COLOUR CONSTANCY THEORIES IN THE REAL WORLD

by

Florian Ciurea
M.Sc., University "Politehnica" of Bucharest, 1997

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the School
of
Computing Science

SIMON FRASER UNIVERSITY

April 2004

APPROVAL

Name: Florian Ciurea

Degree: Doctor of Philosophy

Title of Thesis: Colour Constancy Theories in the Real World

Examining Committee:

Chair: **Richard Vaughan**
Assistant Professor

Dr. Brian Funt
Senior Supervisor
Professor

Dr. Mark Drew
Supervisor
Associate Professor

Dr. Tim Lee
SFU Examiner
Adjunct Professor

Dr. Alessandro Rizzi
External Examiner
Assistant Professor
Dipartimento di Tecnologia dell'Informazione
Università di Milano

Date Approved: April 2ND. 2004

SIMON FRASER UNIVERSITY

Partial Copyright Licence

ABSTRACT

In recent years, several methods to solve the colour constancy problem have been introduced and studied. Colour constancy is an important area in machine vision: it provides a visual system with the capability to compensate for the effect of illumination in a scene. The colours we humans perceive are not easily expressible by a physical apparatus, which is in fact more sensitive to changes in illumination conditions. The human visual system is able to make several compensations and adjustments to ambient illumination conditions, so that we perceive illuminant-independent descriptors of the scene.

This thesis represents a series of experimental approaches to the colour constancy problem. An important part focuses on discussions and concrete implementations of the well-known and controversial Retinex model for human vision and application of the method on real images and choosing its parameters. The Retinex theory of human vision represents an important contribution in colour vision with strong implications in the colour constancy problem. Although the theory has been around for more than three decades, the lack of an efficient implementation and analysis of the effect of its parameters has raised many discussions and scepticism. This part of the thesis provides some common ground in further investigations into the Retinex theory.

A novel method for acquiring a large database for colour constancy research is presented, along with direct applications with improvements to the colour by correlation method. A large image database for colour constancy provides sufficient data to compute some meaningful statistics with respect to the interaction between the colours observed in

the world and the actual measured illuminants in which these colours are recorded. The use of these statistics in the Bayesian approach improves the performance of the colour by correlation method for colour constancy. Another application is in the form of testing the effectiveness on real images of a recently introduced theory for colour constancy using the redness-luminance correlation in scenes.

ACKNOWLEDGEMENTS

First I wish to thank my senior supervisor, Dr. Brian Funt for the trust and continuous support through these years. Working under his direction has been the most important aspect of my experience as a graduate student at Simon Fraser University.

Many thanks go to John McCann for sharing his knowledge about the visual system and for his constant patience. Thanks to my supervisor Dr. Mark Drew for his always frank suggestions and comments.

I would also like to thank Dr. Kobus Barnard and Dr. Vlad Cardei for the useful discussions we had on several occasions. In the lab, thanks to Bill Cressman and Behnam Bastani for making life nicer.

Financially, I acknowledge the support from the Centre for Systems Science, Simon Fraser University, the Ebco-Eppich foundation and the National Sciences and Engineering Research Council of Canada.

Finally, I would like to thank my family for their full support and understanding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION AND THESIS OVERVIEW

This thesis presents several approaches to colour constancy from an experimental perspective. Colour constancy theories fill a gap between illuminant characterization in real scenes and colour appearance. The colour constancy problem is defined as the ability of a visual system to discard the effects of the illumination in a scene. In computational colour constancy – also referred to as illuminant estimation – we would like to achieve an illumination-invariant image: an image that appears to be taken under a canonical illuminant for which the visual system is adapted. Over the years, this problem has proven to be quite hard, even if further simplifying assumptions are considered such as the existence of a single – global illuminant for the scene along with the absence of specular reflections. In essence, the difficulty of the colour constancy problem comes from its under-constrained nature. This very aspect yields the problem naturally to statistical approaches rather than to simple naïve methods. Ultimately, the main obstacle in designing an efficient method for colour constancy lies in our limited knowledge and understanding of the human visual system. In contrast to the computational approach to colour constancy, there is ample evidence that the human visual system is equipped with some local adaptation mechanisms that explain phenomena such as simultaneous contrast. These adaptation mechanisms are simulated by the Retinex model of human vision, which aims to predict human sensations, thus putting the colour constancy phenomena in a different, psychological perspective.

In the thesis, several advances in different related colour constancy theories are proposed: a method to compute chromatic adaptation transforms for perceived illuminant change in simple scenes, a way to characterize the well-known Retinex theory for dealing

with real images. Based on a large colour constancy database, other advances include ways to improving illumination estimation in real scenes based on colour-by-correlation method, and investigating a new hypothesis for illumination estimation and its robustness in real images. All of these are different approaches to colour constancy and their reference to colour constancy is outlined below.

The first two chapters describe the colour constancy problem and how it relates to the problem of recovering reflectance and the chromatic adaptation transforms. An overview of computational models for colour constancy is presented in Chapter 2.

In Chapter 3, a method of using the full spectral information in the chromatic adaptation transforms is introduced. Chromatic adaptation transforms are typically employed by any colour appearance model and they essentially quantify our incomplete colour constancy effects in simple viewing conditions – not complex real scenes. Existing chromatic adaptation transforms compute a universal transformation, based on the CIE tristimulus values of the two illuminants involved. However, with the widespread availability of spectroradiometers, we can compute the real spectra of the two illuminants involved, and using a white-point preserving sharpening method we can tune this transformation for each pair of illuminants, thus potentially achieving better results. Unfortunately, most current colour appearance experiment documenting observers' subjective evaluations do not include full spectral information of the illuminants involved, and therefore real-world testing of this proposed method is still limited.

