we select norm 4). The timing results are summarised in Table 3.3. The SFU A data set has 321 images in total. As we transit from left to right in this graph we see the standard formalism takes more and more time to process the images. For the 321 images in the data set it takes over one second per image to carry out the conventional processing. Using histograms reduces this by a factor of 100 to 0.01 seconds.

In Figure 3.3 we present a comparison of computational time required to derive an

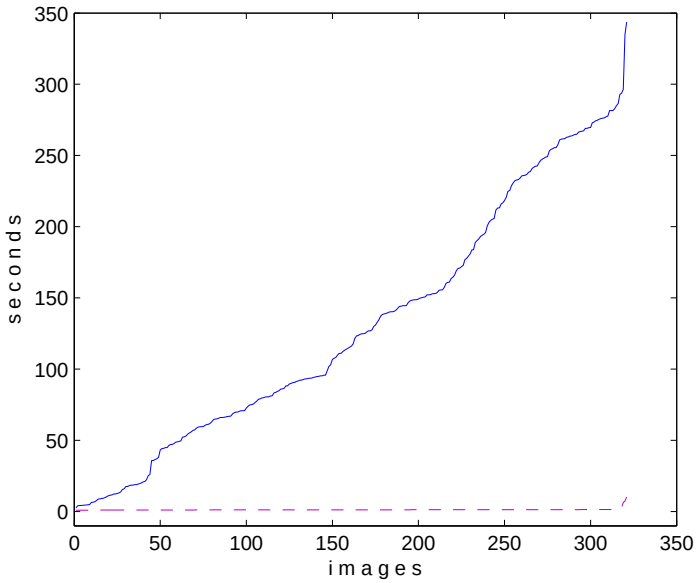

Figure 3.3: SFU A, processing time for different implementations of our constrained approach.

estimation for SFU A. The time presented in the table is given in seconds.

Table 3.1: SoGC processing time, in seconds, for SFU A

Algorithm	min	max	mean	median	Q_1	Q_3
unconstrained	0.12	0.28	0.22	0.22	0.22	0.24
constrained, pixels	2.63	343.59	131.14	116.24	64.25	203.57
constrained, histograms	0.96	10.3	1.33	1.27	1.18	1.32

3.4 Conclusions

In this Chapter we showed that: 1) using quantisation dramatically decreases the processing time for the five data sets used for experiments (the total time for illuminant estimation was reduced to more than 95%), and 2) our redefinition of a constraint based on the Minkowski Family Norm framework achieves more accurate estimation.

Chapter 4

The General ℓ^p Combined Derivative Constrained Minkowski Norm

There are many algorithms for illuminant estimation which can, broadly, be split into three classes: physics-based, statistical algorithms, and a combination of the prevous. Physics-based algorithms work by assuming a known physical process is presented, e.g., a specular highlight [Lee92] [TW96] or mutual interreflection [FDH91] [Tom96], and then an estimate is derived by reasoning about the physics of image formation. In the second class of algorithms, and the class we concern ourselves with in this chapter, an image is viewed as a distribution of measurements and the tools of statistics are applied to estimate the illuminant. Statistical approaches are arguably more general in that their operation does not require a particular physical process to be present in the scene. Examples of simple statistical approaches include Grey World [Buc80] and MaxRGB [Lan77] which respectively estimate the light as the average R, G and B in an image or as the per-channel maximum. The average and maximum of a distribution are particular instantiations of the Minkowski p-norm (with $p = 0$ and $p = \infty$). Across many data sets a p-norm in the range 4-6 reliably outperforms Grey World or MaxRGB colour constancy [FT04].

The input to statistical algorithms need not be the raw image RGBs. Van de Weijer et al. [JGG07], and Gijsenij and Gevers [GG11] have shown that a p-norm summary based on image derivatives (first and second) can deliver improved performance. The derivatives may be taken over a simple 3×3 image neighbourhood or, by incorporating a Gaussian smoothing step, over larger neighbourhoods (over scale). Chakrabarti et al. extended this approach to allow for all (excepting the DC component) terms in the discrete cosine decomposition of 8×8 windows in an image (e.g., using the Jpeg decoding) [CHZ08]. It may not be surprising that algorithms based on a particular p-norm, degree of Gaussian smoothing and spatial derivative deliver better constancy than Grey World or MaxRGB. We are effectively giving the algorithm designer several parameters in which colour constancy performance can be tuned to a particular data set.

Simple statistical algorithms can also be improved by incorporating some additional *weak* physical information. Forsyth [For90] observed that under a reference light the colours of surfaces do not fall in the RGB cube but rather in a convex subset. It follows that we need only to consider illuminant estimates which are consistent with this 'gamut' constraint. In [Fin96] a gamut constraint is also placed on illumination colours: typical lights are bluish, whitish or yellowish (in the gamut) but saturated purples and greens are disallowed (not in the gamut).

In this chapter, we present a new algorithm that falls in the statistical algorithm class. The foundation of our new algorithm is the Minkowki p-norm statistical estimator. We then incorporate an illumination constraint (which determines which lights are feasible) into the Minkowski-norm formulation. Our work improves upon a previous approach reported in the literature [FT05] by fusing two approaches which have recently been presented in the literature and which we believe are complementary in nature. First, we review and then extend the Minkowski p-norm approach so that it incorporates, in a mathematically rigorous way, a constraint on illumination. Second, we incorporate the idea of image derivatives into the Constrained Minkowski norm problem formulation

(since there is evidence that colour constancy on derivatives seems to work better than on the colours themselves).

This next algorithm, *Combined Derivative*, is a simple fixed combination of x and y spatial derivatives together. Rather than laboriously tune our algorithm by choosing the kind of derivatives we use (order and scale) we instead propose a simple combination of first and second derivative information. We do however, optimise over p.

Evaluating our algorithm on five benchmark data sets our *Combined Derivative Constrained*, *CDC*, Minkowski norm algorithm performed very well indeed. It is always the top performer. This is a remarkable result as our constrained Minkowski norm approach is a simple approach and is far less complex than other leading algorithms.

4.1 Background

Over many years several authors have proposed computational theories and algorithms for illuminant estimation. All algorithms work by making some sort of prior assumption about the expected state of the world. For example Land [Lan77] implicitly assumes that every image contains a white patch (or a white per channel e.g., bright yellow and bright blue appear in the scene) and so the maximum $[R\,G\,B]$ responses are a good estimate for the RGB values of the light source. Buchsbaum proposed the Grey World hypothesis that assumes the average reflectance of all surfaces in a scene is achromatic [Buc80]. Under this assumption, the average RGB values in an image are used to approximate the illuminant RGB values. If the shape of spectral reflectances are sufficiently limited (typically this is not the case) Maloney and Wandell [MW86] provide an elegant mathematical algorithm for spectral light recovery. Finlayson and Trezzi developed the "Shades of Gray" algorithm in [FT04], which applies Minkowski Norms to compute illuminant estimation. They showed that Grey World and MaxRGB are represented by the extreme Minkowski norms, where Grey World is equivalent to ℓ^1 and

MaxRGB is equivalent to ℓ^∞. Shades of Gray works well but not as well as algorithms based on more complex inference procedures.

Highly relevant is the recent work on using derivatives in colour constancy (instead of image RGBs). Van de Weijer et al. [JGG07] developed a 'Grey Edge' algorithm which estimates the illuminant colour by taking the average of the absolute values of image derivatives (Grey World over edges). Van de Weijer et al. also incorporated the Minkowski Norm approach into their Grey Edge algorithm. Consistent with the original Shades of Gray algorithm a p norm greater than one and less than infinity was found to work best (the best Minkowski norm algorithm was somewhere between Grey World and MaxRGB). Bianco et al. [BCCS08b] showed that by choosing a particular instantiation of the Grey Edge algorithm (e.g. particular p or particular type of derivative) according to image content improves illuminant estimation. Gijsenij et al. [GGJ09] extended the Grey Edge approach by classifying edge types (material, shadow-geometry, highlights) to formulate a weighted Grey Edge algorithm. Chakrabarti et al. [CHZ08] generalised the Grey Edge algorithm using a Discrete Cosine decomposition in a small image neighbourhood (8×8 pixel window to compute spatial correlation) to capture spatial variation. Combining the terms in the Discrete Cosine expansion in a probabilistic framework was shown to deliver good results.

All the above algorithms (except for the classification step in the Gijsenij and Gevers [GG11] algorithm) considered so far are simple. They are simple in the sense they have low computational complexity and are straightforward to implement. The best performance in colour constancy research is found for *complex algorithms*. Forsyth [For90], Finlayson [Fin96] and Barnard et al. [BMCF02] make the assumption that image RGB values under a reference light fall within a convex gamut. Illuminant estimation in this framework involves finding the mappings that take the gamut of image colours to the reference lighting conditions. The gamut mapping computation involves the intersection of many 3-dimensional convex sets and so is computationally complex. Algorithms

based on probabilistic reasoning e.g. Finlayson et al. [FHH01b] and Brainard and Freeman [BF97] are also more complex but here the complexity is, in part, due to modelling the prior assumptions.

4.1.1 Minkowski ℓ^P Norm for Illuminant Estimation or Shades of Gray

A Minkowski norm is defined in a N dimensional real space. Let $\mathbf{X} = [X_1 ... X_N]^T$ be a vector in $\mathbb{R}^N$. For every $p \geq 1$ the quantity

$$\|\mathbf{X}\|_p = \left\{ \sum_{i=1}^{N} |X_i|^p \right\}^{1/p} \tag{4.1}$$

defines a norm in $\mathbb{R}^N$ [KF75] and X_i represents the ith component of $\mathbf{X}$. Here and throughout this thesis *bold* letters denote vector quantities.

As $p \to \infty$ the pth norm approaches the maximum component of $\mathbf{X}$. Rewriting this equation to normalise it by the number of points in the distribution we have:

$$\mu_p(\mathbf{X}) = \frac{\|\mathbf{X}\|_p}{N^{(1/p)}}. \tag{4.2}$$

Finlayson and Trezzi [FT04] propose an illuminant estimator given by the following equation:

$$\hat{\mathbf{e}} = [\mu_p(\mathbf{R})\, \mu_p(\mathbf{G})\, \mu_p(\mathbf{B})]^T, \tag{4.3}$$

where $\hat{\mathbf{e}}$ is the 3-dimensional RGB vector for the illuminant estimation, and $\mathbf{R}$, $\mathbf{G}$, $\mathbf{B}$ represent the three colour channels in the image.

4.1.2 The Grey Edge hypothesis

Van de Weijer et al.'s Grey Edge hypothesis [JGG07] assumes that the average absolute derivative (1st, 2nd or nth order) in an image is achromatic. Across several studies the average of image derivatives was found to be a better estimator of illuminant colour than the average of the pixels.

Van de Weijer et al. also incorporated the Minkowski-norm idea into the problem formulation [JGG07] and found that a norm greater than one (i.e., not Grey World) and less than infinity (not MaxRGB) delivers the best illuminant estimation. Grey World, Shades of Gray, and Grey Edge are similar and can be combined in a single framework for illuminant estimation methods based on low-level image features:

$$\left(\int \left\| \frac{\partial^n I_k^\sigma(\boldsymbol{x})}{\partial \boldsymbol{x}^n} \right\|^p dx \right)^{1/p} = e_k^{n,p,\sigma} \quad k \in R, G, B \tag{4.4}$$

where k denotes R, G or B.

$$I_k^\sigma(x) = I_k \otimes G^\sigma \tag{4.5}$$

where $I_k^\sigma(x)$ is the convolution of the image I_k with a Gaussian filter G, with σ pixel standard deviation, and n is the order of the derivative ($n = 0$ means no derivative is taken, $n = 1$ indicates first derivative and $n = 2$ second derivative). Finally, $\boldsymbol{e}$ is the 3-dimensional RGB vector of the estimated light source,

$$\boldsymbol{e}^{n,p,\sigma} = \alpha[e_R^{n,p,\sigma} \;\; e_G^{n,p,\sigma} \;\; e_B^{n,p,\sigma}]^t \tag{4.6}$$

where α is a scalar such that $||\boldsymbol{e}^{n,p,\sigma}|| = 1$, i.e., we seek only to recover the RGB of the illuminant up to an unknown multiplicative scalar.

4.2 Constrained ℓ^p Minkowski Norm for Colour Constancy

Let us recall the error calculation for a single light j defined in Section 3.1. Let W be a representative set of RGBs measured for different lights. Let $\boldsymbol{w}_j$ denote the RGB of the jth illuminant in W. Let I be the $N \times 3$ matrix of image RGBs. For an illuminant $\boldsymbol{w}_j = [r, g, b]^T$ let us define the normalised image $L_j : \mathbf{L}_j^R = [R_1/r...R_N/r]^T, \mathbf{L}_j^G = [G_1/g...G_N/g]^T, \mathbf{L}_j^B = [B_1/b...B_N/b]^T$. Our *'percentage error'* formalism of illuminant estimation for a jth light in W is given by:

$$\min_{j,\alpha} \left\{ \sum_{k \in R,G,B} ||\mathbf{1} - \mathbf{L}_j^k \alpha||^p \right\} \tag{4.7}$$

Let us now stack the normalized red, green and blue values as a single long vector denoted $\mathbf{L}$ and using the previous notation n, p, σ to denote the nth order differential operator, pth Minkowski norm and smoothing with σ pixel standard deviation. We define a normalised image vector $\mathbf{F}$, so that $\mathbf{F}_j = \mathbf{L}_j^{n,p,\sigma}$, where $\mathbf{F}_j$ is of size $M \times 1$ ($M = 3N$ where N is the number of pixels). We write this general ℓ^p constrained illuminant estimation as:

$$\min_{j,\alpha} \{ \text{ err} = ||\mathbf{1} - \mathbf{F}_j \alpha||^p \} . \tag{4.8}$$

4.2.1 Solving for α

Let us differentiate Equation 4.8 with respect to p (and assuming p is even) we see that, for a fixed light and F_i denoting the ith component of $\mathbf{F}$:

$$\frac{\partial \text{err}}{\partial \alpha} = p \sum_{i=1}^{M} F_i (1 - F_i \alpha)^{p-1} \tag{4.9}$$

and the second derivative

$$\frac{\partial^2 \text{err}}{\partial \alpha^2} = p(p-1)\sum_{i-1}^{M} F_i^2(1-F_i\alpha)^{p-2}. \tag{4.10}$$

Clearly, (because p was assumed to be even) the above expression is always positive, which implies that $\frac{\partial err}{\partial \alpha}$ is monotonically increasing and this in turn implies that the optimisation has one minimum. We can solve for α by simple gradient descent. A similar (but slightly more complex) argument can be made for odd powers. In either case a simple search suffices to find α.

4.2.2 Combined Derivative Approach

As discussed earlier, through choosing the best Minkowski norm (best p), smoothing parameter (σ), and order of derivative (n), the basic Minkowski-norm algorithm can be tuned to a particular image content. Moreover, published results find quite different parameter combinations work best for individual images [BCCS08b]. Here, we propose a single variant of the algorithm which we call the *Combined Derivative Approach*, *CD*. We do this for two reasons. First, we wish to evaluate how much (or how little) incorporating a constraint in the illuminant helps without the increment being clouded by the question of parameter selection. Second, the fixed-parameter operation typically reflects real-world conditions where cameras typically have a single white point estimation algorithm.

Our **combined derivative** optimisation is written as:

$$\min_{\alpha,j} \left\{ \; ||\mathbf{1} - \mathbf{F}_j^{1,p,1}\alpha||^p + ||\mathbf{1} - \mathbf{F}_j^{2,p,1}\alpha||^p \right\}, \tag{4.11}$$

where we minimise the percentage error for the first derivative operator (x and y derivatives treated separately) and the deviation for the second derivative (x-y second derivatives added together). Both first and second derivatives are calculated after smoothing by

a symmetric Gaussian filter with a standard deviation of one pixel. The only parameter left to optimise for is p. Thus, the above expression brings us back to the simplicity of the Shades of Gray algorithm in terms of its number of free parameters.

There is no *magic* in the above expression: we suggest the above formula only because we observed first, the real power of edges and, secondly, that sometimes first derivatives worked better than second and vice versa. Therefore, it seemed natural to combine first and second derivatives to account for a wider range of information for diverse image content. Thus avoiding the task of tuning parameters to image content which results in a more general algorithm.

4.3 Experiments and Results

To investigate the performance of our general constrained p norm algorithm we test it with five benchmark image data sets: the Simon Fraser calibrated data sets A and B [BMFC02] [Bar02], the Grey Ball data set [CF03b], the Gehler's Cambridge data set [GRB$^+$08] [FS10] and the HDR image data set [FS10]. For each data set there are per image measured answers of the correct light. The set of such lights is used as our illuminant data set.

We follow common practice of determining how close the estimated light colour is to the ground truth light. The distance is measured using angular error. Given the measured light $\mathbf{e} = [R, G, B]^T$, we run the colour constancy algorithm to recover the estimated light $\hat{\mathbf{e}} = [R_e, G_e, B_e]^T$. The angular error between them is defined as

$$\theta = \text{angle}(\mathbf{e}, \hat{\mathbf{e}}) = \cos^{-1}\left(\frac{\mathbf{e} \cdot \hat{\mathbf{e}}}{\|\mathbf{e}\|\|\hat{\mathbf{e}}\|}\right) \tag{4.12}$$

For each data set we test 4 instantiations of our algorithm. First, we run the constrained and unconstrained algorithm using the raw pixels of the image (here we are focussed only on how important the illuminant constraint is in the context of the 'Shades

of Gray' Minkowski norm algorithm). We label these two algorithms as **SoG** and **SoGC** (**C** for constrained). In a second test we wish to evaluate the power of derivatives. Here we test the combined derivative approach without constraining the illuminant, denoted **CD**, and the combined derivative constraint approach, labelled **CDC**.

4.3.1 Simon Fraser data sets A and B

The Simon Fraser calibrated data set is a large data set of many colourful objects measured under different lights. The data set is divided in two groups: the first one, group A, contains 32 scenes of colourful objects; the second, group B, consists of 20 scenes of one object per scene. Group A and B's scenes are captured under 11 coloured lights (not all scenes in set B are imaged under all lights). Details about the images, how they were collected and the illuminants used can be found in [BMFC02] [Bar02].

The results of our experiments on this data set is shown in Figure 4.1 (data set A), and Figure 4.2 (data set B). The figures show mean and median angular error respectively for **SoG** (blue square line), **SoGC** (green dotted line), the combined derivatives **CD** (black star line) and its constrained counterpart **CDC** (magenta triangle line).

We make three observations about the results. First, across p norms, as expected, **SoGC** performs better than unconstrained **SoG**: an illuminant constraint helps. Second, across p norms **CDC** (combined derivative and constrained) works better than using derivatives alone (**CD**). **SoGC** and **CDC** algorithms clearly work best overall. While **SoGC** provides better mean and median angular error than **SoG** and **CD**, **CDC** performance is better across p norms for both mean and median angular error.

In Table 4.1 we highlight the results of our algorithms (for the best norm) and for Gamut Mapping (method described in [For90]) and Gamut Constrained Illuminant Estimation (GCIE) [FHT06]. Recently Funt and Shi showed that MaxRGB's performance for illuminant estimation over preprocessed images leads to more accurate illuminant estimation [FS10]. They used two preprocessing operations: a 5×5 median filtering,

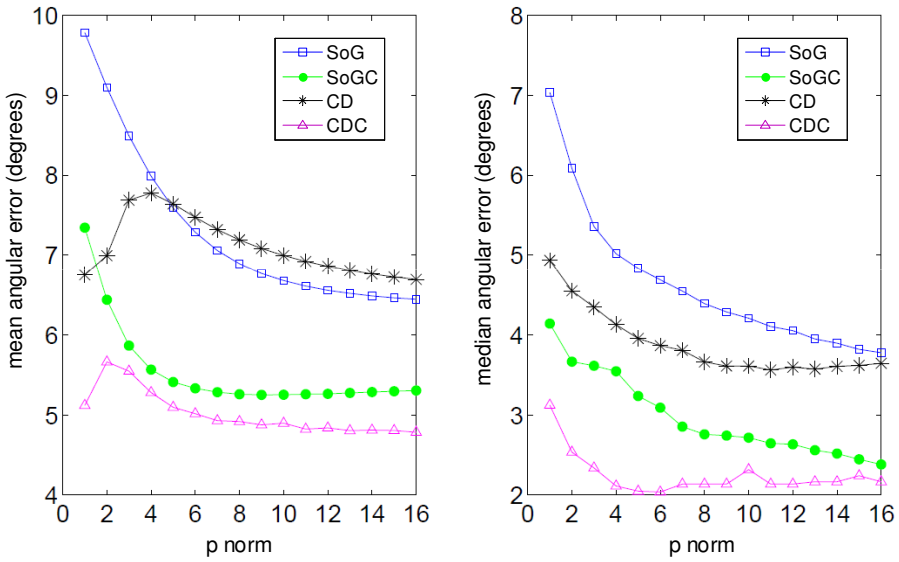

Figure 4.1: A comparison of algorithm performance, in terms of mean and median angular error, is presented using the SFU dataset group A [Bar02]. The square blue line represents **SoG**, **SoGC** is represented by the (dotted green line), **CD** is represented by the star black line), and **CDC** is represented by the magenta triangle line.

and image resizing to 64×64 pixels using bicubic interpolation. We also include these results for comparison. In Table 4.1, Funt and Shi's results are labeled as 'MaxRGB filtering' and 'MaxRGB resizing'. This idea of smoothing as a preprocessing step for colour constancy was originally discussed by Choudhury and Medioni [CM09].

From Table 4.1 it can be observed that our **CDC** algorithm achieves good results and is the best among simple algorithms. For data set A, **CDC** performs as well as the much more complex Gamut Mapping approach, and it is very close in performance to GCIE, but at significantly lower computational cost. The GCIE algorithm works by modelling the gamut of colours that can appear under different lights. Each light (and its plausible gamut) is tested in turn to see how well it accounts for the real image data.

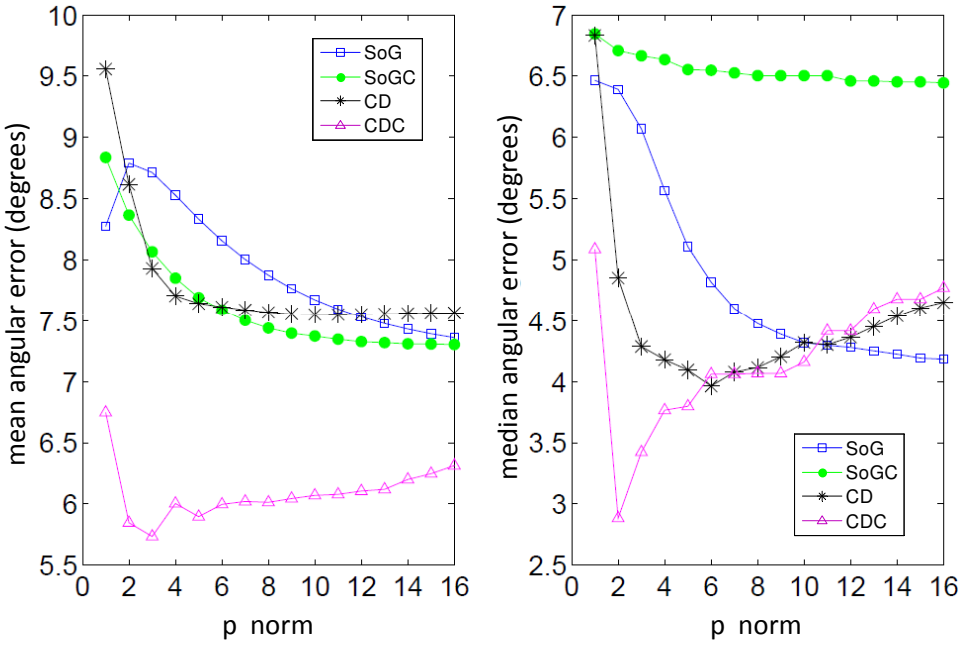

Figure 4.2: A comparison of algorithm performance using the SFU data set [Bar02], group B is presented. The algorithms included are: **SoG** (square blue line), **SoGC** (dotted green line), **CD** (star black line), and **CDC** (magenta triangle line). The performance is measured in terms of mean and median angular error (degrees).

Considerable fitting was carried out to optimise the shape of the gamuts to represent the image data to return the best results. That our simple normalisation procedure works so well without optimisation is encouraging. For data set B, **CDC** outperforms **CD**, **SoG** and **SoGC**, with mean angular error of 5.85 and median angular error of 2.88.

4.3.2 HDR Images, Simon Fraser University

This data set has been compiled by Funt and Shi [FS10] and contains 105 scenes in HDR format. This data set was designed so that data was not clipped and algorithms could be tested on cleaner data (than for other data sets). MaxRGB in particular is known to perform less well in the face of extensive clipping artefacts.

Table 4.1: The results of our experiments on Simon Fraser University data sets A and B [Bar02]. We highlight the best results of our algorithms for each data set. For data set A we also include results for other algorithms presented in [FS10].

Algorithm	angular error (degrees)			
	Group A		Group B	
	Mean	Median	Mean	Median
Grey World	9.8	7.0	7.35	6.48
MaxRGB	9.1	6.5	7.43	4.38
MaxRGB, filtering	5.8	3.4	5.8	3.1
MaxRGB, resizing	5.6	3.1	6.5	3.0
1st order Grey Edge,$p = 7$ A and B	5.6	3.2	7.5	6.1
2nd order Grey Edge,$p = 7$ A and B	5.2	2.7	7.4	4.1
SoG,$p = 7$ A and B	6.4	3.7	7.35	4.18
SoGC,$p = 16$ A and B	5.3	2.3	7.30	6.45
CD,$p = 5$ A and B	7.6	4.2	7.62	3.96
CDC,$p = 5$ A and B	**5.1**	**2.0**	**5.85**	**2.88**
Gamut Mapping	-	2.9	-	-
GCIE v3, 11 lights	-	**1.3**	-	-
GCIE v3, 87 lights	-	2.6	-	-

In order to estimate the ground truth for scene illuminant they placed four mini Macbeth Colour Checkers in each image, each in a different position and angle with respect to the camera and the predominant scene illuminant. Achromatic patches in each colour checker provide an RGB estimate of the scene illuminant. The average of the 4 RGB illuminants provided by each of the colour checkers is used as the scene illuminant.

Overall, it is clear from Figure 4.3 that a constraint on the illuminant does aid algorithm performance. However, the effect is less evident than in the Simon Fraser data set. The lesser performance benefit is, we believe, related to the way the ground truth is 'estimated' (which draws attention to the fact that some of the scenes are really viewed under mixed lighting conditions). While we, ultimately, seek algorithms for these scenes it is not clear how a simple statistical approach developed here can (reasonably) unpick one light from another or find the best compromise for illumination.

Funt and Shi reported results of MaxRGB and Grey World on this data set: Grey World, mean and median angular of 7.9 and 7.3, 'MaxRGB filtering', mean and median angular error of 6.3 and 4.3, and 'MaxRGB resizing', mean and median angular

Table 4.2: Performance of our algorithms using the Funt and Shi's HDR Image data set [FS10] is presented. The performance is measured in terms of the mean and median angular error (degrees). Boldface highlights our algorithm's best result.

	angular error (degrees)	
Algorithm	Mean	Median
Grey World	7.9	7.3
MaxRGB, filtering	6.3	4.3
MaxRGB, resizing	6.3	3.9
1st order Grey Edge	5.98	3.96
2nd order Grey Edge	5.91	4.0
SoG, $p = 5$	6.6	5.1
SoGC, $p = 5$	6.6	4.8
CD, $p = 5$	6	4.1
CDC, $p = 5$	**5.8**	**3.8**

error of 6.3 and 3.9 respectively [FS10]. **CDC**'s best performance is for $p = 9$, with mean angular error of 5.8 and median of 3.62. Our *CDC* delivers the best performance amongst these simple statistical methods.

In Table 4.2 we compare the performance of **SoG**, **SoGC**, **CD**, and **CDC** algorithms with other methods available in the literature. Compared with previously published results both the **CD** and **CDC** algorithms perform best overall. We can observe in the table that our *CDC* delivers the best performance overall.

4.3.3 Ciurea and Funt data set

This is a large image data set that contains 11340 images, commonly used to evaluate colour constancy algorithms. Created by Ciurea and Funt [CF03b], the images are taken from a video sequence recorded in a variety of indoor and outdoor places. A small grey ball was placed in a fixed position at the bottom of the right hand side of the camera. The area of the grey ball in an image is then used to compute the RGB values of the scene illuminant. For our experiments the area of the grey ball is masked so that it does not bias how algorithms perform. Further, following van de Weijer et al. [JGG07] we use a subset of 150 of the images because the whole image set is, in fact, a sequence of

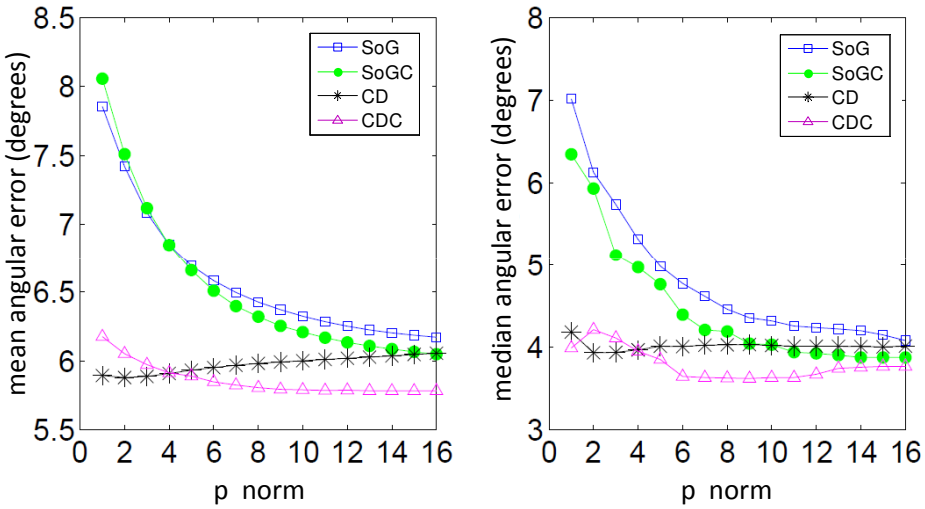

Figure 4.3: A comparison of performance for four algorithms on HDR Funt and Shi's data set [FS10] is presented. Where **SoG** is represented by the square blue line), **SoGC** is represented by the dotted green line, **CD** is represented by the star black line and **CDC** is represented by the magenta triangle line. The performance of the algorithms is measured in terms of the mean and median angular error (degrees).

150 clips. We need only test one frame per clip.

In Figure 4.4 we show the results for **SoG**, **SoGC**, **CD**, and **CDC** algorithms. As before we observe that the constrained approach achieves significantly better performance than the unconstrained method. Even though the effect is less dramatic than for the Simon Fraser data set, it can be observed from the figure that the best results are achieved by **CDC** for norms 3 and higher.

Gijsenij and Gevers's [GG11] best result for Grey Edge is mean angular error of 6.1 and median of 5.2 for second order; from Gijsenij and Gever's parameter tuning approach the best results are mean and median angular error of 5.0 and 3.7 respectively [GG11]; and Chakrabarti et al. 's spatial correlation with mean and median angular error

Table 4.3: A comparison of performance of our new constrained approach using the Ciurea and Funt data set [CF03b] is presented. Results for other state of the art algorithms are also included. For this experiment the 150 image subset proposed in [JGG07] is used.

	angular error (degrees)	
Algorithm	Mean	Median
Grey World	7.9	7.0
MaxRGB	6.8	5.3
1st order Grey Edge	5.9	4.7
2nd order Grey Edge	6.1	4.9
Tuning Approach*1	5.0	3.7
Spatial Corr.*2	6.0	4.4
SoG, $p = 8$	6.4	5.1
SoGC, $p = 4$	6.1	4.3
CD, $p = 2$	5.4	4.4
CDC, $p = 12$	**5.2**	**3.8**

of 6.0 and 4.4 [CHZ10]. While the best results for **CDC** were for $p = 12$ with mean angular error of **5.25** and median of **3.81**, performing as well as the aforementioned algorithms.

In Table 4.3 we summarise the results of our algorithms and compare them with the results presented in [CHZ10] for different algorithms. In all cases our **CDC** approach works best overall.

4.3.4 Gehler data set

This is a set of 568 images from indoor and outdoor scenes taken in Cambridge, England, and is perhaps the most realistic *photographic* data set. Each picture either contains people, places and different objects and many are typical photos (like the ones we all take with our point and shoot cameras). Unlike the images we take ourselves, in every scene there is a Macbeth colour checker placed in a known position. The scene illumination is given by the achromatic patches in the Macbeth colour checker.

The nature of the images in this data set is much debated. It has been pointed out that the images were created using an automatic white balance operation included in the camera used to capture the images. Some algorithms need to work with linear data

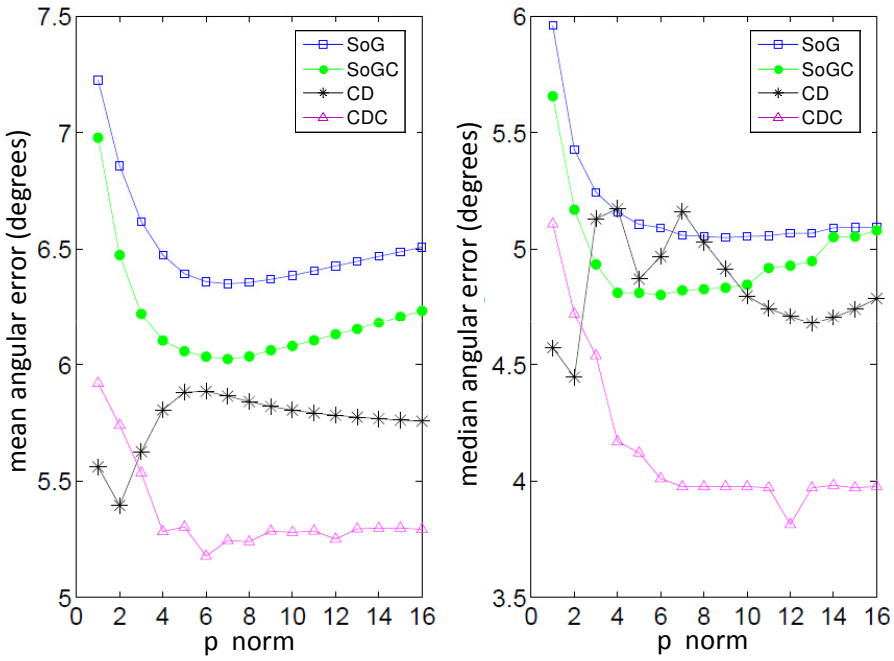

Figure 4.4: A comparison of algorithm performance for our four algorithms on the Ciurea and Funt's Grey Ball data set [CF03b] is presented: **SoG** (square blue line), **SoGC** (dotted green line), **CD** (star black line), and **CDC** (magenta triangle line). The results are presented in terms of the mean and median angular error (degrees). For this experiment the 150 image subset proposed in [JGG07] is used.

so Shi and Funt [SF10] presented a new linear version of this data set.

Figure 4.5 shows results on Shi and Funt's linear version of Gehler's Macbeth data set for our algorithms: **SoG** (square blue line), **SoGC** (dotted green line), **CD** (star black line) and **CDC** (magenta triangle line). It can be observed from the figure that our combined derivatives constrained **CDC** performs better that **SoG**, **SoGC** and remarkably better than its unconstrained counterpart **CD**. Results on linear Gehler's Macbeth data set shows very encouraging results.

Recently Gijsenij et al. presented results for most algorithms [GGJ11] on this data set. In Table 4.4 we compare some of these algorithm performances with our own. We observe from the table that the **CDC** approach performs best among statistical al-

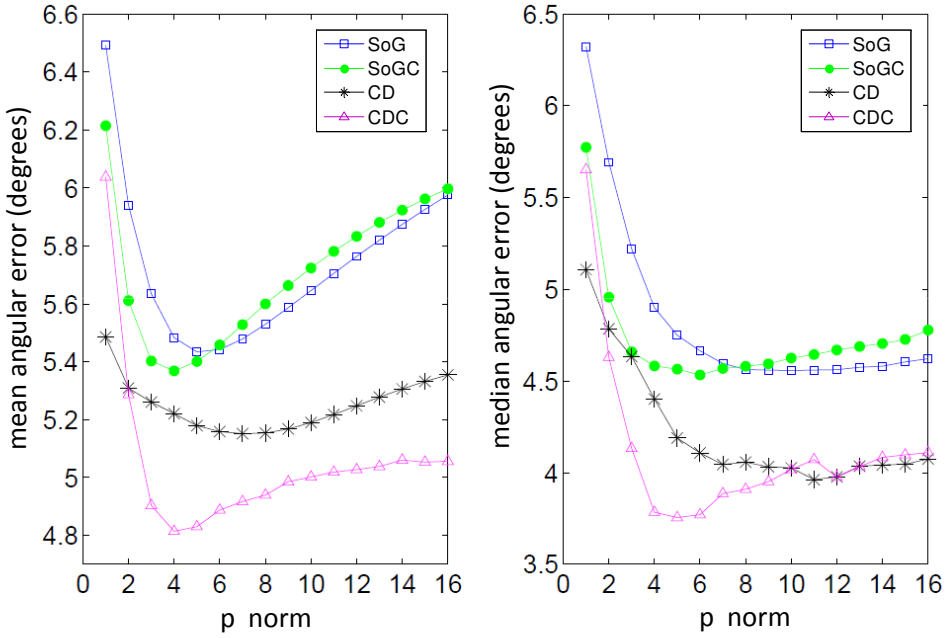

Figure 4.5: Performance for four algorithms on Shi and Funt's linear version [SF10] of Cambridge data set [GRB+08] are presented: **SoG** (square blue line), **SoGC** (dotted green line), **CD** (star black line), and **CDC** (magenta triangle line). Performance is presented in terms of mean angular error (left panel) and median angular error (right panel).

gorithms. Only the approaches which effectively apply different algorithms to different images depending on the scene content perform significantly better [GG11] [CHZ08].

4.4 Conclusion

We have presented a new approach for illuminant estimation. Our algorithm is based on the use of image derivatives and Minkowski norms. We have showed that these two approaches are complementary and by combining them good illuminant estimation is achieved. Furthermore, we add a *weak* illuminant constraint to our preferred *Combined Derivatives Constrained* approach. Experiments on five of the most common image

Table 4.4: A comparison of algorithm performance on Shi and Funt's linear version [SF10] of Gehler data set [GRB$^+$08]. Our new algorithm is compared with other methods available in the literature (* results presented in [GGJ11])

Algorithm	angular error (degrees) Mean	Median
Grey World	6.4	6.3
MaxRGB	7.6	5.7
1s order Grey Edge	5.3	4.5
2nd order Grey Edge	5.1	4.4
Spatial Corr. (without reg)*1	5.9	5.1
Spatial Corr. (with reg)*1	**4.0**	**3.1**
SoG, $p = 10$	5.7	4.6
SoGC, $p = 6$	5.5	4.5
CD, $p = 5$	5.2	4.2
CDC, $p = 5$	**4.9**	**3.8**

data sets proved the good performance of our algorithm.