Chapters 4, 5 and 6 focus on providing concrete implementations of the well-known and controversial Retinex model of human vision and discussions on application of the method on real images and choosing the parameters for the model. The Retinex theory

represents an important contribution to human colour vision with strong implications in the colour constancy problem. Although the theory has been around for several decades, the lack of an efficient implementation and analysis of the effect of its parameters has raised many discussions and scepticism. While the Retinex theory is often being classified as a method for illumination estimation, its scope goes beyond colour constancy, by trying to be a simplified model for the human visual system, or to estimate visual appearance in complex scenes. While Chapter 4 provides concrete implementations for two of the Retinex methods, chapters 5 and 6 focus on tuning the model's parameters to deal with real scenes, based on psychophysical experiments.

In Chapter 7, a novel method for a acquiring a large database for colour constancy research is presented, along with applications and improvements to the colour by correlation method. A large image database for colour constancy provides sufficient data to compute some meaningful statistics with respect to the interaction between the colours observed in the world and the actual measured illuminants in which these colours are recorded. Using these statistics in the Bayesian approach improves the performance of the colour by correlation method for colour constancy.

Finally, Chapter 8 presents a study on the effectiveness on real images of a recently proposed method for illumination estimation based on luminance-redness correlation in a scene. This method is interesting because it relates simple methods like gray world with more advanced techniques such as gamut mapping or colour-by-correlation. However, testing the method on a more diverse hyperspectral database and on real images shows that the proposed correlation is not a robust illumination estimation method.

CHAPTER ONE
INTRODUCTION – COLOUR CONSTANCY

Definition of colour constancy

Colour constancy is defined as the ability of a visual system to diminish – or in the ideal case discard – the effects of the illumination in a scene. With this property, a visual system is able to perceive descriptors of certain objects that are independent of the ambient illumination, such as the objects' surface reflectance. There is good evidence[1, 2, 3] that humans exhibit colour constancy to a relatively high degree and it is believed that the mechanism of colour constancy in the human visual system is part of an evolutionary process. Indeed, if we believe that colour plays an important part in object recognition[4] – given the great variety of natural and man-made illuminants – the human visual system benefits greatly from the mechanism of colour constancy. For instance, a banana appears yellow in various types of illumination, even though the flux of energy reaching our eyes from the fruit will be very different for various lighting conditions. As a result of our own colour constancy mechanisms, we humans usually associate colours with certain natural objects, even though the light energy is different across the illumination conditions.

With the recent advances in imaging systems, colour constancy is becoming a more prevalent problem for the colour imaging devices. We can easily become aware of the colour constancy mechanisms in the human visual system by comparing device outputs of the same scene under different types of illumination. The simple fact that typical photographic film renders accurate colours outdoors in daylight but is problematic in indoor scenes without flash is a direct consequence of our own colour constancy: indoor scenes are more red-yellowish due to the different power distribution of tungsten illumination as opposed to that of daylight for which the film has been calibrated. This aspect has motivated the use of different colour filters in traditional film cameras or camcorders that compensate for the colour cast introduced by various illumination types.

It is difficult to predict the colour appearance of objects in complex scenes, which makes accurate colour reproductions a challenging task. Advances in understanding the mechanisms for colour constancy and compensating for the colour of the illuminant would contribute directly to improvements in the quality of cross-media colour reproductions.

Colour constancy in the human visual system

In the current state of research there is no general consensus on the type of neural activity in our brains that leads to the sensation of colour. Important experimental findings about the organization of the primary visual cortex are provided by Hubel and Wiesel.[5] It is believed that monkeys exhibit the same types of colour sensations that we humans do, although this is more of a speculation. Zeki's studies[6] provide some clues regarding the mechanisms of colour constancy present in monkeys that are in accordance to experiments on humans.[1] Little is known about colour constancy in other species other than goldfish[7, 8], honeybees[9] and some butterflies.[10] Given the limited knowledge about the intrinsic neurological mechanisms of colour vision and colour constancy in humans, most of what is known comes from experimental studies on the subjective colour appearance.

There is evidence that humans do not compensate fully for the colour of the illuminant.[11, 12] Quantifiable measures of human colour constancy are based on asymmetric colour matching experiments. Brainard and Wandell[13, 14] initially reported that in laboratory conditions, subjects' performance on colour constancy is pretty conservative, compensating for about half of the true illumination change. More recently, Brainard and Brunt's psychophysical exmperiments[2, 3] show that in near-natural scenes, colour constancy in humans is actually quite good. Unfortunately, a real analysis on human colour constancy in complex, natural scenes is not yet available, partly because this process is tedious and not easily quantifiable.

Not only is the colour constancy mechanism in humans imperfect, it sometimes fails completely. An example is provided is the case of low-pressure sodium street lighting[15] – practically a monochromatic light source – under which scenes appear to be characterized only by lightness, as in black-and-white imaging. In these scenes a strong yellowish cast is still noticeable. Another example is provided by Maximov's shoeboxes[16] in which observers

are asked to report the colours in a collage of papers viewed in a shoebox. Two versions of the same arrangement of coloured papers were prepared; the reflectances of the papers in the second display differ from the ones of the first by a systematic amount. If the second display is illuminated by a different light source that is carefully chosen to compensate for the systematic differences in two displays, the two collages appear identical. A simpler argument for the non-absolute definition of colour constancy itself is the very case of a single surface occupying the whole visual field, illuminated uniformly by an unknown light source. Such limiting case is unsolvable from the colour constancy perspective, for humans, and for any imaging device in fact. We need several surfaces in a scene for the colour constancy problem to be even well posed. This limiting case emphasizes an important aspect in the quest to solving for the strongly underdetermined problem of colour constancy: the inherent statistical nature of the problem itself. In the following sections, we will investigate computational models exploiting this very aspect.