The *Combined Derivatives Constrained* method is in the simple class of statistical algorithms, yet its almost excellent performance across all data sets is either better or close to the performance of the best complex algorithms.

Chapter 5

Extending the Constrained Minkowski approach using Gamut Mapping

The chapter will start with an introduction to gamut mapping, continuing with its formalisation and principles. Next, a review of the state of the art is presented, where the most important aspects of gamut mapping are discussed. We then incorporate *CDC*, *Combined Derivative Constrained*, for a new version of gamut mapping. We test the new approach in the image sets more widely used for colour constancy algorithms and present our results and findings. Experiments using real and synthetic images provide evidence that our new approach achieves good colour constancy while requiring a lower computational cost than previous algorithms. We finish the chapter with conclusions and future work.

5.1 Colour Constancy

A camera's response to surfaces, or observed colour, depends both on the surface properties and the lighting conditions illuminating the scene. An invariant colour descriptor is defined here as the camera response to surfaces under a known light. Colour con-

stancy can be defined as the problem of finding colour descriptors that are independent of the lighting conditions, and colour constancy algorithms attempt to solve this problem.

Gamut mapping algorithms, following from Forsyth [For90], were based on a diagonal model of illumination change where an estimate of the illuminant is recovered in the form of a mapping transforming observed colours to invariant colour descriptors. In Forsyth's original algorithm a 3-dimensional diagonal model mapping RGB colour observations to colour descriptors was proposed [For90]. Then Finlayson [Fin96] modified Forsyth's gamut mapping and proposed a 2-dimensional chromaticity algorithm to map observations to descriptors.

In gamut mapping a canonical gamut represents sensor responses to all possible surfaces under a canonical light. An image is captured for a scene under unknown lighting conditions, or in other words, sensor responses to a set of surfaces are imaged under an unknown light. These sensor responses are represented by an image gamut. Then a mapping that transforms the image gamut to fit inside the canonical gamut is computed and effectively estimates the unknown illuminant. Once an estimate of the illumination is recovered, an equivalent image of the scene under a reference light can be determined by applying this mapping to the image data.

Gamut mapping can achieve good illuminant estimation but its implementation is normally complex and computationally intensive (at least compared with simple ideas like Grey World and Grey Edge). A great deal of research has been carried out both to exploit its potential and to try to solve its weaknesses. In the following sections we provide a detailed analysis of the gamut mapping approach, its more important aspects and the most relevant variations to it since first introduced in 1990 [For90]. We then present a new algorithm of gamut mapping by using a new constrained approach and by including our *Combined Derivative Constrained*, *CDC* approach, presented in Chapter 4.

5.2 Gamut Mapping

Gamut mapping was first presented by Forsyth [For90] when observing that all possible colours viewed under one light are restricted to a specific area in the RGB space. Under this assumption any change in the light will be reflected in the RGB space of these observable colours. In other words, any variation in this RGB space reflects a variation in the light. This change of the RGB colour space can be used to determine the change in illumination. For example let us take Figure 5.1: a set of surfaces are imaged under two different lights (Figure 5.1 (c), and (d)). The gamut of colours for these surfaces under the two different lights are presented in Figure 5.1 (a) and (b) for each light respectively. Indeed, we can observe the gamuts are slightly skewed reflecting each illuminant, i.e., the area in the RGB space of this set of surfaces change to reflect the change in the light.

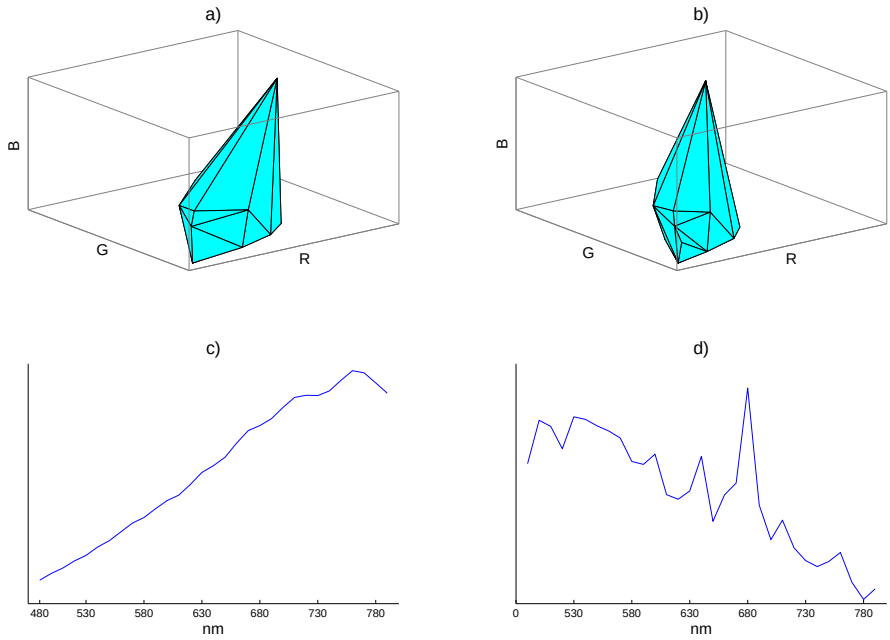

Figure 5.1: A set of surfaces illuminated by two different lights. The gamut of colours is skewed in the direction of the illuminant

Gamut mapping aims to find mappings that transforms an unknown light gamut so that it is contained within the canonical gamut. Such a mapping represents the change in illumination and one can apply the inverse of this mapping to the canonical light to estimate the unknown light (e.g. if the canonical light is $\mathbf{w}$ and D maps unknown to canonical, the Gamut Mapping estimate of the light is $D^{-1}\mathbf{w}$).

However, there are many diagonal maps taking an image gamut inside the canonical gamut. Forsyth's original gamut mapping solves for all possible mappings, and then selects the mapping with maximum volume as the solution. At the time of its publication, Forsyth's Gamut Mapping outperformed all other state of the art algorithms for colour constancy. A significant amount of research has been done since Forsyth first presented Gamut Mapping [Fin96] [FH98] [Bar00] [FHH01a] [FX03] [FHX05] [FHT06] [GGJ07] [GGJ10] [MF10].

5.2.1 Formal definition of Gamut Mapping

Forsyth assumed that the observed colours in the scene depends on the scene illuminant. This assumption is based on the observation that surfaces cannot reflect more light than it is cast on them. For example, if a sensor response to a surface is strongly red, the illuminant lighting the surface cannot be deep blue. From this observation and by imposing some constraints in the world, a scene illumination can be estimated starting from the observed colours in the scene.

These constraints on the world are quite restrictive since they are based on the Mondrian world assumptions and they do not correspond to scenes in the real world. The constraints for Mondrian world are:

1. Surfaces are matte and frontally presented. There are no shadows.

2. There is only one illumination and this is uniform across the scene.

3. All surfaces are Lambertian (and all reflection is diffuse). There are no fluorescent surfaces in the scene.

Gamut mapping imposes a constraint on possible illuminants since it assumes that the space of possible colours depends on the light. Sensor responses to all **physically** realisable surface reflectances under a canonical light form the **canonical gamut**. This is defined to be a convex set of RGB responses obtained by imaging all reflectances under the canonical illuminant. In practice a good approximation can be formed by considering a representative subset of surface reflectances [Fin96]. The canonical gamut serves as a constraint for ruling out certain illuminants.

Sensor responses to surfaces under an unknown light form the **image gamut**. Suppose that a few reflectances are viewed under unknown light conditions generating image colours. We know that under the canonical light these same reflectances would induce responses which fall inside the canonical gamut. If we can state a formal relationship between these two gamuts we will have some knowledge of the unknown light. This is only valid if certain restrictions are imposed in the world, and if colour responses under different lights can be approximated by a linear transform [For90] [MW86].

Let us represent an image taken under an unknown light o, with a set I^o of n sensor responses, $I^o = \{\mathbf{p}_1^o, \mathbf{p}_2^o, ..., \mathbf{p}_n^o\}$. For this image, we would like to determine the corresponding set of sensor responses $\mathbf{I}^c$ which would be observed under a reference light, c. Solving this problem implies determining a mapping $\mathcal{F}$ such that:

$$\mathbf{p}_i^c = \mathcal{F}^{o,c}(\mathbf{p}_i^o), i = 1...n \tag{5.1}$$

For a reference light, c, we define the canonical gamut denoted by $\Gamma(C)$. A change of illumination implies a different gamut of observable image colours and this new gamut is related to the first one by a mapping $\mathcal{F}$. That is, if $\Gamma(C)$ denotes the canonical gamut under the reference light, c, and $\Gamma(I)$ denotes the image gamut under the unknown light, o, then:

$$\mathbf{p}^o \in \Gamma(O) \leftrightarrow \mathcal{F}^{o,c}(\mathbf{p}^o) \in \Gamma(C) \tag{5.2}$$

The gamuts $\Gamma(C)$ and $\Gamma(I)$ taken together, provide sufficient constraint for solving the set of mappings that transform $\Gamma(I)$ inside $\Gamma(C)$ and these form the gamut mapping solution to colour constancy.

It follows then that colour constancy was a problem of parameterising the linear transforms which take the image colours into the canonical gamut. Forsyth's MWEXT algorithm does just that:

$$\forall \mathbf{p}^o \in \Gamma(I^o), M\mathbf{p} \in \Gamma(I^c) \tag{5.3}$$

The 3x3 map M is a possible solution to the colour constancy problem if and only if each RGB response vector $\mathbf{p}$ in the image gamut, $\Gamma(I)$, is mapped inside the canonical gamut $\Gamma(C)$.

To this point, characterising the solutions to Equation 5.3 is problematic in three respects. First, in order for gamut mapping to work the following assumptions are imposed: the surfaces should be matte, have uniform illumination and have no specularities. In general these conditions are too restrictive. Second, linear maps have nine parameters. This implies that solving Equation 5.3 is a nine-dimensional problem. Forsyth [For90] found this to be "computationally laboriously". Third, not all linear maps correspond to plausible changes in illumination.

To overcome these three problems Forsyth [For90] proposes that mappings should be restricted to three parameter diagonal matrices. When Forsyth proposed this restriction it was deemed rather arbitrary, but shortly after Finlayson et al. [FDF94a] showed that it is generally reasonable (if a diagonal model is not a physically realistic model of light change it can be made to hold by an appropriate change of sensor basis). This led to the second version of Forsyth's gamut mapping, known as CRULE. It solves for all those diagonal matrices D which take the image gamut into the canonical gamut.

$$\forall \mathbf{p}^o \in \Gamma(I), D\mathbf{p}^o \in \Gamma(C) \tag{5.4}$$

The uth point in I^o can be mapped to the vth point in I^c by applying a diagonal matrix $D^{v,u}$ such that:

$$D^{v,u}\mathbf{p}_u^o = \mathbf{p}_v^c \tag{5.5}$$

The set of diagonal mappings, $\mathcal{D}^{c,o}$, that simultaneously map all points of $\Gamma(I)$ into $\Gamma(C)$ is the intersection of all the mappings transforming the individual points of $\Gamma(I)$ into $\Gamma(C)$:

$$\mathcal{D}^{c,o} = \bigcap_{i=1}^{n} \mathcal{D}_i^{c,o} \tag{5.6}$$

Let us illustrate with the example shown in Figure 5.2 how this idea works (where for ease of illustration we look at mapping two component colours instead of three). Let us suppose that we want to map the three coloured points in Figure 5.2 (b) to the canonical gamut in Figure 5.2 (a). The mappings transforming each vertex in (b) inside the canonical gamut in (a) are represented in (c). The intersection of the three mappings is the grey area in (c).

To choose a good candidate from $\mathcal{D}^{c,o}$ and thereby provide a unique estimate of the unknown illuminant, Forsyth employs the heuristic of selecting the mapping with the maximum volume [For90] and consequently the mapping with a wider range of colours.

With the diagonal restriction, colour constancy is simpler and has lower computational cost. However, gamut mapping is still computationally expensive and only an approximate solution to Equation 5.4 is actually calculated [For90].

Forsyth tested his CRULE algorithm on images which meet Mondrian world conditions [For90]: the scenes contained flat matte reflectances under uniform illumination. For such scenes, CRULE was capable of delivering good colour constancy. Achieving colour constancy in such conditions was Forsyth's aim when he presented gamut mapping in 1990. However, even though research has been done to develop the ability

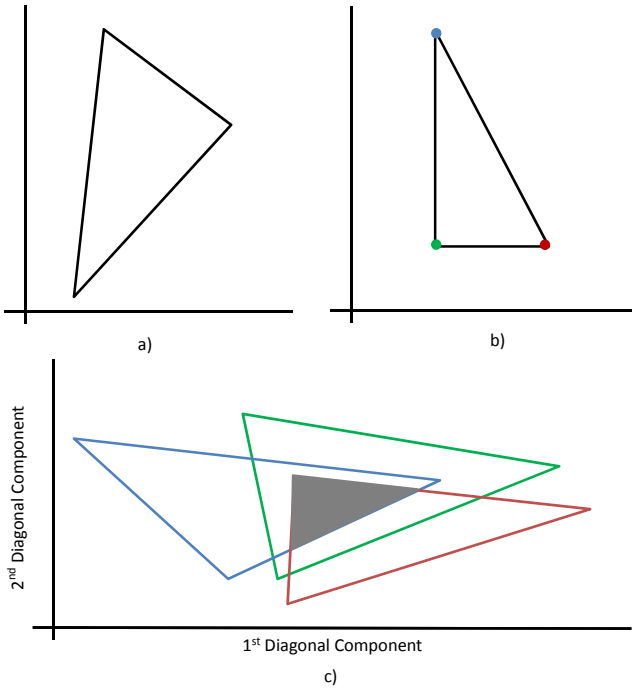

Figure 5.2: Intersection of feasible mappings, where a) represents a canonical gamut, b) represents an image gamut, and c) is the intersection of mappings transforming the three points in b) inside the canonical gamut in a)

of the algorithm to deal with "more interesting objects than matte, frontally presented surfaces" [For90], there is still scope for further improvements. Different versions to Forsyth's original version have been proposed to tackle different aspects of gamut mapping.

More recent versions of gamut mapping deal with more realistic assumptions about the sort of surfaces that can be found in the real world and with non-uniform illumination [Fin96] [Bar00] [FHH01a] [FX03] [FHX05] [FHT06] [GGJ07] [GGJ10]. Specifically, algorithms by Barnard and Funt and Tominaga et al. can work when specularities, fluorescent surfaces, or coloured metals are presented in the images.

5.3 CDC and a generalised gamut mapping approach

We are interested in the principles that Gamut Mapping, GM, provides about colour information on images. CDC estimates the illuminant using statistical theory but no canonical gamut constraint is used. However, GM has been shown to be very powerful colour constancy algorithm. Thus, here we fuse them into a new colour constancy algorithm, Gamut Combined Derivative Constrained, GCDC.

Our new approach for colour constancy is based on the following assumptions:

1. We have some knowledge about physically realisable lights in the world that allows us to build an illuminant set. This set contains all possible and physically realisable lights. Let us define a set of RGB values corresponding to real lights, W, such that the jth element in the set is of the form: $\mathbf{w}_j = [r_j, g_j, b_j]$ (constraint in illumination from CDC). The same illuminant chromaticity presented several times corresponding to the same light but at slightly different intensities. This is an important nuance. In common with most algorithms for colour constancy we assume the input image has a maximum value of 1. It could be that the diagonal matrix $[1,\ 1.33,\ 2]$ cannot be a solution to colour constancy (corresponding to the light colour $[1,\ 3/4,\ 0.5]$. But, the mapping $0.9[1,\ 1.33,\ 2]$ does work i.e., it is necessary to assume that the prevailing light is bright in order to map the image gamut inside the canonical gamut. That is, we have found that only the brightest lights need to be considered.

2. The observed colour in a scene are constrained by the illuminant conditions (Assumption of Gamut Mapping).

3. Under a diagonal model for illumination change, each $\mathbf{w}_j$ in the illuminant set defined in Assumption 1, corresponds to a mapping, D_j, that transforms an image gamut, such that:

$$D_j = diag\left(\left[\frac{1}{r}, \frac{1}{g}, \frac{1}{b}\right]\right) \tag{5.7}$$

We then define the set of m mappings such that $\mathcal{D} = [D_1, D_2, ..., D_m]$. Note that here we assume that the canonical light is $[1, 1, 1]$.

4. Given assumptions 2 and 3 we could determine which lights from the illuminant set are consistent with a given test image. That is, to determine if the illuminant $\mathbf{w}_j$ could be the scene illuminant, we evaluate if its corresponding mapping, D_j, transforms the test image gamut to fall inside the canonical gamut. Only those lights with corresponding mappings that transform the image gamut inside the canonical gamut are considered candidates for the scene illuminant. Thus we discard lights that cannot be the scene's illuminant (derived from Asumption 2), and just keep those lights that are consistent with the test image, i.e., we form a set, W', of lights corresponding to these mappings. W' is a subset of W.

5. Once we have W' we want to select a single light as the scene's illuminant. We define an error criteria to be optimised, so that the light that minimises it is the best candidate and so selected as the illuminant for the given test image. This criterion is provided by our *CDC* approach according to Equation 4.11. As a point of detail, Assumption 1 may return more than one light brightness (for the same chromaticity of light). This is no concern to CDC. So, W' contains only one element (for each unique chromaticity).

5.4 Experiments and Results

We test our algorithm on different sets of images both synthetically generated and real images, using the five best known and widely used image datasets to test colour constancy algorithms. We show that our results on images taken under controlled conditions (SFU A and SFU B) are very good and our approach outperforms other state of the art algorithms. This is expected since these images have been created under very specific requirements to test the theory behind a model. Yet, when testing our algorithm

on real images, the performance is also good and most times better than state of the art illuminant estimation algorithms.

Simon Fraser University, SFU, sets A and B

In Chapter 4 we showed that our *CDC* algorithm outperforms the state of the art algorithms for SFU A and B. So we first want to determine if adding an extra constraint in the illuminant, GCDC, helps illuminant estimation. In Figure 5.3 results for an experiment using the SFU 11 illuminant set (see Figure 3.1 for the 11 and 87 illuminant set chromaticities) are presented. We present mean and median angular error as a function of p norms. The results for our GCDC are presented for norms from one to sixteen. We can observe that the performance of GCDC is considerably improved when compared with our *CDC* approach. Furthermore, we can observe from the figure that the angular error as a function of the p norm is quite stable in terms of its median and considerably smaller, reduced even by half.

Results of GCDC using the actual lights is presented in Figure 5.3. We can observe from the figure that the performance of GCDC over the p norms is also stable. The angular error in terms of median angular error is quite small, this is expected since for this experiment only the 11 lights used to create the scenes are included in the illuminant set. However this is good evidence of the good performance of GCDC if a consistent and relevant illuminant set is provided, which is ideally the result of including the constraint in illumination.

We next want to test our *constrained* approach in the case when the available information about the feasible illuminants is more general (in the previous experiment we test the case when only the true lights were included in the illuminant set). Thus, we carry out a further experiment for SFU A using the SFU 87 illuminant set and present the results in Figure 5.4. We can observe from the figure that even though the mean and median angular errors are not improved, our *constrained* approaches, **CDC** and **GCDC**

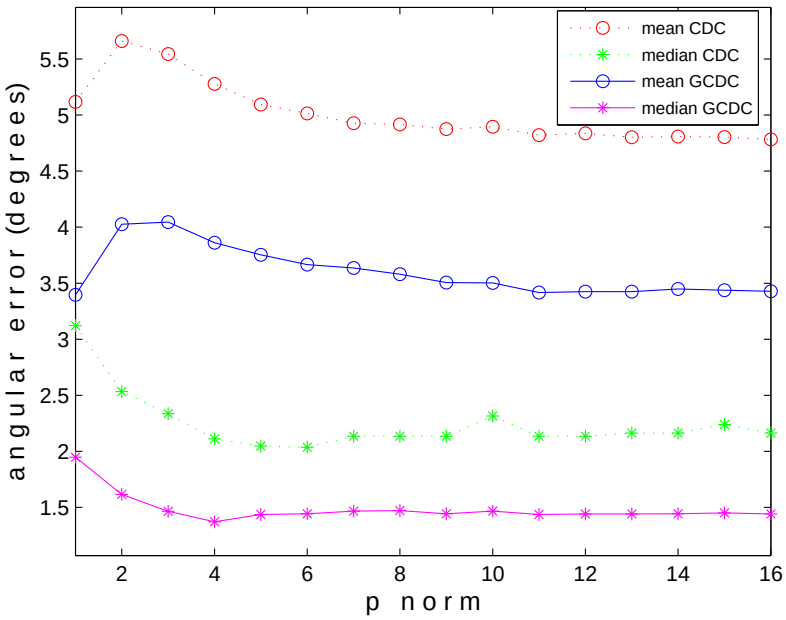

Figure 5.3: A comparison of algorithm performance on the SFU A dataset [Bar02] for: CDC (red circle represents the mean angular error while the green star line represents the median) and GCDC (blue circle and magenta star represent the mean and median angular error respectively). For this experiment, the SFU 11 illuminant set was used.

achieve the most accurate results over norms. These results are still better than other state of the art algorithms (see Table 5.1). Furthermore, we can observe from these results the importance of a well characterised illuminant set to achieve a reasonably good estimation.

Results for the experiments in SFU B using the SFU 11 and 87 illuminant sets are presented in Figure 5.5 and in Figure 5.6 respectively. The images for the SFU group B are similar to the ones in SFU group A and so good performance of our algorithm is expected. These images were created synthetically and we know that the diagonal model for illumination change holds for this sort of images, so we expect our algorithm to deliver a good and stable illuminant estimation, as it is shown in the figure.

In Table 5.1 the results for the SFU set, groups A and B, are presented. We can

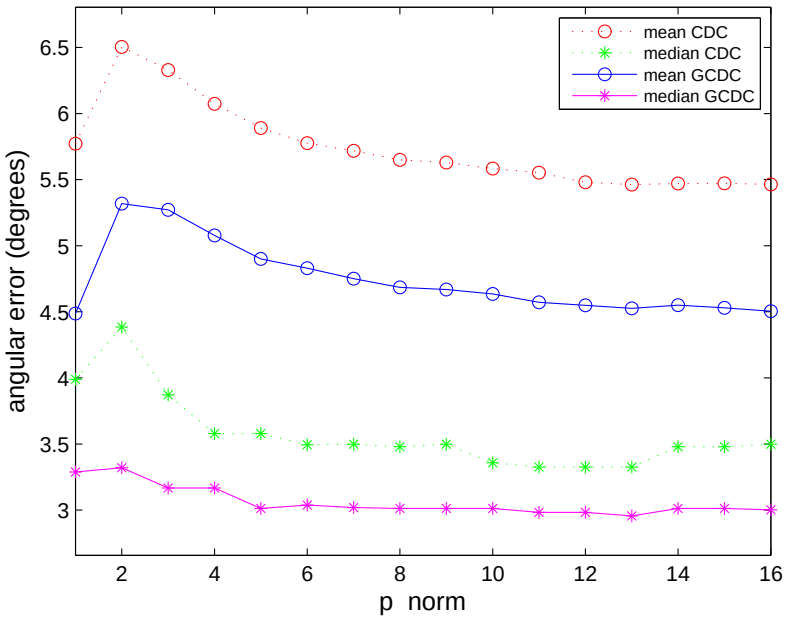

Figure 5.4: A comparison of algorithm performance on the SFU A dataset [Bar02], with the SFU 87 illuminant set, for: CDC (the mean and median angular error are represented by the red circle and green star lines respectively) and GCDC (blue circle and magenta star lines represent the mean and median angular error respectively)

observed that our new GCDC outperforms Gamut Mapping and GCIE v3 for 87 lights results for SFU A: original Gamut Mapping achieves a median angular of 2.9, GCIE v3 for 87 lights 2.6, while our GCDC achieves 1.3. GCIE v3 for 11 lights achieves a median angular error of 1.3, while for norm 4 and also 11 lights we obtain a median angular error of of 1.1.

For SFU B, the best result for CDC is for norm 5, with a mean angular error of 5.85 and a median angular error of 2.88, while our GCDC achieves the best result for norm 2, with a mean angular error of 3.19 and a median of 1.0.

Ciurea and Funt Grey Ball dataset

Results for Ciurea and Funt dataset are presented in Figure 5.7. We compare the performance of our *CDC* algorithm against GCDC, where plain blue and green lines rep-

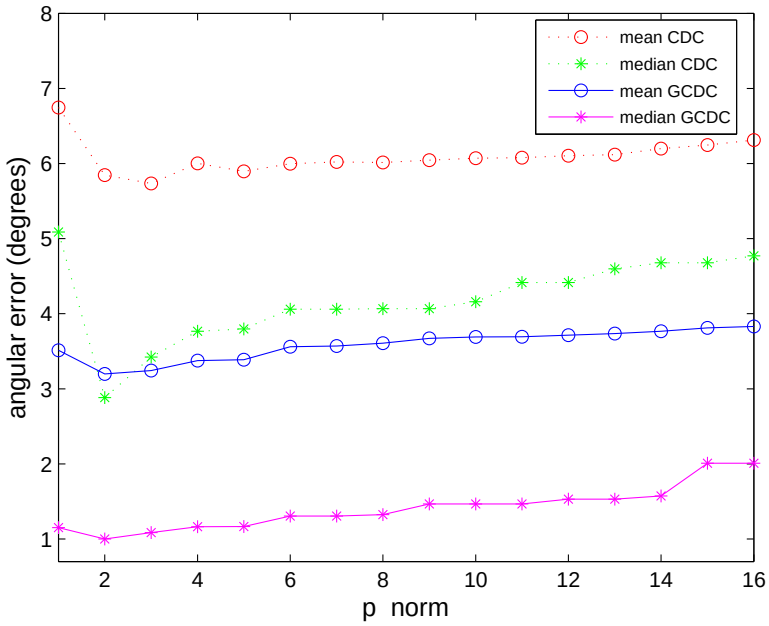

Figure 5.5: A comparison of algorithm performance on the SFU B dataset [Bar02] for: CDC (mean angular error is presented by the red circle line and the median by the green star line) and GCDC (mean and median angular error are represented by the blue circle and magenta star lines respectively). For this experiment the SFU 11 illuminant set was used.

resent the mean and median angular error respectively for *CDC*, and the star blue and green line represent mean and median angular error for GCDC. There has been much discussion about the nature of the images in this datatset, especially regarding the question if the images have been provided with gamma correction or not. In these experiments we used the images as provided by the authors (ready for display) [CF03b] to be able to compare our results with those presented by other works. For this dataset the results show again that the new GCDC achieves significantly better results than its predecessor *CDC* algorithm. This is a good evidence of the improvement by adding the constraint in illumination.

In Table 5.2 we compare the results for our GCDC with other algorithms presented in the literature and our previous version of CDC. We can observe that our GCDC achieves better performance in terms of both mean and median angular error, for norm

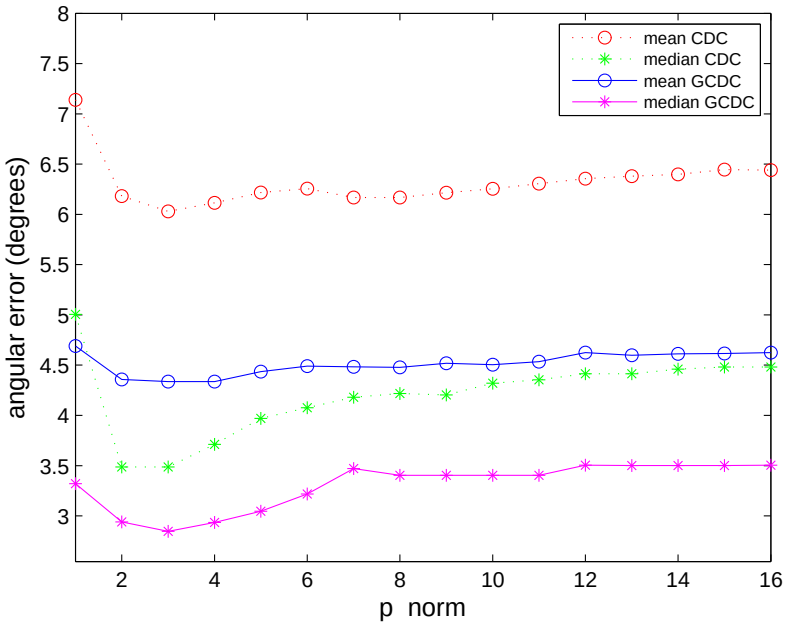

Figure 5.6: A comparison of algorithm performance on the SFU B dataset [Bar02], with the SFU 87 illuminant set, for: CDC (red circle and green star lines represent the mean and median angular error respectively) and GCDC (blue circle and magenta star lines represent the mean and median angular error respectively)

9, with 5.1 and 3.9, respectively. While the previous CDC version achieved, for norm 10, a mean angular error of 5.3 and median of 4.3. Our *CDC*s algorithms work best overall.

Gehler dataset

As in Chapter 4 we use Funt and Shi's linear version of the Gehler dataset. In Figure 5.9 we presented a comparison between the performance of *CDC* and that of *GCDC*. We can observe from the figure that in terms of the mean angular error, GCDC is of very little help. For norms larger than 4, we can see a small improvement in terms of median angular error. We suspect Gamut Mapping does not improve our *CDC* algorithm for this set as the range of illuminants is modest (SFU A and B have much more chromatic lights). Gamut Mapping, typically, does not hold for many lights in this set.

Table 5.1: Result of Experiments on Simon Fraser University data sets A and B [Bar02]. We highlight the best results of our algorithms for each data set. For data set A we also include results for other algorithms presented in [FS10].

Algorithm	angular error (degrees)			
	Group A		Group B	
	Mean	Median	Mean	Median
Grey World	9.8	7.0	7.35	6.48
MaxRGB	9.1	6.5	7.43	4.38
MaxRGB, filtering	5.8	3.4	5.8	3.1
MaxRGB, resizing	5.6	3.1	6.5	3.0
1st order Grey Edge,$p = 7$ A and B	5.6	3.2	7.5	6.1
2nd order Grey Edge,$p = 7$ A and B	5.2	2.7	7.4	4.1
SoG,$p = 7$ A and B	6.4	3.7	7.35	4.18
SoGC,$p = 16$ A and B	5.3	2.3	7.30	6.45
CD,$p = 5$, A and B	7.6	4.2	7.62	3.96
CDC, $p = 5$ A and B	5.1	2.0	5.85	2.88
Gamut Mapping	-	2.9	-	-
GCIE v3, 11 lights	-	1.3	-	-
GCIE v3, 87 lights	-	2.6	-	-
GCDC, v11, $p = 4$ A and $p = 2$ B	**3.8**	**1.3**	**3.19**	**1.0**
GCDC, v87, $p = 5$ A and $p = 3$ B	5.0	3.0	4.4	2.9

Table 5.2: Result of Experiments with the Ciurea and Funt data set [CF03b]. The 150 image subset proposed in [JGG07] is used for this experiment.

Algorithm	angular error (degrees)	
	Mean	Median
Grey World	7.9	7.0
MaxRGB	6.8	5.3
1st order Grey Edge	5.9	4.7
2nd order Grey Edge	6.1	4.9
SoG, $p = 8$	6.4	5.1
SoGC, $p = 4$	6.1	4.3
CD, $p = 2$	5.4	4.4
CDC, $p = 10$	5.3	4.3
GCDC, $p = 9$	**5.1**	**3.9**

Gijsenij et al. [GGJ11] presented results for this linear version of Gehler dataset for most common algorithms for colour constancy and we include them in Table 5.3 in addition to our previous results for CDC. We also include the results for our new GCDC, which achieves the best result for norm 5, with a mean angular error of 4.8, and a median of 3.7. In this dataset our *CDC*'s algorithms perform best overall for basic statistical algorithms as well, being outperform just by the more complex pixel-based

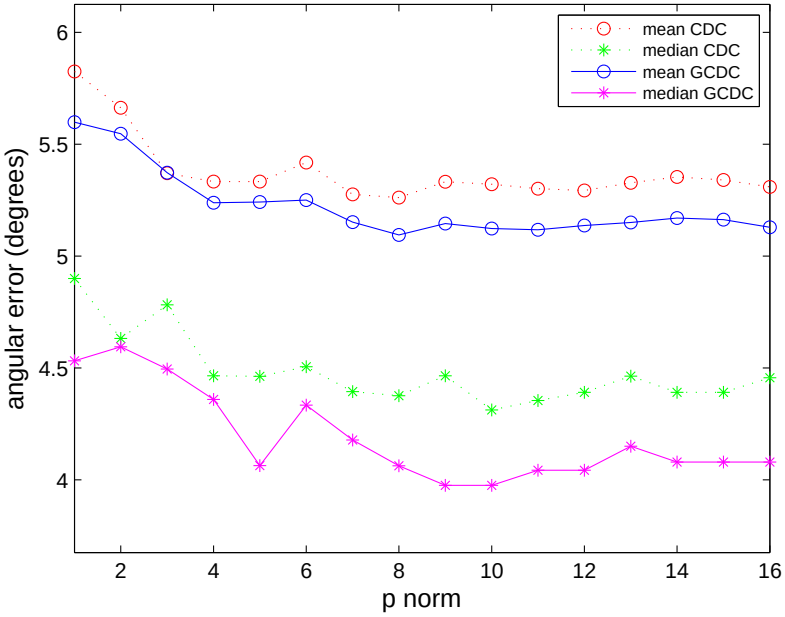

Figure 5.7: A comparison of algorithm performance on the Ciurea and Funt Grey Ball dataset [CF03b] for: CDC (red circle and green star lines represent the mean and median angular error respectively) and GCDC (mean and median angular error are represented by the blue circle and the magenta star lines respectively). For this experiment the 150 image subset proposed in [JGG07] is used.

gamut mapping.

Table 5.3: The perfomance of our algorithms on Funt and Shi's linear version [SF10] of Gehler data set [GRB$^+$08], and compared with other methods (* results presented in [GGJ11])

Algorithm	angular error (degrees) Mean	Median
Grey World	6.4	6.3
MaxRGB	7.6	5.7
1st order Grey Edge	5.3	4.5
2nd order Grey Edge	5.1	4.4
SoG, $p = 10$	5.7	4.6
SoGC, $p = 6$	5.5	4.5
CD, $p = 5$	5.2	4.2
CDC, $p = 5$	4.9	3.8
Pixel-based Gamut Mapping*	4.1	2.5
Edge-based Gamut Mapping*	6.7	5.5
GCDC, $p = 5$	**4.8**	**3.7**

Figure 5.8: Gehler dataset: cases of outdoor images where including GM to CDC helps

HDR dataset

For this dataset we compare the results of CDC with our new GCDC, presented in Figure 5.10. The results are again shown in terms of the mean and median angular error for norms 1 to 16. From the figure we can observe that by adding the constraint in the illumination in the form of GCDC results in a stable performance for illuminant estimation and helps, slightly, to improve the accuracy, even though it is less dramatic than that of the SFU set.

We are the first to admit the performance increment here is small. Like in Gehler's dataset the performance increment is small thought, in this case, more significant.

In Table 5.4 we compare results presented by Funt and Shi in this dataset for Grey World, with mean and median of 7.9 and 7.3, "MaxRGB filtering "mean and median

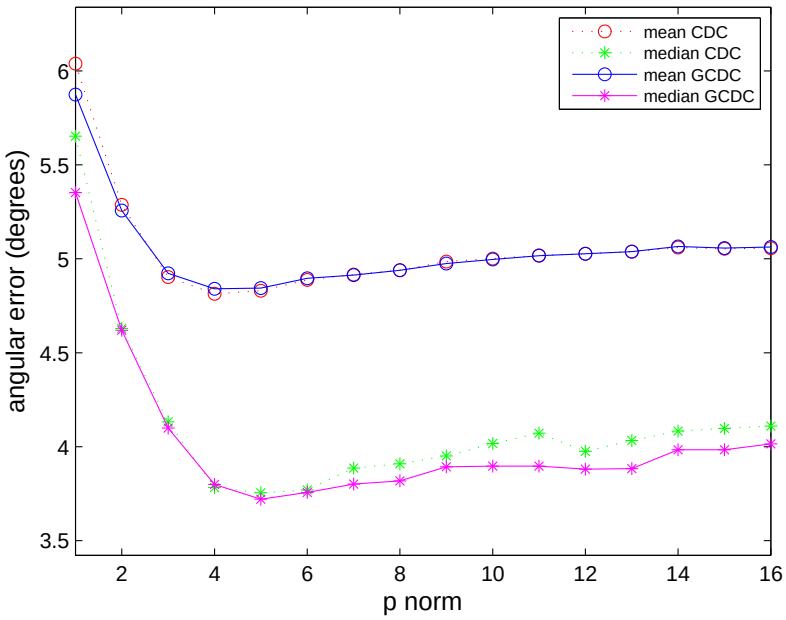

Figure 5.9: A comparison of algorithm performance on the Gehler dataset [GRB$^+$08] (linear version provided by Funt and Shi [SF10]) for CDC (mean and median angular error are represented by the red circle and green star lines respectively) and GCDC (blue circle and magenta star lines represent the mean and median angular errors respectively).

of 6.3, 4.3, "MaxRGB resizing "mean and median of 6.3 and 3.9, and the results of our *CDC* presented in Chapter 4 for norm 5, mean and median of 5.8 and 3.8, and compare them with our *GCDC* best results, for norm 6, with mean and median angular error of 5.6 and 3.2 respectively. From the table, we can observe once more that our *CDC*'s algorithms, original and GCDC versions, perform best overall over these statistical methods.

Finally, in Table 5.5 we resume the results for all five data sets used for the experiments in this chapter. We compare the median angular error for norms four, five and six, and compare them with the best angular error and their corresponding p norm. We can observe that for SFU A, Gehler and HDR data sets, the most accurate estimation is given by a norm in the range [4,6]. Furthermore, for SFU B and Grey Ball, even though the best result was not obtained with a norm in this range, the variation in angu-

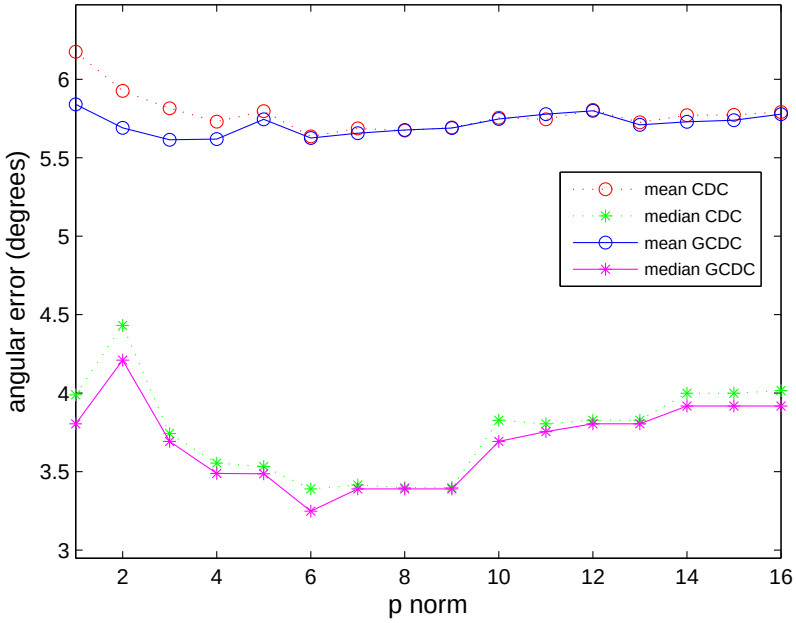

Figure 5.10: A comparison of algorithm performance on Funt and Shi's HDR dataset [FS10] for: CDC (the mean and median angular error are represented by the red circle and green star lines) and GCDC (blue circle and magenta star represent the mean and median angular error respectively).