Colour constancy, recovering reflectance and metamerism

In a simplified two-dimensional world of matte surfaces illuminated uniformly, the following equation describes the response of a device capturing the physical surface-illuminant interaction:

$$r^{(k)} = \int_{\lambda} R^{(k)}(\lambda) \cdot E(\lambda) \cdot S^{x,y}(\lambda) d\lambda \qquad (1)$$

Here, $R^{(k)}(\lambda)$ gives the spectral sensitivity function of the k-th sensor, $E(\lambda)$ is the spectral power distribution of the illuminant, and $S^{x,y}(\lambda)$ is the surface reflectance as a function of wavelength. Applied to all pixels (x, y) in the scene, the equation describes the problem of recovering reflectance: recover $S^{x,y}(\lambda)$, given the set of observed sensor responses $r^{(k)}$. The problem can be reformulated as recovering $E(\lambda)$, the illuminant spectral power distribution, from the sensor responses $r^{(k)}$. Considering n distinct surfaces in the scene and the typical case of k=3 sensor classes, the set of equations have 3n knowns, in the form of sensor responses. In the discrete case of s samples, the equation becomes:

$$r^{(k)} = \sum_{i=1,s} R^{(k)}(\lambda_i) \cdot E(\lambda_i) \cdot S^{x,y}(\lambda_i) \quad (2)$$

In this form we can easily see the under-constrained nature of the problem.[17] We have k×n knowns (the set of sensor responses for each surface) and n×s (the set of reflectances) + s knowns (the illuminant power distribution). Given that the typical number of sensor classes k is much lower than the number of samples required to characterize surface reflectances and illuminants, we must use additional knowledge about the illuminants and surfaces in order to solve for this problem. This set of equations are the core of a model of linear methods for colour constancy,[18] based on a reduced dimensionality of reflectances and illuminants – which will be presented further. For s = 3, the problem of recovering reflectance is reduced to the colour constancy problem, which still remains underdetermined. Because the human visual system has only three cone receptor classes, we can only hope to achieve colour constancy for the three-samples case. The discrete sampling of the continuous space of surface reflectances and spectral power distributions of illuminations in three classes gives rise to a well-known phenomena in colour science called "metamerism": the possibility of perceiving the same colour from different surface reflectances under a particular illuminant. In particular, two metameric samples may appear to have the same colour under a certain illumination, yet will look different under another illumination condition.

The dichromatic reflection model

The light reflected from a surface is the result of an interaction of different physical processes and the full model of reflectance is rather complex. The dichromatic reflection model due to Shafer[19] is a simple model that explains several important aspects of the physics involved in light reflecting off dielectric surfaces – non-conducting materials. Dielectrics are characterized by a clear substrate covering the colorant particles. Through the reflection process, incident light is transformed in three forms of energy:

a) light reflected from the interface;
b) light reflected from the body;
c) heat absorbed by the material.

Once the rays of light arrive at the interface, they get reflected off the clear substrate in the form of interface reflection. The remaining energy enters the material, where it is scattered between the material particles and eventually comes out in the form of body reflection. Depending on the material, some of the energy will also be transformed into heat. Light reflected from the interface is concentrated around a small angular cone. Interface reflections are also called specular reflections. These types of reflections can only be seen under certain viewing angles, when the viewer is situated in their angular cone. Another interesting property of specular reflections is that, because they reflect light equally from all wavelengths, they provide important cues about the colour of the illuminants. However, very often these are not a reliable source of information about the illuminants either because they are not necessarily visible from the camera's perspective or, in cases where they are available, their brightness often exceeds the dynamic range of the capturing device.

In contrast with interface reflection, body reflection is diffuse and usually emerges equally in all directions. Surfaces that do not exhibit interface reflection are called matte, they are characterized by body reflection only, which is independent of the viewing angle and that of the incident light.