Table 5.4: The performance of our algorithms on HDR Funt and Shi's Image data set [FS10] in terms of mean and median angular error.

	angular error (degrees)	
Algorithm	Mean	Median
Grey World	7.9	7.3
MaxRGB, filtering	6.3	4.3
MaxRGB, resizing	6.3	3.9
1st order Grey Edge	5.98	3.96
2nd order Grey Edge	5.91	4.0
SoG, $p = 5$	6.6	5.1
SoGC, $p = 5$	6.6	4.8
CD, $p = 5$	6	4.1
CDC, $p = 6$	5.7	3.4
GCDC, $p = 6$	**5.6**	**3.2**

lar error across the norms tested is not significant. These results support the argument that a norm in the range [4,6] will provide a good illuminant estimation for different image content. These results are encouraging when striving to find a general illuminant

estimator.

Table 5.5: Results of experiments for all five data sets using GCDC. The best norm result per data set is compared with the results for norms in the range [4,6].

Algorithm	Median angular error (in degrees)			Best	Median angular error
	norm 4	norm 5	norm 6	norm	(in degrees)
SFU A, v11	1.3	1.4	1.5	4	1.3
SFU A, v87	3.2	3.0	3.1	5	3.0
SFU B, v11	1.3	1.4	1.5	2	1.0
SFU B, v87	3	3.1	3.3	3	2.9
Grey Ball	4.3	4.0	4.3	9	3.9
Gehler	3.8	3.7	3.75	5	3.7
HDR	3.5	3.5	3.2	6	3.2

5.5 Conclusions and Future Work

In this Chapter we proposed a fusion of our *CDC* approach and Gamut Mapping into a new algorithm for colour constancy. We showed in the previous chapter that our *CDC* algorithm is a stable algorithm providing good results. Gamut Mapping on the other hand provides extra information about an image's illuminant and so it is incorporated here to extend the original *CDC* approach.

We do not use gamut mapping as such to estimate the illuminant scene but use it to determine if a light, from an illuminant set, is a candidate to be the illuminant scene of a test image. Thus we rule out lights that are not consistent with a test image, reducing the size of the illuminant set. We proved that by doing so we optimise *CDC* computations, since *CDC* computation time depends directly on the size of the illuminant set, and its performance is improved.

GCDC maintains the simplicity of a statistical algorithm, while exploiting a constraint in illumination inspired by Gamut Mapping principles. Experiments in synthetic and real images provide enough evidence of the algorithm good performance.

The new algorithm provides a general framework for illuminant estimation and so future work would naturally lead to include additional information such as results of

spatial filters. Moreover in terms of the illuminant set, future work could aim to build a general illuminant set. Aditionally, future work could also consider the case of images created using multiple illuminants such are the cases of HDR and Gehler datasets.

Chapter 6

The OSA Illuminant Spectrum Recovery Competition 2011

In this Chapter, a final evaluation of a *constrained* Minkowski algorithm is presented. We want to investigate how our approach, and the principles on which it is based, performs in an experiment relevant to what is known about how the human visual system responds to light. Specifically, we evaluate our approach when applied to a problem of illuminant spectrum recovery proposed by the organisers of the Optical Society of America, *OSA 2011 Fall Vision Meeting*.

In this Chapter, we present the context of the illuminant spectrum estimation problem. We also show how computational illuminant estimation (in the form of our algorithm) can be adapted to this human vision experiment. We begin by recalling that all our algorithm development assumes a diagonal model of illumination change. The "cone response" images used in the OSA experiment are known not to adhere to the diagonal model. Rather, our idea is that via a process called *Spectral Sharpening* [FDF94b] [FS00a] [FS00b], cone responses are transformed to a new "sharp" basis where the diagonal model does hold. So, we include spectral sharpening and we also describe its principles and how we include it on the spectrum recovery task. The OSA

experiment placed over 30 algorithms in competition with each other. Each algorithm's performance is given a single score (described later). Here we show that if care is taken to use the correct sensor basis, then a simple constrained Minkowski norm approach outperforms all other algorithms.

6.1 The competition

The OSA's Fall Vision Meeting is an annual meeting designed to motivate scientific discussion among researchers on key aspects in vision science, such as, the visual system, colour vision, perception, application areas and clinical vision. During the Fall Vision Meeting in 2011, David H. Brainard and Alex R. Wade presented the *Spectrum Recovery Competition 2011*. The aim of the competition was to estimate (using computational algorithms) the illuminant spectra for a set of 10 cone response images. That is, an algorithm will have the set of 10 test images as input, compute a series of operations and finally provide 10 illuminant spectra (one for each scene). Shortly after the competition closed on July 2010, all details about it (instructions, rules), the data, the final participant's scores, and the correct answers were made available on line [BW11] for further reference. This provides an excellent opportunity to test future (and our) algorithms in the context of illuminant spectrum recovery.

The competition ran for a period of several weeks. Every week the participants were allowed to submit their illuminant estimations and a score for these submissions was provided. The score was computed as the sum of the squared differences between the relative spectrum of the illuminant spectrum estimations and the actual illuminant spectra. The participants could then revise their estimations (tweak their algorithms) to submit a new set of guesses the following week and all participants could see each other scores.

There were no restrictions regarding the type of algorithm that could be used and

so the participants could hold to any assumptions they would find relevant for the available data. Moreover, the weekly feedback provided a continuous evaluation of one's algorithm.

The available data consisted of:

The test images. 10 "cone generated" images (rendered from 10 hyperspectral images) from which 10 spectrum estimates of their illuminant need to be recovered. The images, "rendered"for display, are shown in Figure 6.1.

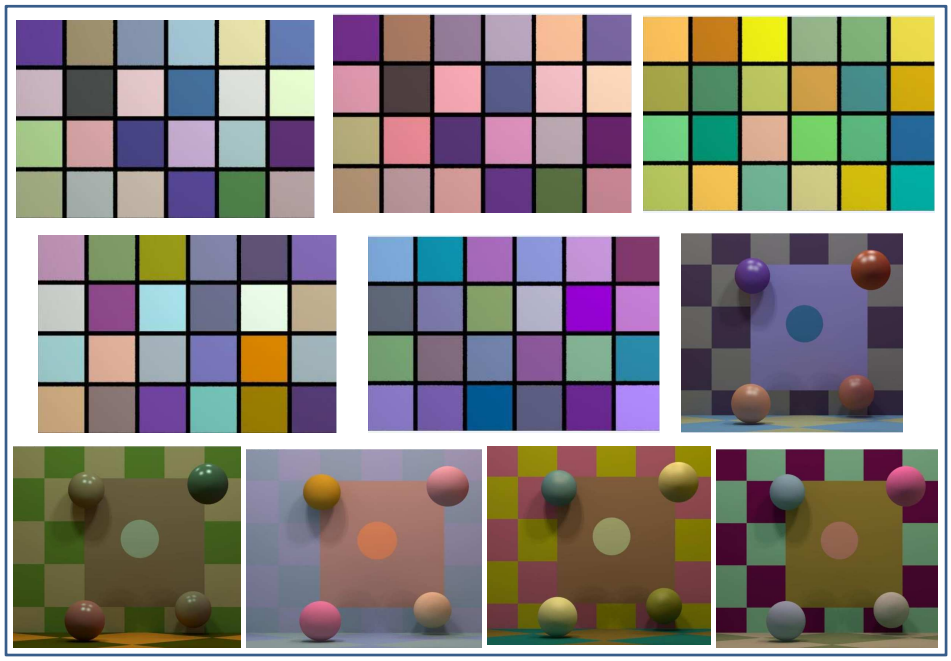

Figure 6.1: Hyperspectral images for the Spectrum Recovery Competition, as part of the OSA's Fall Vision Meeting 2011. The competition was organised by David H. Brainard and Alex R. Wade [BW11]

Three illuminant basis functions. Every scene's illuminant was created as a weighted sum of the three illuminant basis functions (presented in Figure 6.2). The 10 SPDs of

the scene's illuminants are shown in Figure 6.3.

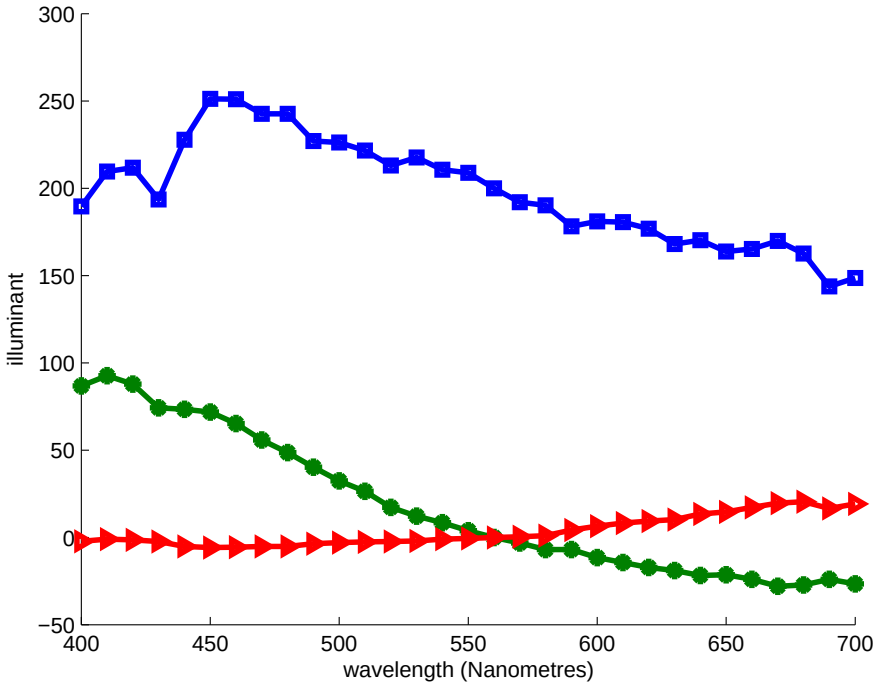

Figure 6.2: The three illuminant basis functions provided for the competition. The spectrum power distribution of each illuminant corresponds to the range from 400 nm to 700 nm sampled every 10 nm.

Three reflectance basis functions. The scenes created for the experiments contain objects with both diffuse and specular reflectances. The diffuse components are created as a weighted sum of three reflectance basis functions (Figure 6.4).

Cone Sensitivities. The **LMS** cone sensitivities are also provided. This functions are sampled every 10 Nanometres from a range of 400 to 700 Nanometres. The cone fundamentals are shown in Figure 6.5.

Calibration Image. Considering that some calibrated algorithms need to compute some sort of training (of image data over known conditions), a scene under a reference

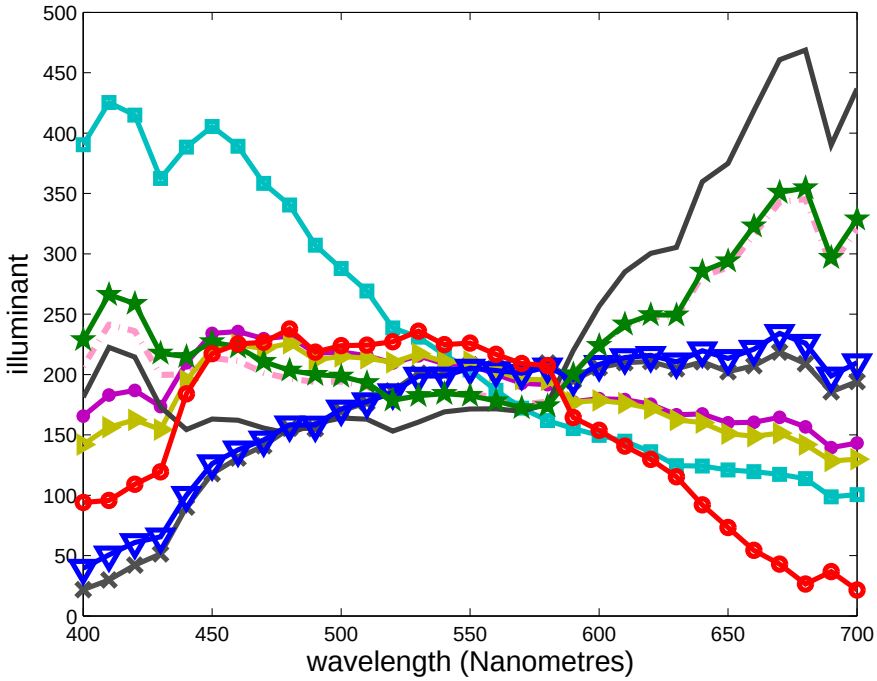

Figure 6.3: The 10 Illuminant SPD for the illuminant spectrum recovery competition

light and the SPD of the light was also made available (Figure 6.6).

Sample code. A sample code was provided as example of how to manipulate the data and compute the spectrum estimation for a particular scene. Moreover, the score for *Grey World* was given as reference (clearly we wish to do better than *Grey World*) which achieved a score of 3544.88 (details are in [BW11], the precise meaning of this "score" is not important but the idea is that if we run an algorithm with a score 1750, this tell us the algorithm is "twice as good" as Grey World). Remarkably by the end of the competition *Grey World* came on 13th place out of 38 participants. That Grey World did so well offers a salutary lesson for the community: simple statistical algorithms generally offer reasonable performance and are not fragile. We presume that many of the 25 algorithms which performed worse were adopting "advance reasoning".

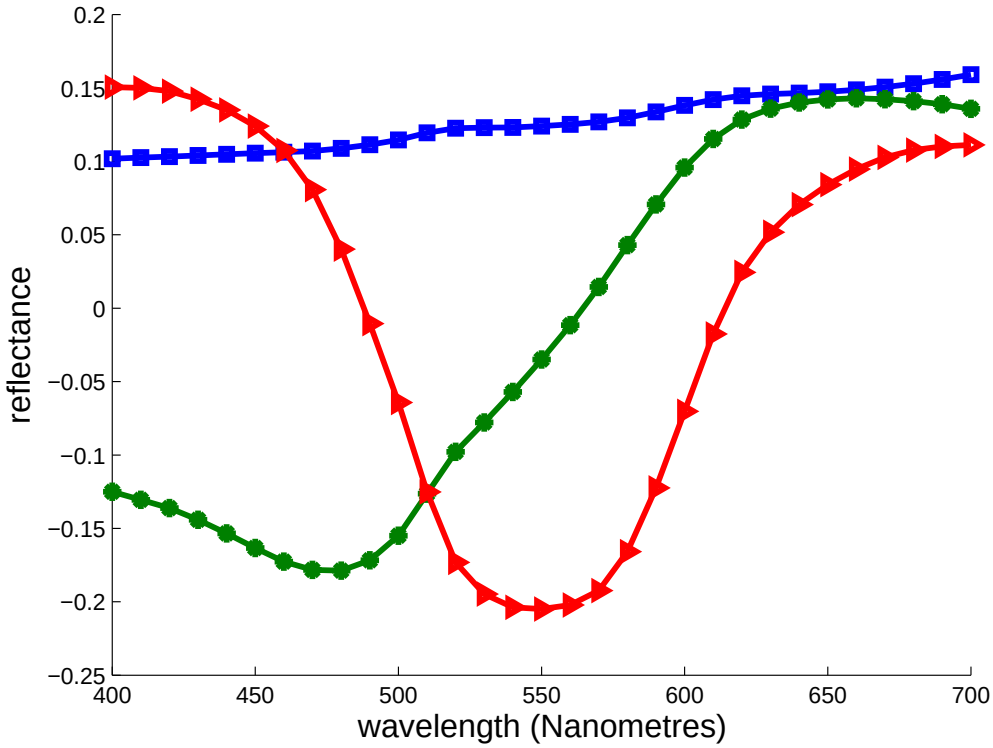

Figure 6.4: The three surface reflectance basis functions. The figure presents the spectral power distribution for each basis function, sampled every 10 nm, in the range from 400 to 700 nm.

Other approaches (participants that came second and fifth position) used to compute the illuminant spectra in this competition included the use of Bayesian algorithms (see Section 2.2.1), placing constraints over surface reflectances and illuminant spectra (used by the algorithm that achieved the fourth place) and exploiting information from specular highlights present in the scenes (see Section 2.4).

The final scores for the participants in the competition are presented in Table 6.1. The winning approach, *the leopards*, obtained a score of 597.92, the *ales*, obtain a score of 1224.6, the third place achieved a score of 1299.71, and fourth and fifth places obtained scores of 1331.71 and 1635.85 respectively. It is worth commenting about the leading solution. It turns out by the nature of the experiment, i.e., where estimates are submitted

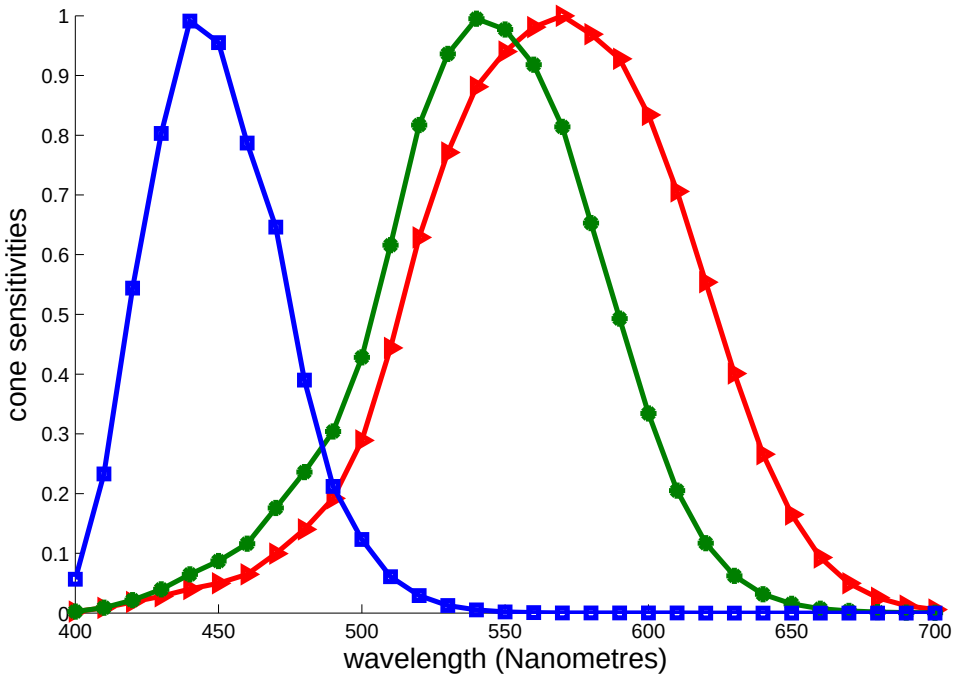

Figure 6.5: The LMS cone fundamentals for the spectrum range from 400 to 700 nm, sampled every 10 nm

to the organisers on a weekly basis and the scores returned, that by this paradigm one can adopt a testing procedure where you can discover the exact answer for all images. That the leopards score 592 (and not 0) is only because they could not get enough feedback to discover all the lights exactly. The best conventional (and implementable) illuminant estimation algorithm is, thus, the one at the second position in the table.