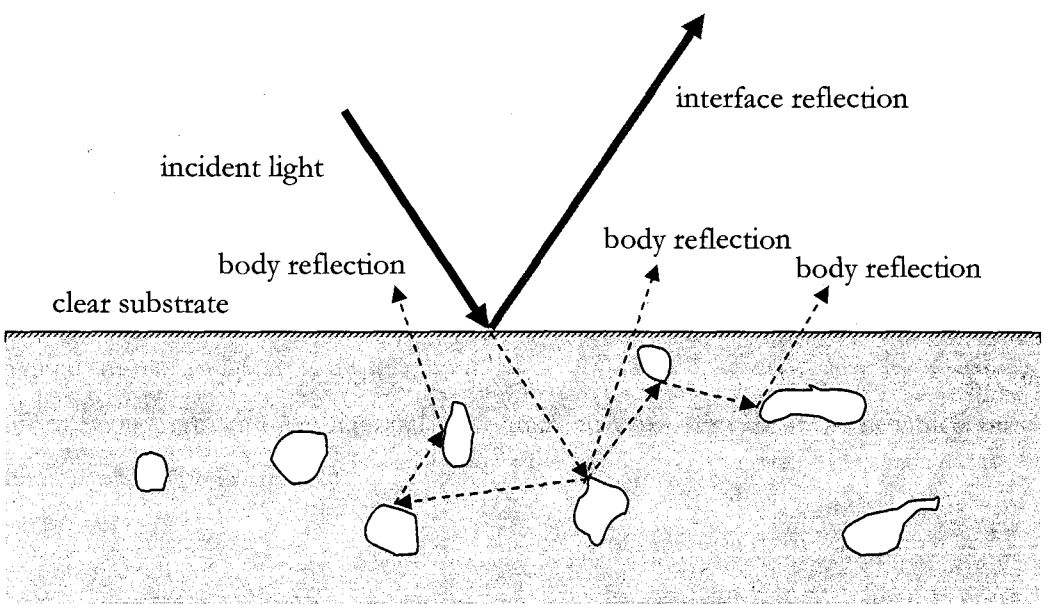


Figure 1. The dichromatic reflection model

The dichromatic reflection model states that the total radiance of light reflected off a dielectric surface is roughly the sum of the radiance in the form of interface reflection plus the radiance as body reflection. The dielectric model does not apply to metals and fluorescent surfaces. Fluorescent materials have a rather particular property, in that they absorb energy at lower wavelengths and emit it at higher wavelengths.

The diagonal model of illumination change

In the quest for illumination-invariant scenes, computational models of colour constancy typically consist of two chained processes:

a) estimate the scene illumination;
b) correct for the scene illumination such that an illumination-invariant image is obtained

While it is true that estimating the scene illumination is the hardest problem of the two, finding a suitable model for illumination change has an interesting history.

A simple form of modeling illumination change represents a linear transformation, which, for a three-sensor device can be written as follows:[20]

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{3}$$

(L', M', S') and (L, M, S) represent the device response for the same physical location (pixel) under the two different illuminants. It turns out that the simple linear model is quite good at modeling illumination change and is justified by the simple laws of additive colour mixture. A general transformation is represented by a full 3 x 3 matrix, but by restricting the transformation matrix to be diagonal we obtain an even simpler form in which only three parameters are to be estimated:[21]

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} k_L & 0 & 0 \\ 0 & k_M & 0 \\ 0 & 0 & k_S \end{bmatrix} \cdot \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (4)$$

In this form the model is known as the von-Kries adaptation rule[22] and the diagonal terms are well-known as von-Kries coefficients. Von Kries suggested that in the human vision system, the visual pathways have a mechanism of adjusting to the illumination by scaling the signals individually on each of the short, medium and long cones. While challenged at times (Burnham, Wassef), the rule has continued to be adopted as a general law of expressing illumination change.

The diagonal model is preserved in a perspective chromaticity space:[23]

$$\begin{bmatrix} l' \\ m' \end{bmatrix} = \begin{bmatrix} k_l & 0 \\ 0 & k_m \end{bmatrix} \cdot \begin{bmatrix} l \\ m \end{bmatrix} \quad (5)$$

$$\text{where } l = \frac{L}{M}, m = \frac{M}{S}, k_l = \frac{k_L}{k_M}, k_m = \frac{k_M}{k_S}$$

Worthey and Brill[24, 25] have explored the limitations of the diagonal model and the conditions under which it holds. The diagonal model of illumination change works quite well in the case of narrow-band and non-overlapping sensors. This is easily understandable because in the extreme case of sensors being modeled by delta functions at different wavelengths the diagonal model would hold perfectly. Based on this observation, Finlayson et al[26] have proposed a method of transforming the sensor functions by means of a linear transformation in which the diagonal model of illumination change holds more exactly. This method is called "spectral sharpening". Therefore, to the extent of the diagonal model, the resulting sharpened sensors are optimal for colour constancy. Three methods for spectral sharpening are proposed in[26]:

a) sensor based sharpening proposes to find the linear transformation such that the sensors are optimally narrow; this transformation is only function of the sensors functions themselves

b) database sharpening finds the optimally linear transformation for a given set of reflectance data and a set of illuminants
c) perfect sharpening is a formulation in a world of two-dimensional illuminant spectra and three-dimensional reflectance spectra

Given the current performance of colour constancy algorithms, the diagonal model of illumination change gives good enough approximation and, from a practical perspective, there is little to be gained by considering alternate models. Where available, spectral sharpening gives equivalent results to using a full 3 x 3 linear transformation.

A practical aspect of the diagonal model is that it reduces the problem of colour correction to the problem of having a good white (hence the notion of white balance in imaging devices). The colour of a standard white reflectance present in the scene is usually sufficient information for determining the illuminant colour. This very aspect is at the root of colour appearance models.[27, 28]

Chromatic adaptation transforms

Chromatic adaptation transforms model illumination change in simple viewing conditions. More specifically, they relate the change in cone tristimulus values (XYZs) for a certain patch when viewed under a different, "reference" light from a "test" light. The "reference" and "test" lights are specified in terms of tristimulus (XYZ) values, as is the corresponding patch. We can think of chromatic adaptation transforms as the consequence of the human visual system's limitations in achieving perfect colour constancy: chromatic adaptation transforms predict perceived illumination change.

Historically, Lam was the first to provide a chromatic adaptation transform, also known as the Bradford transform. In his experiment[29] he used 58 wool samples to model the various degrees of colour constancy in the context of illuminant change from standard D65 illuminant to standard illuminant A. Lam used a memory matching experiment, in which he asked the observers to describe the appearance of the samples in the Munsell coordinate system. He then transformed the Munsell values in the CIE tristimulus

coordinates. Lam devised a chromatic adaptation transform that would explain the change in tristimulus values, based on the following assumptions:

1) the transformation preserves achromatic constancy for neutral samples;
2) the transformation is the same for different pair of illuminants;
3) the transformation is reversible: going from one illuminant to another and back should leave the result unchanged.

The Bradford transform is still at the root of colour appearance models, that aim to predict colour sensations under more complex viewing conditions.

CHAPTER TWO
OVERVIEW OF COMPUTATIONAL MODELS FOR COLOUR CONSTANCY

Linear models for colour constancy

As previously shown, the colour constancy problem is underdetermined. The fact that humans seem to be pretty good at recovering illuminant-independent descriptors from the product of reflectance, illuminant and sensor response functions is evidence that there must be something about the actual interplay between the surfaces and illuminants that is exploited by our vision system. In particular, there is wide evidence that the spectral reflectances lay in a finite-dimensional space. Cohen's measurements[30] of the set of Munsell chips argued that the first three of the principal component decomposition support over 99% of the variation in the data. Krinov has continued reflectance measurements[31] on a selection of natural reflectances. Based on Krinov's data, Dannemiller[32] concluded that a number of three functions represent natural reflectances accurately. However, further analysis by Brown indicates that Krinov has not actually measured natural reflectances, but rather large terrains. Further analysis[33] concludes that the dimensionality of natural reflectances could be as high as five. Maloney[34] fitted a large number of surface reflectances including the ones measured by Krinov and Cohen and found that the a linear model of five to six dimensions can represent them accurately.

Maloney and Wandell have exploited the finite-dimensionality of illuminants and reflectances in a linear approach to solving colour constancy.[18, 20] They show that for a number N of photoreceptors, if the dimensionality of illuminants is N-1 and the dimensionality of reflectances is N, solving for colour constancy is transformed into a determined problem. Unfortunately, for the typical case of N=3 classes of photoreceptors, the method is not very successful, as the dimensionality of reflectances is usually higher.

Gray world

The gray world is probably the simplest of the colour constancy algorithms. It works by computing the statistics of the scene and then comparing these statistics with some predetermined values, followed by adjustment of the scene such that the new statistics match the predetermined ones. One of the simplest statistics is the average device (R, G, B) of the scene. The gray world algorithm makes two implicit assumptions about the real world:

a) all scenes properly colour balanced have the same "normal" statistics
b) the deviation from the actual statistics of the scene to the normal statistics is due to the scene illumination

The simplest form of the gray world approach assumes that each scene should have an average (R, G, B) of 50 % gray. As an algorithm for colour constancy, the scene illuminant is characterized by the ratio of the actual average found in each of the (R, G, B) channels versus the expected average (50%), using the diagonal model for illumination change. In its initial formulation, the idea is due to Buchsbaum,[35] based on an "adaptation level theory" introduced by Helson.

Buchsbaum postulates: "the system arrives at the illuminant estimate assuming a certain standard common spatial average for the total field. It seems that arbitrary natural everyday scenes composed of dozens of colour sub-fields, usually none highly saturated, will have a certain, almost fixed spatial spectral reflectance average. It is reasonable that this average will be some medium gray, which comes back to Helson's principle".

Gershon et al.[36] further improved on the method in several ways. The choice of gray has been derived statistically from natural occurrences of reflectances and illuminants in the world. Because the algorithm works on sensor responses directly, it is more efficient when the "normal" gray is computed through the device or using the device sensor spectral sensitivity functions. Another aspect of Gershon's extension is the introduction of segmentation as a pre-processing step in the gray world estimate of the scene. The purpose of the segmentation is to compute descriptors of surfaces as opposed to descriptors of pixels, which are inherently less representative of the actual surfaces present in the scene.

However, as it is well known, segmentation in itself is a hard problem and not very robust on real scenes.

While certainly very appealing for its simplicity, the gray world algorithm is essentially naïve, as its two main assumptions are flawed for many natural scenes: the actual departure from the "normal" gray in properly balanced scenes is not necessarily due to an illuminant cast, as is the case for many scenes depicting deep-blue sky or predominantly greenery, autumn foliage, etc. In practice, the gray world algorithm has the tendency to artificially de-saturate colours in natural scenes in order to achieve the "normal" gray.

Retinex and lightness models

The Retinex theory was introduced by Land[37] as a model of human vision after a series of experiments with colour appearance led him to believe that something is inherently wrong with the general idea of how colour is perceived. Land was surprised to notice that in certain circumstances there is almost no correlation between the incident flux of light coming from certain areas in his experimental displays and the perceived colour – which in turn seems to be highly correlated with the reflectance of the coloured patches. His experiments represent rather a confirmation of the mechanism of colour constancy in the human visual system, and they do not necessarily undermine the classical laws of colour mixture and colour appearance as Land has originally thought.[38] Land's theory is called "Retinex", from "retina" and "cortex", where most of the human visual processing takes place.

There are several important components of the theory that are based on experimental observations: First of all, the theory embraces the von Kries adaptation rule of illumination change. The motivation for this essential part of the method was given by experiments that support correlation between colour appearance and the so-called scaled integrated reflectance of a given studied area. The scaled integrated reflectance is defined as the ratio of integrated radiance of a certain patch and the integrated radiance of the white patch in the scene. Another important aspect on which the theory relies on is the smooth spatial variation of the illumination in real scenes. Retinex discounts the illumination based on this spatially smoothness constraint. In its first formulation,[37] the algorithm computes random paths in the scene. Along each of these paths the algorithm compute ratios of

adjacent pixels that are then propagated along by a multiplicative operation. Any given ratio product is reset to 1 if it becomes supra unitary (Retinex reset operation) or is simply ignored if its value falls behind a certain threshold (Retinex threshold operation). For a certain pixel, ratio products along paths from several directions are then averaged together to give the final pixel value. The reset operation implements the scaling on each independent photoreceptor class and the threshold operation implements the smoothness constraint on illumination spatial variation.