6.2 Computational Illuminant Estimation for a problem of Human Vision

In Chapter 4 we have presented an new constrained approach for colour constancy which proved to deliver reasonable estimation of an scene's illuminant. Furthermore, the results of the experiments showed that our algorithm's performance is stable when

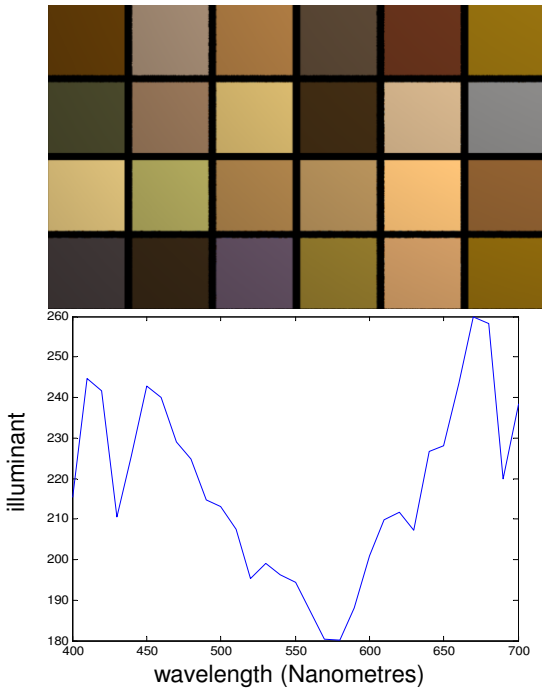

Figure 6.6: The calibration image for the competition is presented in the top side, and its corresponding illuminant's spectral power distribution is presented in the bottom panel. Such calibration image and corresponding light could be used for calibrated algorithms

evaluated in different image data sets. We believe that at this point, a natural next step is to investigate whether (or not) our algorithm can be used for a practical application. Specifically, in the context of the illuminant spectrum recovery task, an experiment for human vision. Thus, in this Chapter we want to investigate if our *constrained* framework for illuminant estimation can be applied so that we compute the spectrum of the illuminants for the test images.

Our constrained algorithm belongs to the same category (statistical algorithms) as the Grey World algorithm. However, it has proved to provide a more accurate estimation and hence, we would expect our algorithm to beat the Grey World's score of 3544.88.

Table 6.1: OSA Spectrum Recovery Competition 2011: The final Scores. The wining team was *the leopards* for a set of 38 participants, with a final score of 597.92. The Grey World algorithm occupies the 13th position at the end of the competition.

	Algorithm	Score
1	the leopards	597.92
2	ales	1224.95
3	abe	1299.71
4	pinkStinks	1331.71
5	alaf	1635.85
6	jili	1702.12
7	simbian	1792.64
8	uciduo	2048.92
9	hamed	2750.31
10	mirch	3163.52
11	oaksey	3322.71
12	yichang	3532.46
13	Grey World	3544.88
...	...	...
38	IliaLebedev	18656.95

6.2.1 Spectral Sharpening

The principles of the illuminant estimation algorithms that we have presented in this thesis are based on a diagonal model of illumination change, which has been proved not to always hold when using in human vision. A disadvantage of the diagonal model is that the cone sensitivities have broad functions and the model do not perform well in these cases. However, it has been shown that by using spectral sharpening and working with narrower-band cone sensitivities improves the accuracy of illuminant estimation [FDF94b] [Fin95].

Narrow-band sensors are those that are sensitive to just one wavelength. Spectral sharpening methods find linear combinations of broad sensors (for example the cone fundamentals) that are maximally sensitive in a given wavelength interval $[\lambda_1, \lambda_2]$, and which are narrower. Using sharp sensors helps algorithms for colour constancy (based on the diagonal model of illumination change) to achieve better estimation than using

broad sensors [FDF94b].

Finlayson and Susstrunk showed that **The Sharp Adaptation Transformation**, which was derived by using a sensor-based spectral sharpening method, is a good spectral sharpening transformation [FS00a] and [FS00b]. Not only does it provide more narrow sensors. It also models psychophysical corresponding colour data. We incorporate the Sharp Adaptation Transform to derive sharp cones for our experiments. Numerically, this transform is represented by:

$$Msharp = \begin{bmatrix} 1.2694 & -0.0988 & -0.1706 \\ -0.8364 & 1.8006 & 0.0357 \\ 0.0297 & -0.0315 & 1.0018 \end{bmatrix} \tag{6.1}$$

Note that when refer to *Sharp*, we refer to the sensors illustrated in Figure 6.7, which are derived by multiplying the CIE XYZ matching function with Equation 6.1 (which themselves are a linear combination of the cones). We can observe from the Figure 6.7 that cones are actually narrower but there is still some overlapping.

6.2.2 Illuminant set

An illuminant set needs to be define prior running our *constrained* algorithm. We selected to test our algorithm with two illuminant sets. We use the 99 Granada University's measurements of day light [HRNL01]. To a second set we add the actual correct answers published with the results to the competition.

The Granada illuminant set. This dataset contains 99 SPD measurements of day light at different times of the day in Granada, Spain. Results using this dataset set are labelled as *v99*. When including the real answers we our algorithm as *v99+10*.

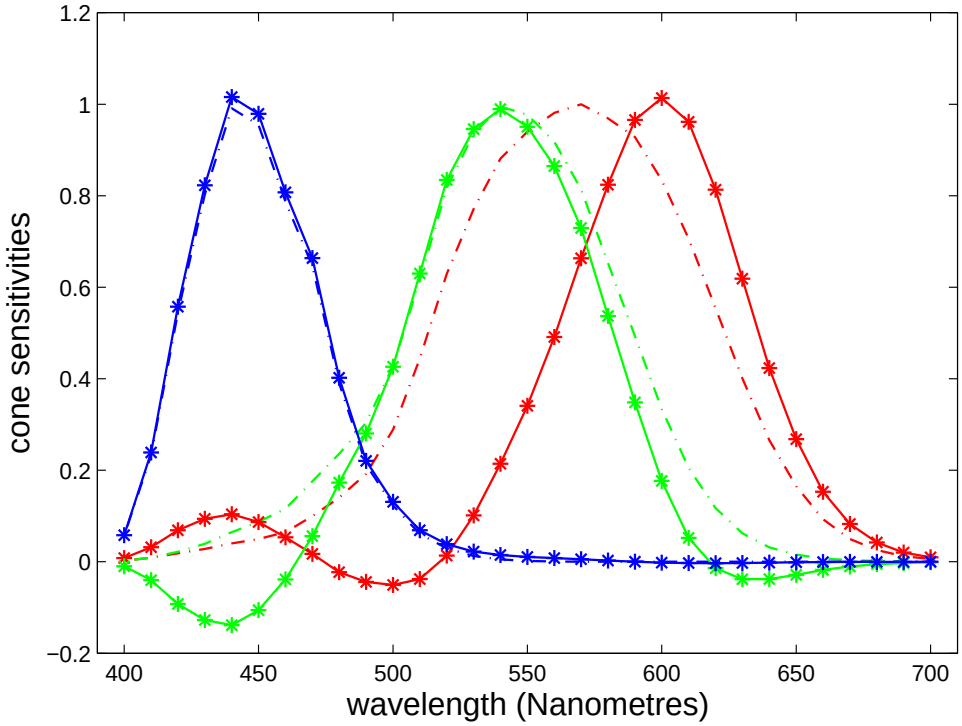

Figure 6.7: The LMS cone fundamentals (discontinuous lines) are presented. The LMS sharp cones (star continuous lines) transformed by Equation 6.1 are also included.

6.3 Experiments and Results

In this section we present the results of experiments using **SoG**, **SoGC**, **CD** and **CDC** for spectrum recovery (i.e. simple Shades of Gray with a constraint on the illuminant, the Minkowski norm approach working on derivatives and constrained version thereof). All these algorithms return an RGB estimate of the light. As part of the competition the standard method for mapping this RGB to spectral power distribution is supplied as a Matlab function. Thus, we simply use the same competition code to turn our RGB estimates into spectra. The test images and their corresponding illuminant spectrum are presented in Figure 6.8.

We first want to investigate if by including spectral sharpening will help in this context. Thus, we first test **SoG** using the cone fundamentals and with sharp cone sens-

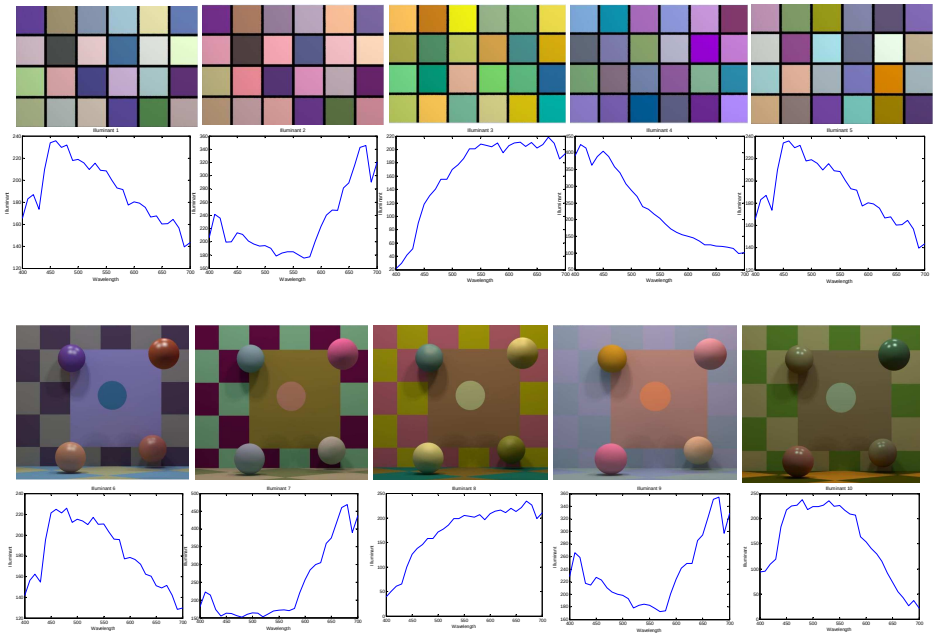

Figure 6.8: OSA 2011 Illuminant Spectrum Recovery Competition: the ten test images and their corresponding illuminant SPDs

itivities to compare their respective performance. The results for this test is presented in Figure 6.9, where we present the score obtain for **SoG** for norms from one to 32. The blue continuous line represents the performance of applying **SoG** over cone sensitivities and the star magenta line represents the performance when using sharp cone sensitivities.

We can observe from the figure that a remarkable improvement is achieved when using sharp cone sensitivities from norm 2 and onwards. Furthermore, large norms, from 15 and onwards, the score remains constant, much more constant than the performance of the cone sensitivities.

It has been shown in previous works [BMCF02] [FS10] that by applying a preprocessing step illuminant estimation can be improved. Thus, in our next experiment we

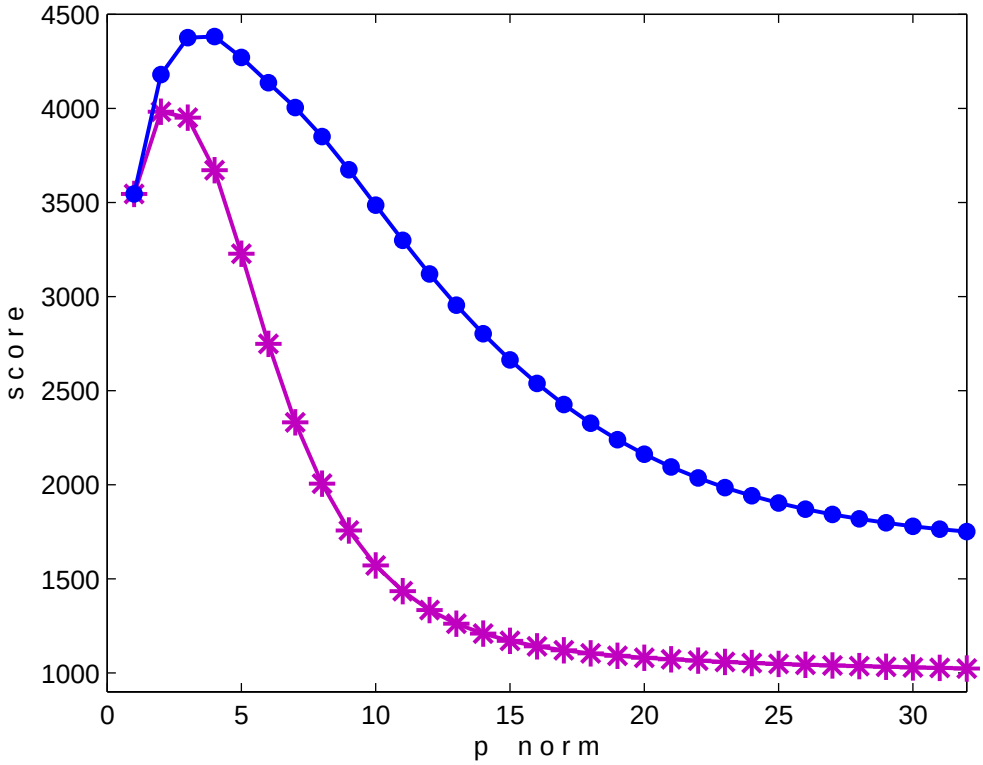

Figure 6.9: A comparison of **SoG** performance using cone sensitivities (the circle blue line) and sharp cones (the star magenta line).

incorporate a pre-processing step to **SoG** by applying a Gaussian smoothing operation before the estimation. In our experiments we use a Gaussian filter with different σ values varying from 1 to 10 pixels. We found that the performance for **SoG** for different σ values varies having the extreme errors obtained for $\sigma = 1$ and $\sigma = 10$. In Figure 6.10 we present these results. The figure shows the performance of **SoG** for three cases: with no pre-processing step, for a smoothing operation with $\sigma = 1$ and with $\sigma = 10$. This experiment was first done using cone sensitivities.

This same previous experiment for different σ values for the Gaussian filter was also done using sharp cone sensitivities. As in the previous case, the performance for **SoG** with $\sigma = 1$ provided the smallest error while case with $\sigma = 10$ provided the

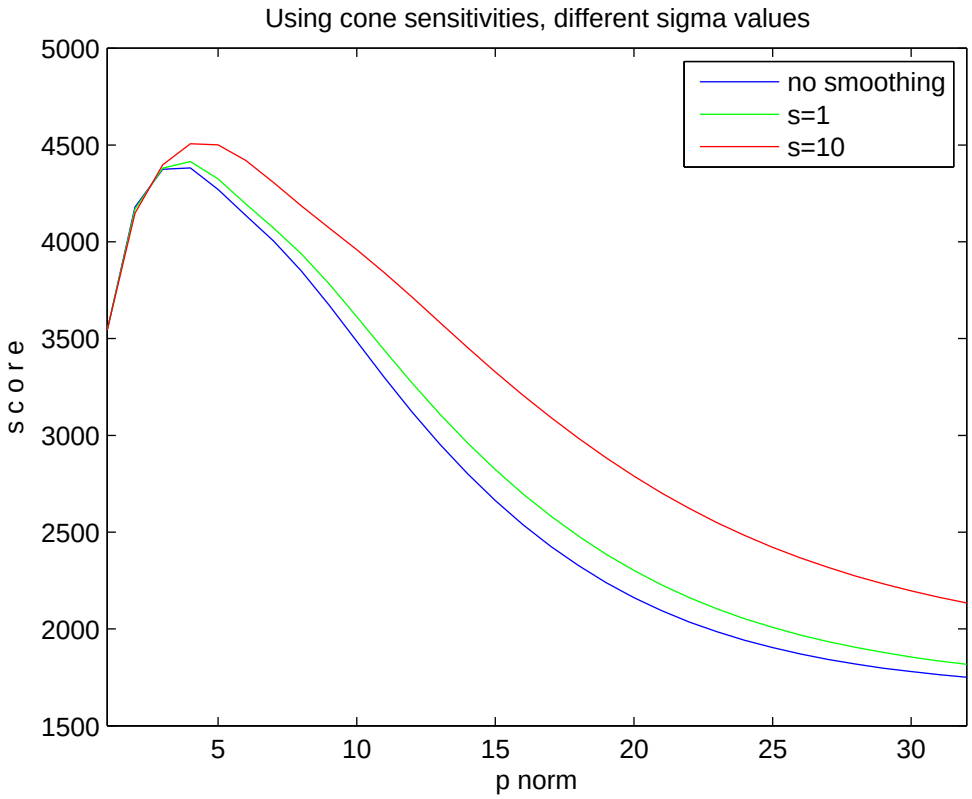

Figure 6.10: Using cone sensitivities. A comparison of **SoG** performance with a pre-processing smoothing operation: $\sigma = 1$ (the green line) and $\sigma = 10$ (the red line). The results for **SoG** with no pre-processing are included for reference (the blue line).

larger errors for values from one to 10. In Figure 6.11 we present the results for this experiment.

It can be observed from the figure that the smoothing effect on the score is more significant when using sharp cone sensitivities, specially for large norms. Even though it helps, specially for larger norm, we can see from the figure that for norms 22 and larger the smoothing operation (with $\sigma = 1$) helps to improve **SoG** performance. In the case of using cone sensitivities this is not the case since the best results over the norms is always given for the case where no pre-processing was applied.

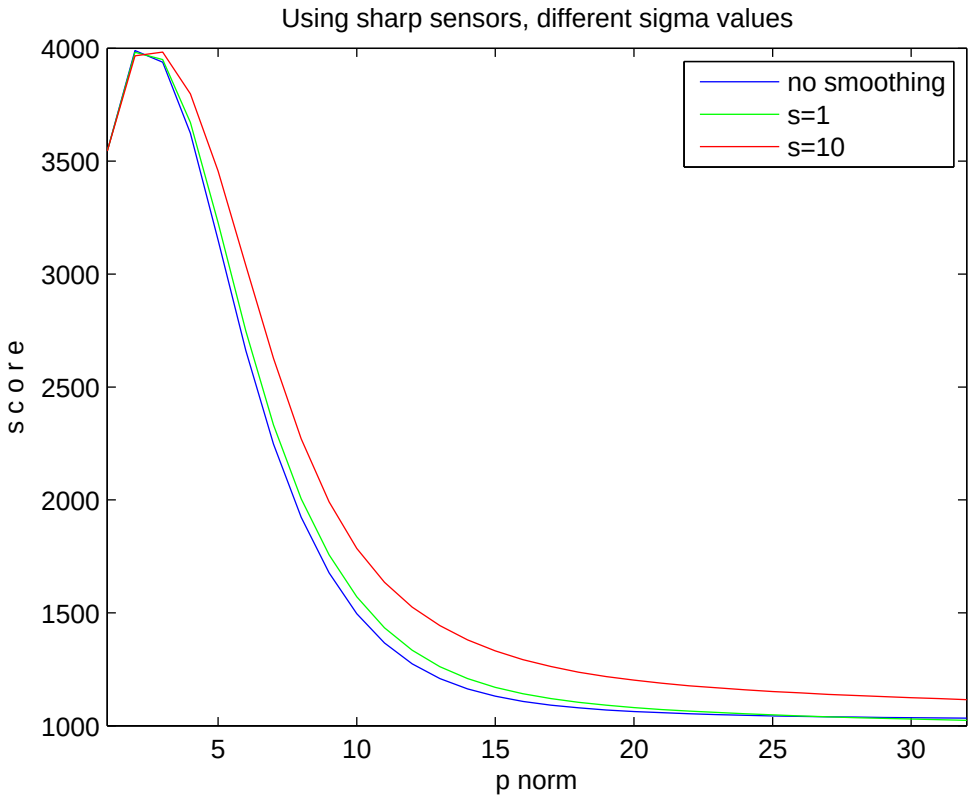

Figure 6.11: Using sharp cones. A comparison of **SoG** performance with a pre-processing smoothing operation: $\sigma = 1$ (the green line) and $\sigma = 10$ (the red line). The results for **SoG** with no pre-processing are included for reference (the blue line).