In later implementations,[39, 40, 16] Land and his colleagues discarded the threshold operation. Interestingly, with this simplification, it is only the choice of paths that makes this algorithm worthwhile. For infinite paths, the Retinex process becomes the simple approach of scaling each photoreceptor channel by the maximum. In this particular case, the Retinex model essentially becomes the white-patch algorithm.

Retinex differs from many colour constancy methods in that it does not aim to find a single chromaticity for the scene illumination. Instead, it adjusts the image colours in a non-global manner as is necessary since the model attempts to match human visual response. It is a theory for the human visual system processing rather than simply a computational method for colour constancy. Supported by a series of experiments,[1, 16] the theory rejects the idea of any correlation between colour appearance and scene average, therefore coming in opposition to the gray world algorithm and favouring the von Kries model.

In a detailed analysis, Brainard and Wandell[41] show that in certain circumstances the Retinex model predicts poorly. The method was also extended[42, 43, 44] and recently reformulated as an optimization problem.[45] Historically, the model has been quite controversial and because of different variants and control parameters the possibility of a concrete experience with the algorithm in real scenes has been limited.

As an alternative method aimed at recovering lightness Horn[42] proposes to differentiate the logarithm of the image, ignore small gradients, and then integrate to obtain the "lightness image" up to a constant of integration. For reasons of simplicity we illustrate the process in 1-D, but processing a two dimensional image is similar, except for the integration. Integration in 2-D is less straightforward.[46]

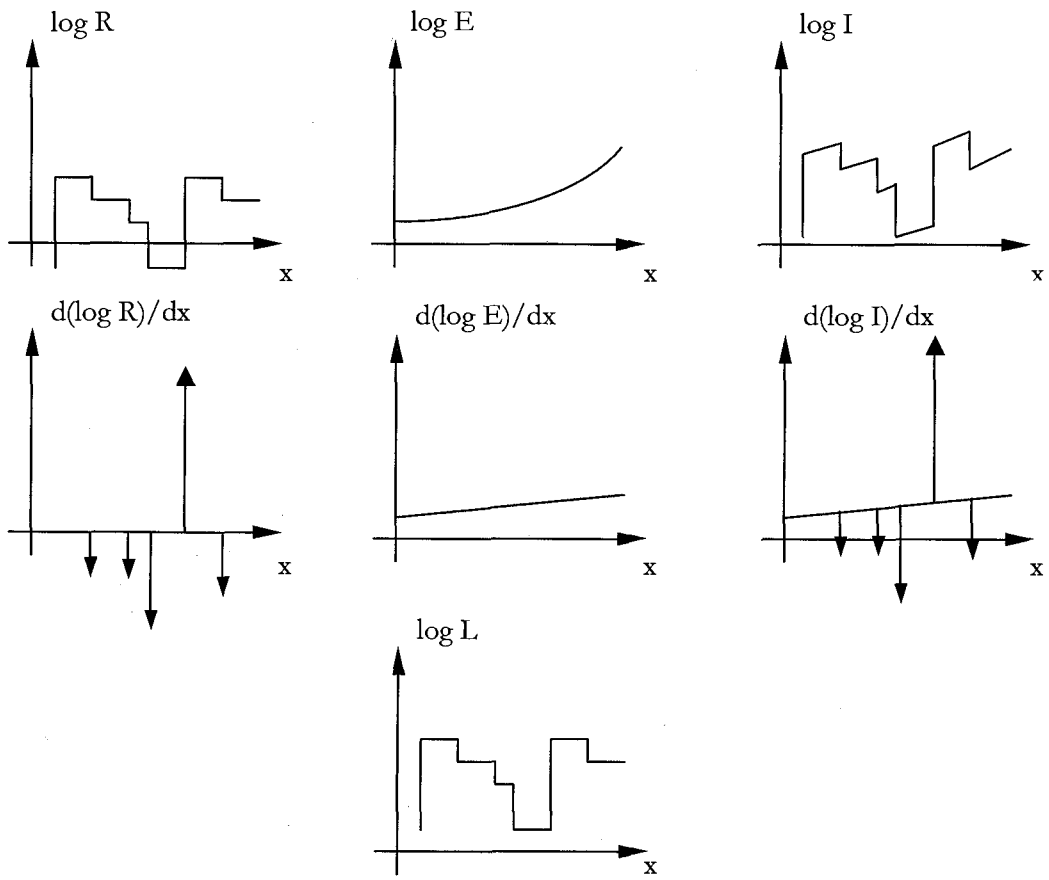


Figure 2. Horn's lightness method of recovering reflectance

Gamut mapping

The gamut mapping approach to colour constancy is due to Forsyth.[47] The algorithm exploits the fact that for a particular imaging device – characterized by its sensor response functions – only certain pixel values are realizable under certain illuminants, given real surfaces. For instance, under a reddish illuminant, a very saturated blue or green are hardly achievable. The algorithm is based on the idea that the shift in the observed colours due to an unknown illuminant is reflected by a related transformation on the colour gamut. The gamut of an image is defined as the collection of all its pixels. It is easy to show that the collection of all realizable pixels under a certain illuminant is a convex set. Indeed, if two pixels P_A and P_B are present in an image under illuminant I_1, then it is possible to observe the pixel whose value is a linear combination of the two: $tP_A+(1-t)\cdot P_B$, for $t \in [0, 1]$. It follows that the image gamut – taken under an unknown illuminant – is a subset of the image convex hull, which is further a subset of the convex hull that would be obtained under all possible reflectances viewed under the same unknown illuminant. After introducing the notion of canonical gamut – the convex hull of all possible surfaces observable through the imaging device under a canonical illuminant, the colour constancy problem is then solved in two steps: 1) identifying the solution set: finding all possible linear mappings M such that the image convex hull can be transformed into the convex hull of the canonical illuminant and 2) choosing a solution from the solution set. There are a variety of strategies for choosing the possible linear mapping. One possibility is to choose the one that maximizes the new gamut. This is justified by the observation that in the majority of the properly balanced scenes, the gamut is maximal. Any colour cast would further reduce the gamut in some way. The most practical implementation of the gamut mapping algorithm uses diagonal transformations for reasons of simplicity.

Forsyth's original gamut mapping algorithm works in the (R, G, B) space. Finlayson et al. propose to work in a chromaticity space[48, 23] and reports better results. A chromaticity space makes the algorithm more robust to varying illumination. The choice of chromaticity space chosen is important because it preserves convexity: (R/G, G/B). The solution is further improved by placing additional constraints on the illuminant set, based on some a priori knowledge about the probability to be encountered in the real world.

The method was further generalized[49] by considering varying illumination which in turn poses additional constraints on the,solution set.

Bayesian colour constancy

This method was introduced by Brainard and Freeman[50] and is based on having some a priori knowledge about the distributions of illuminants and reflectances in real scenes. At the core of the method lies the Bayesian framework:

$$p(x \mid y) = \frac{p(y \mid x) \cdot p(x)}{p(y)} \tag{6}$$

$p(x \mid y)$ denotes the posterior probability of observing x given y, $p(x)$ is the prior probability of x alone, $p(y)$ is the prior probability of y alone, and $p(y \mid x)$ is the probability of observing y given x. The authors formulate the Bayesian approach using a finite dimensional model of surfaces and illuminants. Although interesting, the proposed Bayesian approach is far too complex to be used successfully in real scenes because the dimensionality of the solution is linearly dependent with the number of surfaces in the scene. Another practical limitation of the algorithm is the requirement of having a properly segmented image.

Colour by correlation

A more practical statistical-based method is colour by correlation.[17] In essence, the algorithm is a discrete formulation of the Bayesian method by Brainard and Freeman and combines elements from the gamut mapping approach. At the core of the method lies the notion of a correlation matrix that expresses the interdependence between illuminants and chromaticities in the form of probability distributions computed according to Bayes' rule. Based on observed chromaticities in the image, the illuminant estimation is thus reduced to a voting scheme, in which an illuminant gets a vote from each chromaticity that has a non-zero entry for it (that is observable under the given illuminant).

The correlation approach simplifies Bayes' rule, by assuming that all illuminants and all chromaticities are equally likely.

$$p(E \mid c) = \frac{p(c \mid E) \cdot p(E)}{p(c)} \cong p(c \mid E) \qquad (7)$$

In this equation, p(E) and p(c) represent the prior probabilities of observing the illuminant E and the chromaticity c alone. Therefore, with the further assumptions mentioned above, the probability that the illuminant is E given that we observe chromaticity c $p(E \mid c)$ is the same as $p(c \mid E)$, the probability that we see the chromaticity c under illuminant E.

The above formula applies to each chromaticity observed in the current scene, for which the illuminant E is to be determined. An important further assumption is being made in the computation of the joint probability $p(E \mid C_{im})$, the probability that the illuminant is E when we observe the set of chromaticities C_{im} in the image:

$$p(E \mid C_{im}) = k \cdot \prod_{c \in C_{im}} p(c \mid E) \qquad (8)$$

The joint probability is computed as a product of all separate probabilities, assuming independency. In reality, we expect that the probabilities of chromaticities to be non-independent in real scenes. This is an important limitation of the model.

Recent experiments[51] have showed that working in the full (R, G, B) space as opposed to the (r, g) chromaticity space leads to better estimations. A unique and useful feature of the colour by correlation algorithm with respect to illuminant estimation is an estimate of the confidence in the solution, as justified by the probability correlation-matrix.

Given its simplicity, the colour by correlation algorithm is fast, as it is based on a look-up-table. However, its accuracy can be improved by computing actual prior probabilities and further more, by including models of joint probabilities in the computation of the overall probability that the scene illuminant is E, $p(E \mid C_{im})$.

Neural network for colour constancy

Good results have been reported by a neural network approach to colour constancy.[52, 53, 54] An immediate advantage of a neural network approach is the possibility of modelling non-

linear formulations of the colour constancy problem. The neural network is fed with a binary version of the chromaticity histogram of the scene to be colour corrected. The network is trained on a number of calibrated images in which the illuminant is documented, typically by using a photographic gray card carefully placed in the scene. In the absence of a large calibrated dataset, an alternative method is also proposed in which images are colour corrected by the means of a simple gray world algorithm and then used as a training dataset. For this purpose, selected images that include a large number of objects and colours are used, for which the simple gray world assumption holds and the algorithm works properly.