We then want to investigate if including derivatives in the image helps for the estimation. In Figure 6.12 we summarise the performance for **Shades of Gray** using derivatives and the performance of corresponding variants where the light is constrained. For this illuminant constraint we use the Granada data set. Immediately, we can conclude that a constraint on an illuminant is powerful. Yet, equally apparent is the fact that the derivative-based approach is not working. We propose the explanation of this is simple. Looking back to the images in Figure 6.1, we see that in the colour chart images, the individual patches are surrounded by black. Moreover, there is not sufficient edge information in the rest of the images either. This makes **all** the edge information calculated in these images unreliable. Yet, **Shades of Gray Constrained** (on RGB col-

our distributions) works well. However, the best result is only slightly better than the unconstrained version. This tell us simply that **Shades of Gray** estimation is always close to one of the Granada Lights.

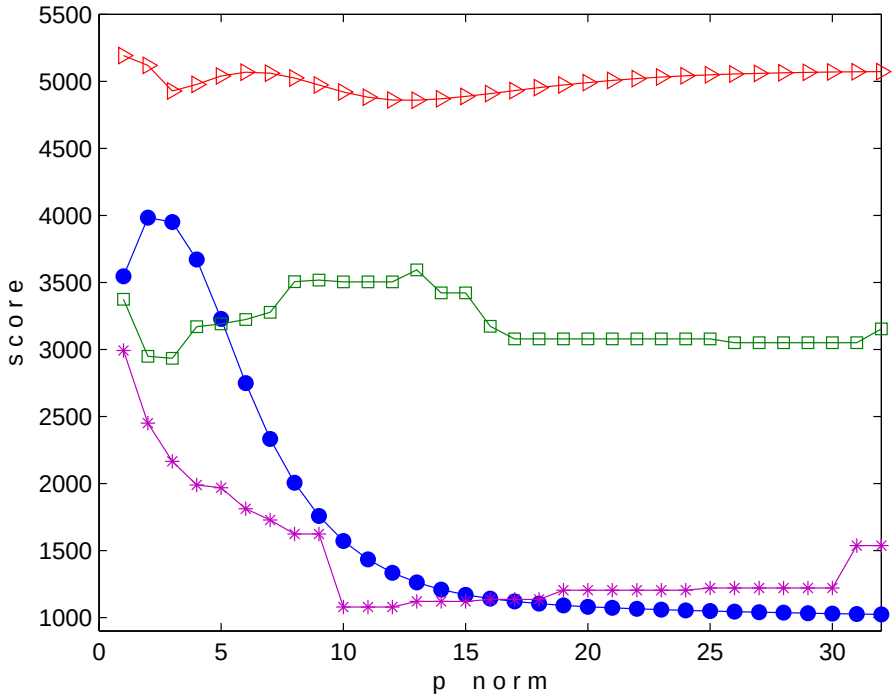

Figure 6.12: The results for sharp cone sensitivities using the Granada Daylights for SoG (circle blue line), SoGC (star magenta line), CD (triangle red line) and CDC (square green line).

In Figure 6.13 the results for sharp cone sensitivities are presented for norms from one to thirty two for the second data set, the Granada Daylights and the actual Daylights used in the experiment. The best results achieved in this experiment are: **SoG** with a score of 1023.6, **SoGC** with a score of 297. Here we awe the adding a constraint on the illuminant does make a significant difference. Another conclusion is that the Daylights used in the OSA experiment are not representative of typical everyday light.

Finally, in Table 6.3 we summarise and present the best scores for the algorithms tested in this experiment: for **SoG**, the best scores were 1033 and 1023.6 with a smooth-

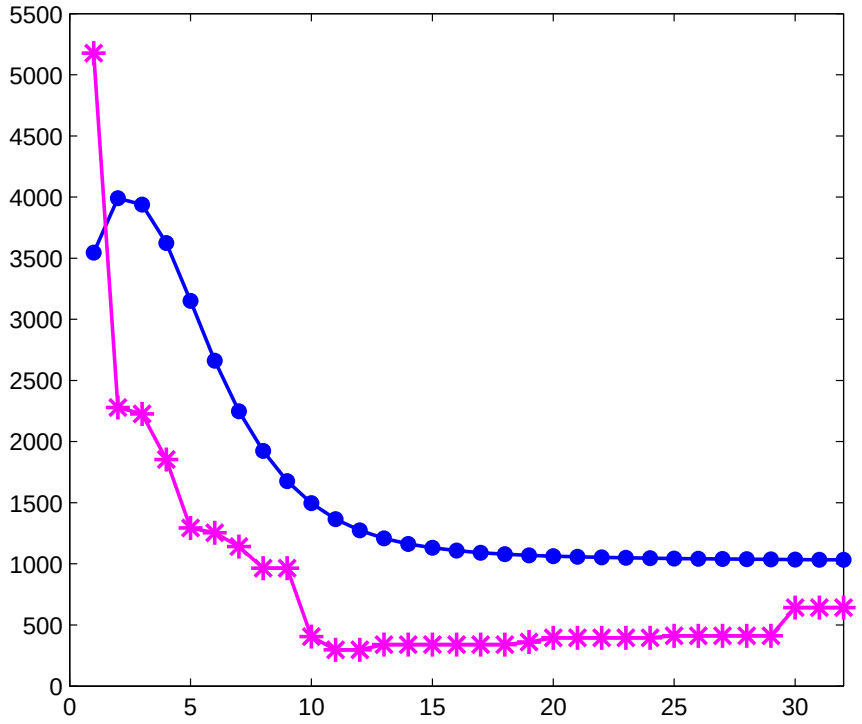

Figure 6.13: The results for sharp cone sensitivities using the Granada Daylights and the actual Daylights used in the experiments for SoG (circle blue line) and SoGC (star magenta line)

ing using $\sigma = 1$, both cases for $p = 32$, while for **SoGC**, using the Granada Daylight data set, the best score is 1078.8 for $p = 11$. Again, we can see how the constraint helps to improve the illuminant spectrum estimation. Furthermore, for **SoGC** the results are comparable with those of the algorithms participating in the competition, even the results of the winning algorithm.

6.4 Conclusions

We provided the details of the implementation of our *constrained* approach for spectrum recovery in a human vision test. We show that even when the nature of the test images

Table 6.2: OSA Spectrum Recovery Competition 2011: The final scores and results using the Granada Daylight data set.

Algorithm	Score
the leopards (1st place)	597.92
ales (2nd place)	1224.95
abe (3rd place)	1299.71
pinkStinks (4th place)	1331.71
Grey World (13th place)	3544.88
SoG, $p = 32$	1033
SoG, $p = 32$, $\sigma = 1$	1023.6
SoGC, $p = 11$, v99+10	297

there is not enough information for our *Constrained Minkowski* approach to outperform other algorithms, by adding the constraint in illumination improves performance in this context.

Furthermore we showed that by implementing spectral sharpening is essential for our algorithm to deliver the best performance. Our final results outperform the best results for algorithms that participated in the experiment, the winning algorithm included.

Chapter 7

Conclusions

The principal aim of this thesis is to propose an efficient statistical algorithm for illuminant estimation. First, we presented a review of the state of the art for illuminant estimation algorithms, highlighting the relevance and potential of the framework based on the Minkowski Family Norms. We then presented our new algorithm and experiments on different image datasets. The results showed that the new approach delivered good illuminant estimation, most of the time outperforming the state of the art algorithms and sometimes achieving as good estimation as much more complex algorithms.

In this Chapter we summarise the main contributions of this thesis together with a discussion of the possible applications of the work.

7.1 Contributions

Adopting a Reformulation of a constrained approach based on the Minkowski Norm Framework for Illuminant Estimation. An illuminant estimation based on Minkowski Norms uses the assumption that basic statistics in an image can be used to infer the illuminant of the scene. It has been shown that this framework integrates the basic and well known statistical algorithms, the **Grey World** and the **MaxRGB**,

proving that they are norm based approaches [FT04]. From the category of statistical approaches, the Minkowski-based algorithm is the most general formulation and remarkably it achieves the best performance. In Chapter 3 we revisited this approach and modified the original error optimisation for illuminant constraint based on the Minkwoski Family Norm framework [FT05]. In the approach by Finlayson and Trezzi, an illuminant set is created a priori and then a criterion is defined to select the best illuminant from the set. This criterion establishes the consistency of an illuminant with regard to a test image. Finlayson and Trezzi calculate this consistency as an 'error measurement' using a distance error in the form of a Minkwoski norm. This error is calculated for every light in the set and the light with the smallest error is selected.

We found that this 'error measurement' was biased by either lights with very large RGB values or by achromatic lights. We considered these situations and reformulated the error expression as follows: a test image is normalised by an illuminant ('dividing out' the illuminant), so that a 'white surface' will be mapped to the vector [1 1 1]. The error measurement will then add the deviations from this vector [1 1 1] for all pixel vectors. Furthermore, a change in intensity was also taken into account in the error formalism by adding an intensity smoothing factor. We carried out experiments using one of the standard image datasets (SFU A) to evaluate our new constrained approach. The results showed that our new formalism is more accurate than the previous approach.

Improving Computational Efficiency One of the key aspects of an illuminant estimation algorithm is their performance in terms of computational efficiency. This is highly relevant for applications where an illuminant estimation needs to be calculated in real time, for example, in tasks such as white balance. In Chapter 3 we show how our method can be implemented to run much faster than our previous constrained Minkowski formalism.

The time to estimate an illuminant for our constrained approach is directly affected

by the calculation of the error measurements since an error distance is calculated for every pixel in the image. This means that the computation increases based on the number of pixels in the image. Thus, to reduce the calculation time, we proposed the use of histograms to describe the colour distribution in an image and then calculate the error for every illuminant. In Chapter 3 we run our algorithm using both RGB pixel values and histograms. We found that the simple strategy of using histograms significantly affects the processing time required to compute the distance measurements. The total time for illuminant estimation was reduced dramatically, in some cases the processing time was reduced to more than 95 percent. The simple insight of using quantisation and frequency histograms is essential for the Minkowski Norm approach to be used in practice.

Combined Derivative Constrained Approach It has been shown that edges in an image can provide extra information about the light illuminating a scene. Edges in an image can be found using derivatives (typically first and second order). Gijsenij et al. [JGG07], showed that, for some image content, derivatives of first order achieved good estimation while second order derivatives achieved good estimation for a different image content. Moreover, applying a Gaussian filter with a smoothing factor (pixel standard deviation) can improve the results. Thus finding a good estimation for specific image content is directly affected by the combination of derivative order and smoothing factor. Indeed, the previous work on edge-based constancy which purports to deliver leading estimation does so by tuning the order and smoothing used to the data set.

We believe that an ideal colour constancy algorithm would be able to obtain illumination -invariant colour images despite the sort of scene or content that the image contains (the "tuning"required in previous research is a weakness of that approach). In Chapter 4 we proposed a new edge-based algorithm for illuminant estimation, the *Combined Derivative*, *CD*, approach. Our new *CD* algorithm includes x- and y- de-

rivatives and a Laplacian operator over images, providing a fixed derivative approach for illuminant estimation. We take into consideration first and second order derivatives simultaneously to exploit the use of differences in images avoiding tuning our algorithm to image content. Lastly, we included an illuminant constraint (which was defined in Chapter 4), to arrive at a new algorithm, the *Combined Derivative Constrained*, *CDC*, illuminant estimation.

We carried out experiments to evaluate our new *CDC* approach using five different datasets for colour constancy algorithms. Results were shown that our *CDC*'s performance is stable across all of them, moreover, we did not need to tune the algorithm per dataset. *CDC*'s results outperformed other statistical algorithms, in some cases achieving almost as good results as more complex approaches (i.e., approaches which need much more sophisticated processing compared with the simple statistical algorithms which are the focus of this thesis).

Gamut Combined Derivative Approach. Gamut Mapping algorithm is one of the most important algorithms for colour constancy to date and so we presented a detailed description of state of the art for Gamut Mapping in Chapter 1. The principle of Gamut mapping assumes that the observed colours in an image are constrained by the lighting conditions in the scene. The reddest red colour cannot occur under the bluest light. Gamut mapping works by finding a mapping that takes image colour within the set, or gamut, of typical surfaces viewed under a fixed reference light. We believe important when trying to solve the problem of illuminant estimation in general and so we proposed to include it in the form of a 'gamut' constraint to our CDC algorithm. We proposed to use this constraint to determine what lights, from an illuminant set, are consistent with a test image. This reduces the set of plausible lights for which *CDC* tests. Thus, the fusion of these two algorithms resulting in our new approach, the *Gamut Combined Derivative Constrained*, *GCDC*. This new algorithm is presented in Chapter 5.

Our new approach benefits from a dividing the illuminant estimation in two steps: first, it selects the feasible lights for a given test image (discard lights that are not consistent with the image). Second, our *CDC* approach selects the best illuminant for the image. By adding the the 'pre-selection' step, we are reducing the number of elements to consider for *CDC*, and so in consequence we are reducing the number of calculations carried out select the illuminant.

We tested our new algorithm and compared our results with our previous *CDC*'s results and found that performance is improved. For the SFU datasets (A and B), where images where created under laboratory conditions and we can ensure that the diagonal model for illumination change holds, the improvement is remarkable. For the rest of the datasets improvement is modest, though often significant. We conclude that: a) the 'gamut' constraint helps, b) the performance of our new *GCDC* algorithm is stable across datasets, and c) that *GCDC* remains the simplicity of being a statistical algorithm, outperforming other statistical state of the art algorithms

An experiment in the context of human vision Finally, we wanted to investigate if our computational illuminant algorithm could be applied more generally. For this aim we tested our *constrained* algorithm in the context of human vision. We carried out experiments for the OSA Illuminant Spectrum Recovery Competition. In Chapter 6 we give the details of the experiment and the results obtained.

We show that by using our *constrained* approach on **spectral sharpened** cone sensitivities achieves good illuminant spectrum recovery in terms of this Human Vision Test. The results of our *constrained* algorithm outperforms other computational algorithms (if our algorithm had entered the competition it would have been in first place, out of 32 entries!).

7.2 Future Work

Possible applications from this work include the general framework provided by the constrained approach adopted in Chapter 4 and 5 (*CDC* and *GCDC*). We indeed proposed the use of derivatives in images (first and second order) into a Minkowski framework for illuminant estimation. However, our constrained approach is not limited to use only such statistics and higher order statistics such as DCT or wavelengths could be also included.

In the field of digital photography, most digital cameras have means to carry out colour correction based on the type of scene illumination, where automatic illuminant selection could benefit the illuminant estimator presented in this work. Furthermore, the UEA colour group has close relationships with the industry and another future work would involve the commercialisation of the algorithms presented in this thesis into real processing pipelines.

Appendix A

Future work: Using brightest pixels

In this appendix we report on some recent work we carried out with colleagues elsewhere which, in effect, encompass our first ideas of how to extend the work presented in this thesis. Implicit in the Minkowski norm idea that "bright is right" and make it even more important to our formalism. Recently, Funt and Shi [FS10] showed that the MaxRGB algorithm worked better on significantly downsampled thumbnail images (it is advantageous to smooth the input). Rather than use a maximum we wish to test a Minkowski norm operating on the brightest pixels in the thumbnail. And, then we repeat the experiment where a thumbnail is not used. We considered the following cases:

Case 1. Using RGB values a 64×64 resized image

Case 2. Using the 5% brightest pixels of a 64×64 resized image. That is, 204 pixels out of 4096 pixels.

Case 3. Using the 5% brightest pixels in an image (full size).

For this experiment, we use the SFU calibrated image set and Funt and Shi's HDR dataset. The results are presented in Figure A.1 and in Figure A.2. And, as this is on-

going (and so-future) work we have only tested the simple Shades of Gray formalism.

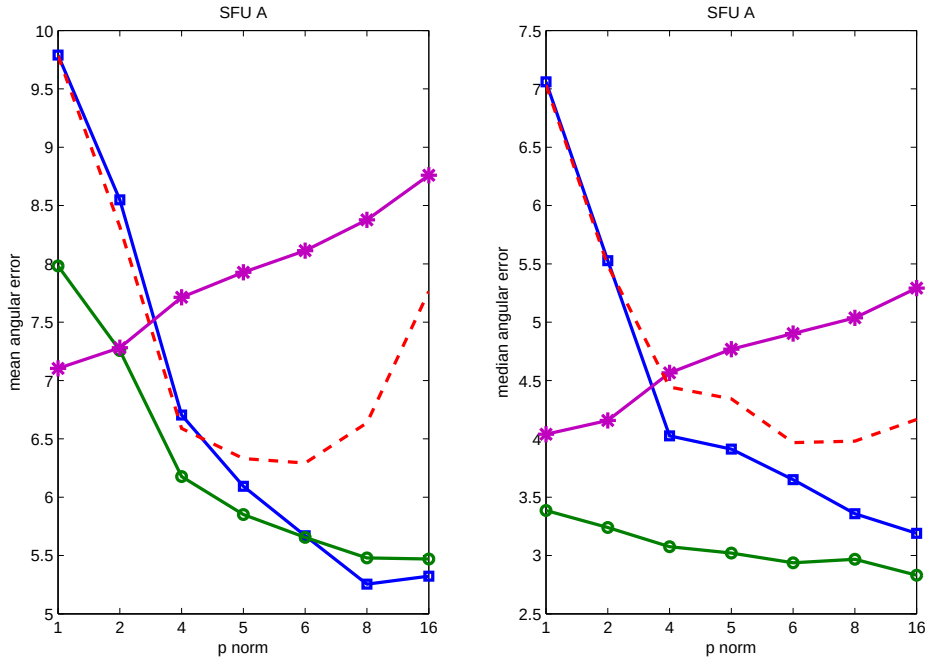

Figure A.1: Left and right panels show mean and median angular error for the SFU A data set for 4 varieties of Shades of Gray. Blue line: thumbnails Shades of Gray. Green line: 5% of thumbnail. Magenta shows performance using top 5% of the original (not downsampled) image. Red line is the normal Shades of Gray.

Our experimental results are shown in Figure A.1. The best performing algorithm (green line) is for Shades of Gray operating on the 5% brightest pixels of the thumbnail. This seems to validate Funt and Shi's [FS10] insight (that huge amounts of smoothing helps). Next best performance is for Shades of Gray operating on the whole thumbnail (blue line). Interestingly, the brightest pixel idea does not work on the full resolution image (magenta). Here, we hypothesise that this is because we are, more or less, making the Shades of Gray operator more like MaxRGB which we know is not a robust statistical estimator. For reference, the red line shows normal Shades of Gray (for full resolution images).

Figure A.2 show the same experiment for the HDR image data set where the data in

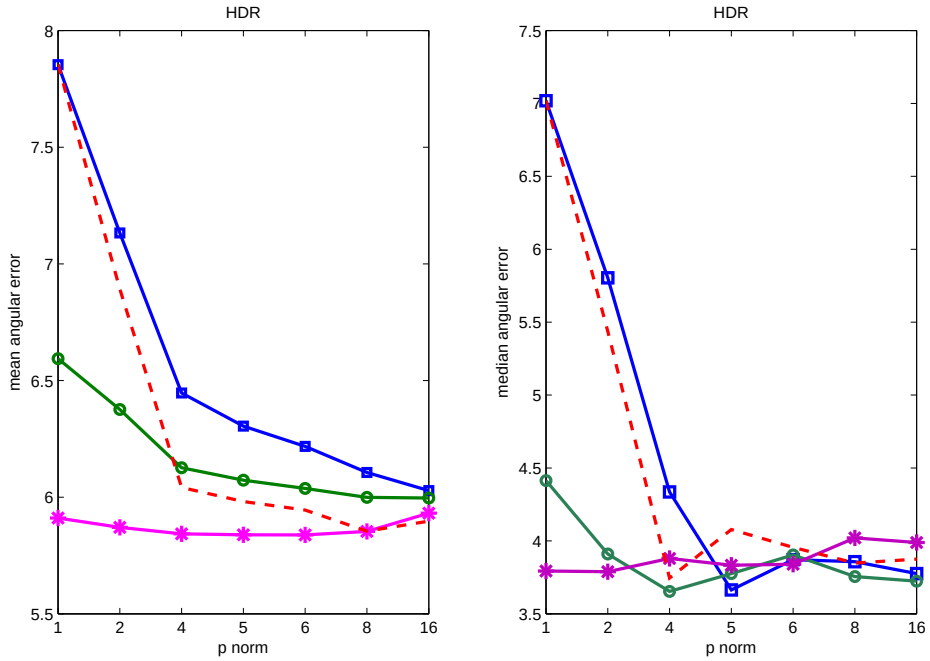

Figure A.2: Left and right panels show mean and median angular error for the HDR data set for 4 varieties of Shades of Gray. Blue line: thumbnails Shades of Gray. Green line: 5% of thumbnail. Magenta shows performance using top 5% of the original (not downsampled) image. Red line is the normal Shades of Gray.

the brightest regions is better preserved. Arguably, here the best performance is for the Shades of Gray algorithm working on the full resolution image but using the top 5% brightest pixels (at least for the mean error). But, the bright thumbnail Shades of Gray also works well. Taken together our experiments over both data sets, further validate the premise behind the robustness of the Shades of Gray framework and that preprocessing prior illuminant estimation helps.

References

[AG09] Marcel P. Lucassen Arjan Gijsenij, Theo Gevers. Perceptual analysis of distance measures for color constancy algorithms. JOSA A, 2009.

[AGA07] V. Agarwal, A. Gribok, and M. Abidi. Machine learning approach to color constancy. Neural Networks, 20:559–563, 2007.

[AGKA06] V. Agarwal, A. Gribok, A. Koschan, and M. Abidi. Estimating illumination chromaticity via kernel regression. in IEEE International Conference on Image Processing, pages 981–984, 2006.

[Bar95] Kobus Barnard. Computational Color Constancy: Taking Theory into Practice. PhD thesis, Simon Fraser University, 1995.

[Bar99] Kobus Barnard. Practical Color Constancy. PhD thesis, Simon Fraser University, School of Computing, 1999.

[Bar00] Kobus Barnard. Improvements to gamut mapping color constancy algorithms. In Proceedings European Conf on Computer Vision, 2000.

[Bar02] Kobus Barnard. Data for computer vision and color constancy [online], http://www.cs.sfu.ca/ colour/data/, 2002. http://www.cs.sfu.ca/ colour/data/.

[BCCS08a] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini. Improving color constancy using indoor-outdoor image classification. IEEE Transactions on Image Processing, 17:2381–2392, 2008.

[BCCS08b] Simone Bianco, G Ciocca, C Cusano, and R Schettini. Improving color constancy using indoor-outdoor image classification. IEEE Trans on Image Processing, 17:2381–2392, 2008.

[BCF02] K. Barnard, V Cardei, and B Funt. A comparison of computational color constancy algorithms - part i: Methodology and experiments with synthesized data. IEEE Trans on Image Processing, pages 972–984, 2002.

[BF97] D.H. Brainard and W.T. Freeman. Bayesian color constancy. J. Opt. Soc. Am. A, 14:1393–1411, 1997.

[BF99] K. Barnard and B. Funt. Color constancy with specular and non-specular surfaces. In Proc. IS&T/SID's Color Imag. Conf., pages 114–119, 1999.

[BGS08] S. Bianco, F. Gasparini, and R Schettini. Consensus-based framework for illuminant chromaticity estimation. Journal of Electronic Imaging, 17, no. 2, 2008.

[BK69] B Berlin and P Kay. Basic Color Terms: Their Universality and Evolution. Berkley, CA: University of California Press, 1969.

[BMCF02] Kobus Barnard, Lindsay Martin, Adam Coath, and Brian Funt. A comparison of computational color constancy algorithms; part two: Experiments with image data. IEEE Transactions on Image Processing, 11(9):985–996, 2002.

[BMFC02] Kobus Barnard, Lindsay Martin, Brian Funt, and Adam Coath. A data set for color research. Color Research & Application, 27(3):147–151, 2002.

[Buc80] G. Buchsbaum. A spatial processor model for object color perception. Journal of the Franklin Institute, pages 310:1–26, 1980.

[BW11] David H. Brainard and Alex R. Wade. Osa fall vision meeting, the spectrum recovery competition, 2011, 2011.

[CF99] V. Cardei and B. Funt. Committee-based color constancy. IS&T/SID's Color Imaging Conference, pages 311–313, 1999.

[CF03a] F Ciurea and B Funt. A large image database for color constancy research, 2003.

[CF03b] F. Ciurea and B.V. Funt. A large image database for color constancy research. Eleventh Color Imaging Conference, pages 160–164, 2003.

[CFB02] Vlad C. Cardei, Brian Funt, and Kobus Barnard. Estimating the scene illumination chromaticity by using a neural network. JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A, 19(12):2374–2386, 2002.

[CHZ08] Ayan Chakrabarti, Keigo Hirakawa, and Todd Zickler. Color constancy beyond bags of pixels. Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1-12:2509–2514, 2008.

[CHZ10] Ayan Chakrabarti, Keigo Hirakawa, and Todd Zickler. Computational color constancy with spatial correlation. Technical report, Harvard University, Computer Science Group, 2010.

[CHZ12] Ayan Chakrabarti, Keigo Hirakawa, and Todd Zickler. Color constancy with spatio-spectral statistics. IEEE Trans. Pattern Anal. Mach. Intell., 2012.

[CM09] Anustup Choudhury and Gérard Medioni. Color constancy using denoising methods and cepstral analysis. IEEE Int. Conf. on Image Processing, 1-6:1617–1620, 2009.

[DIS95] M. D'Zmura, G. Iverson, and B. Singer. Probabilistic color constancy. in Geometric Representations of Perceptual Phenomena. Lawrence Erlbaum Associates, pages 187–202, 1995.

[Ebn09] M Ebner. Color constancy based on local space average color. Machine Vision and Applications, vol. 20, no. 5:283–301, 2009.

[FBM98] Brian Funt, Kobus Barnard, and Lindsay Martin. Is machine colour constancy good enough? In Proceedings of the 5th European Conference on Computer Vision, pages 445–459, 1998.

[FDF94a] Graham D Finlayson, Mark Drew, and B Funt. Color constancy: generalised diagonal transforms suffice. JOSA, A, 1994.

[FDF94b] Graham D. Finlayson, Mark S. Drew, and Brian Funt. Spectral sharpening: sensor transformations for improved color constancy. J. Opt. Soc. Am. A, 11:1553–1563, 1994.

[FDH91] Brian Funt, Mark Drew, and Ja Ho. Color constancy from mutual reflection. Int. Journal on Computer Vision, 6:5–24, 1991.

[FF08] C. Fredembach and GD Finlayson. The bright-chromagenic algorithm for illuminant estimation. J Imaging Sci. Technol. 52 040906, 2008.

[FFB95] Graham D Finlayson, Brian Funt, and Kobus Barnard. Color constancy under varying illumination. Proccedings of the fifth International Conference on Computer Vision, pages 720–725, 1995.

[FH98] G. Finlayson and S. Hordley. A theory of selection for gamut mapping colour constancy. In Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on, pages 60 –65, jun 1998.

[FH00] G. Finlayson and S. Hordley. Improving gamut mapping color constancy. Image Processing, IEEE Transactions on, 9(10):1774 –1783, oct 2000.

[FHH01a] G.D. Finlayson, S.D. Hordley, and P.M. HubeL. Color by correlation: a simple, unifying framework for color constancy. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(11):1209 –1221, nov 2001.

[FHH01b] G.D. Finlayson, S.D. Hordley, and P.M. HubeL. Color by correlation: a simple, unifying framework for color constancy. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(11):1209 –1221, nov 2001.

[FHM05] GD Finlayson, S. Hordley, and P. Morovic. Colour constancy using the chromagenic constraint. IEEE Conference on Computer Vision and Pattern Recognition (IEEE 2005), pages 1079–1086, 2005.

[FHT06] G. Finlayson, S. Hordley, and I. Tastl. Gamut constrained illuminant estimation. International Journal of Computer Vision, 67:93–109, 2006.

[FHX05] G.D. Finlayson, S.D. Hordley, and R. Xu. Convex programming colour constancy with a diagonal-offset model. In Image Processing, 2005. ICIP 2005. IEEE International Conference on, volume 3, pages III – 948–51, sept. 2005.

[Fin95] Graham D Finlayson. Coefficient Colour Constancy. PhD thesis, Simon Fraser University, 1995.

[Fin96] Graham D. Finlayson. Color in perspective. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18:1034–1038, 1996.

[FNA04] David H. Foster, Sergio M. C. Nascimento, and Kinjiro Amano. Information limits on neural identification of colored surfaces in natural scenes. Visual Neuroscience, 21(003):331–336, 2004.

[For90] D. A. Forsyth. A novel algorithm for color constancy. International Journal of Computer Vision, 5:5–35, 1990. 10.1007/BF00056770.

[FS00a] Graham D Finlayson and Sabine Süsstrunk. Performance of a chromatic adaptation transform based on spectral sharpening. pages 49–55, 2000.

[FS00b] Graham D Finlayson and Sabine Süsstrunk. Spectral sharpening and the bradford transform. pages 236–243, 2000.

[FS10] Brian Funt and Lilong Shi. The rehabilitation of maxrgb. IS&T/SID's Color Imaging Conference, pages 256–259, 2010.

[FT04] G.D. Finlayson and E. Trezzi. Shades of gray and color constancy. IS&T 12th Color Imaging Conference, pages 37–41, 2004.

[FT05] G.D. Finlayson and E. Trezzi. Illuminant estimation using constrained minkowski norms. IEE International Conference on Visual Information Engineering, pages 337–342, 2005.

[FX03] Graham D Finlayson and R Xu. Convex programming color constancy. In Workshop on Color and Photometric Methods in Computer Vision. IEEE, October 2003.

[FX04] B. Funt and W. Xiong. Estimating illumination chromaticity via support vector regression. in IS&T/SID's Color Imaging Conference. IS&T - The Society for Imaging Science and Technology, pages 47–52, 2004.

[GB94] G.D.F and B.V.Funt. Color constancy with shadows. Perception, 23:89-90. Special Issue on the 17th European Conference on Visual Perception, Eindhoven., 1994.

[GG07] Arjan Gijsenij and Theo Gevers. Color constancy using natural image statistics. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 1–8, 2007.

[GG11] Arjan Gijsenij and Theo Gevers. Color constancy using natural image statistics and scene semantics. IEEE Trans. Pattern Anal. Mach. Intell., 33:687–698, 2011.

[GGJ07] Arjan Gijsenij, Theo Gevers, and Joost, Van De Weijer. Color constancy by derivative-based gamut mapping. In Peter Belhumeur, Katsushi Ikeuchi, Emmanuel Prados, Stefano Soatto, and Peter Sturm, editors, Proceedings of the First International Workshop on Photometric Analysis For Computer Vision - PACV 2007, page 8, Rio de Janeiro, Brasil, 2007. INRIA.

[GGJ09] Arjan Gijsenij, Theo Gevers, and Joost van de Weijer. Physics-based edge evaluation for improved color constancy. IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, 1-4:581–588, 2009.

[GGJ10] Arjan Gijsenij, Theo Gevers, and Joost, van de Weijer. Generalized gamut mapping using image derivative structures for color constancy. International Journal of Computer Vision, 86:127–139, 2010.

[GGJ11] A. Gijsenij, T. Gevers, and J van de Weijer. Computational color constancy: Survey and experiments. Image Processing, IEEE Transactions on, 20(9):2475 –2489, sept. 2011.

[GJT86] R. Gershon, A.D. Jepson, and John K. Tsotsos. Ambient illumination and the determination of material changes. JOSA, A, pages 1700–1707, 1986.

[GJT87] R. Gershon, A. Jepson, and J Tsotsos. From [r,g,b] to surface reflectance: computing color constant descriptors in images. International Joint Conference on Artificial Intelligence, pages 755–758, 1987.

[GRB+08] Pether Gehler, Carsten Rother, Andrew Blake, Toby Sharp, and Tom Minka. Bayesian color constancy revisited. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1-12:3291–3298, 2008.

[Hea91] Glen Healey. Estimating spectral reflectance using highlights. Image and Vision Computing, 9:333–337, 1991.

[HF06] Steven D. Hordley and Graham D. Finlayson. Reevaluation of color constancy algorithm performance. J. Opt. Soc. Am. A, 23:1008–1020, 2006.

[HHFD97] Paul M. Hubel, Jack Holm, Graham D. Finlayson, and Mark S. Drew. Matrix calculations for digital photography. IS&T and SID's 5th Color Imaging Conference, pages 105–111, 1997.

[Hor06] S D Hordley. Scene illuminant estimation: Past, present, and future. Color Res. Appl., 31(4):303–314, 2006.

[HRNL01] Hernandez-Andres, J, Romero, J, Nieves, J. L., and Lee, R. L. Jr. Color and spectral analysis of daylight in southern europe. J. Opt. Soc. Am. A, 18(6):1325–1335, Jun 2001.

[HS94] Glenn Healey and David Slater. Global color constancy: recognition of objects by use of illumination-invariant properties of color distributions. Journal of the Optical Society of America, A, 1994.

[JGG07] Joost van de Weijer, Th Gevers, and A Gijsenij. Edge-based color constancy. IEEE Trans on Image Processing, pages 2207–2217, 2007.

[JSV07] J, van de Weijer, C. Schmid, and J. Verbeek. Using high level visual information for color constancy. in IEEE International Conference on Computer Vision, 2007.

[JZQ11] Guo-Quan Jiang, Cui-Jun Zhao, and Jun-Yan Qi. The research of image segmentation based on color characteristic. In in Proceedings of the 2011 International Conference on Machine Learning and Cybernetics, 2011.

[KF75] A.N. Kolmogorov and S.V. Fomin. Introductory Real Analysis. Dover Publications, 1975.

[Kob00] Barnard Kobus. Improvements to gamut mapping colour constancy algorithms. In Computer Vision - ECCV 2000, volume 1842 of Lecture Notes in Computer Science, pages 390–403. Springer Berlin - Heidelberg, 2000.

[KSK] Gudrun J. Klinker, Steve A. Shafer, and Takeo Kanade. The measurement of highlights in color images. In Gleen E. Healey and Steven A. Shafer and Lawrence B. Wolf, editor, Color, pages 309-332. Jones and Bartlett, Boston, 1992.

[Lan77] E. H. Land. The retinex theory of color vision. Scientific American, pages 108–129, 1977.

[Lee92] Hsien-Che Lee. Method for Computing Scene Illuminant Chromaticity from Specular Highlights. Jones and Bartlett, Boston, 1992.

[LGG+09] R. Lu, A. Gijsenij, T. Gevers, K. van de Sande, J. Geusebroek, and D. Xu. Color constancy using stage classification. in IEEE International Conference on Image Processing, 2009.

[LJBS86] Hsien-Che Lee, Edwin J, Breneman, and Carl P Schulte. Modeling light reflection for computer color vision. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12:402–409, 1986.

[MF10] M. Mosny and B. Funt. Cubical gamut mapping colour constancy. in IS&T's European Conference on Color in Graphics, Imaging and Vision, 2010.

[MW86] L.T. Maloney and B.A. Wandell. Color constancy: A method for recovering surfaces spectral reflectances. J. Opt. Soc. Am. A, pages 3:29–33, 1986.

[NFF02] Sergio M. C. Nascimento, Flavio P. Ferreira, and David H. Foster. Statistics of spatial cone-excitation ratios in natural scenes. J. Opt. Soc. Am. A, 19(8):1484–1490, Aug 2002.

[NSRG10] V. Nedovic, A. Smeulders, A. Redert, and J. Geusebroek. Stages as models of scene geometry. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010.

[PBV10] C. A. Parraga, R. Baldrich, and M. Vanrell. Accurate mapping of natural scenes radiance to cone activation space: A new image dataset. CGIV 2010/MCS'10 - 5th European Conference on Colour in Graphics, Imaging, and Vision - 12th International Symposium on Multispectral Colour Science, Society for Imaging Science and Technology, 2010.

[PVCV09] C. A. Párraga, J. Vazquez-Corral, and M. Vanrell. A new cone activation-based natural images dataset. Perception, 2009.

[RLG09] T. Gevers D. Xu V. Nedovic R. Lu, A. Gijsenij and J. Geusebroek. Color constancy using 3d stage geometry. in IEEE International Conference on Computer Vision, 2009.

[RML03] C. Rosenberg, T. Minka, and A. Ladsariya. Bayesian color constancy with nongaussian models. in Advances in Neural Information Processing Systems, 2003.

[RNI04] R.Tan, K. Nishino, and K. Ikeuchi. Color constancy through inverse-intensity chromaticity space. Journal of the Optical Society of America, A, 21:321–334, 2004.

[RNK+09] E. Rahtu, J. Nikkanen, J. Kannala, L. Lepisto, and J Heikkil. Applying visual object categorization and memory colors for automatic color constancy. in Proceedings of the International Conference on Image Analysis and Processing, 2009.

[Sap99] G. Sapiro. Color and illuminant voting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21:1210–1215, 1999.

[SF10] L. Shi and B. V. Funt. Re-processed version of the gehler color constancy database of 568 images. availabe on-line: http://www.cs.sfu.ca/ colour/data/, 2010.

[Sha85] S. A. Shafer. Using color to separate reflection components. Color Research and Application, 10:210–218, 1985.

[SHF05] G. Schaefer, Steven D. Hordley, and Graham D. Finlayson. A combined physical and statistical approach to colour constancy. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 148–153, 2005.

[Süs05] Sabine Süsstrunk. Computing Chromatic Adaptation. PhD thesis, University of East Anglia, 2005.

[SSS92] B Smith, C Spiekermann, and R Sember. Numerical methods for colorimetric calculations: Sampling density requirements. Color research and applications, 17(6):394–401, 1992.

[TCRK01] Y. Tsin, R. Collins, V. Ramesh, and T. Kanade. Bayesian color constancy for outdoor object recognition. in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 1132–1139, 2001.

[TEW01] Shoji Tominaga, Satoru Ebisui, and Brian A. Wandell. Scene illuminant classification: brighter is better. J. Opt. Soc. Am. A, 18(1):55–64, Jan 2001.

[Tom96] Shoji Tominaga. Surface reflectance estimation by the dichromatic model. Color Research and Application, pages 104–114, 1996.

[TW96] Shoji Tominaga and Brian A. Wandell. Standard surface-reflectance model and illuminant estimation. J. Opt. Soc. Am. A, pages 576–584, 1996.

[VAA09] A. Koschan B. Abidi V. Agarwal, A. Gribok and M. Abidi. Illumination chromaticity estimation using linear learning methods. Journal of Pattern Recognition Research, 4:92–109, 2009.

[VCVBT12] Javier Vazquez-Corral, Maria Vanrell, Ramon Baldrich, and Francesc Tous. Color constancy by category correlation. IEEE Transactions on Image Processing, 21(4):1997–2007, April 2012.

[VPVB09] Vazquez-Corral, J., C. A. Párraga, M. Vanrell, and R Baldrich. Color constancy algorithms: Psychophysical evaluation on a new dataset. Journal of Imaging Science and Technology, 2009.

[WXL09] N. Wang, D. Xu, and B. Li. Edge-Based Color Constancy via Support Vector Regression. IEICE Transactions on Information and Systems, 92:2279–2282, 2009.

[XF06] W. Xiong and B. Funt. Estimating illumination chromaticity via support vector regression. Journal of Imaging Science and Technology, 50(4):341–348, 2006.

[XSF+07] W. Xiong, L. Shi, B. Funt, S. Kim, B. Kan, and S. Lee. Illumination estimation via thin-plate spline interpolation. in IS&T/SID's Color Imaging Conference, 2007.