An important aspect of the method is the fact that, by feeding the network with image histograms, a model of joint probabilities is inherently included, unlike in the colour-by-correlation approach.

Methods using higher-order statistics

Recent experiments[55, 56, 57] suggest that the human visual system uses some higher-order statistics inferred from the current scene in order to compensate for the colour of the illuminant and achieve colour constancy. Second-order statistics of the scene such as the interplay between surfaces and illuminants could yield additional information about the illuminant chromaticity. Simple statistics such as the ones employed by the gray world assumption could not distinguish between a reddish room in white illumination and a white room under reddish illumination. The redness-luminance correlation computed across all pixels in the scene could be a distinguishing factor between the two scenes. Because of the distribution of the natural illuminants on the red-blue axis (often referred to as "warm" for reddish versus "cold" for bluish illuminants), a redness component of the illumination would account for most of its variation in chromaticity for natural light sources.

In probability theory, the correlation (or correlation coefficient) between two variables is defined as the ratio of their covariance by the product of their standard deviation. The correlation can vary from 1 in case of an increasing linear relationship to -1 in case of a decreasing linear relationship. A zero correlation indicates that the two variables are independent.

In essence, the redness-luminance correlation is supposed to give insight as to the colour of the illuminant, much the same as the gamut of observable colours under the current scene gives clues about the departure from the canonical gamut.[58] In fact, Mausfeld and Andres[57] propose that the mean and the covariance matrix, as first and second-order statistics respectively, may give valuable information about the shape and form of this gamut. Experiments by MacLeod and Golz[55,56] indicate that the luminance-redness correlation could be used, together with the mean of the sensor responses across the image, to disambiguate between reddish scenes under white light and white scenes under reddish light (Figure 3).

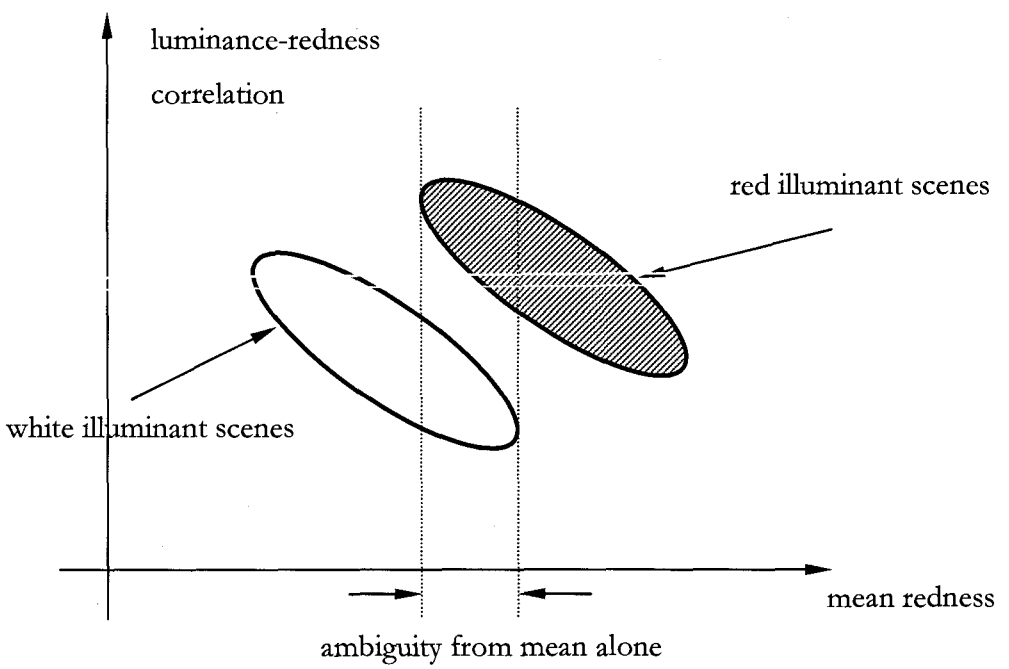


Figure 3. Redness-luminance correlation used in conjunction with the mean redness to determine the illuminant redness

In a two-dimensional space defined by the luminance-redness correlation and the mean image redness, the probability distribution of the white illuminant scenes can actually be distinct from the probability distribution of the red illuminant scenes. On the mean redness axis alone, there is a region where the two probability distributions overlap.

Methods based on specular reflections

The algorithms presented so far assume that the scenes are composed of matte surfaces, that is, that the spectral reflectance of the objects is independent of the incident angle of illumination – or the interface component can be neglected. This is usually a safe assumption, although occasional specularities present in scenes can raise problems. For these algorithms, specular reflections are treated as noise, simply because they do not correlate well with the body reflectance alone.

In real scenes, specular reflections are encountered, and can even provide important cues to the nature of the illuminants. A simple approach to colour constancy based on specular reflections was introduced by Lee.[59] At the heart of Lee's method lies the dichromatic model of reflectance and the observation that the set of possible chromaticities due to a certain surface that are visible in a scene can be expressed as a linear combination of the chromaticity of the body reflection and that of the interface reflection. Furthermore, in a chromaticity space, each distinct dielectric surface in a scene is defined by a line segment. One end of each of these line segments is defined by the chromaticity of the interface reflection – which is in fact the illuminant chromaticity. The success of the method relies in finding the intersection of the light segments corresponding to dielectric surfaces, much like the Hough transform. Geometrically, two unique surfaces can identify the illuminant, unless the two corresponding line segments are collinear. In practice, more surfaces are desired because of inherent noise.

The method is inherently dependent on correctly identifying the surfaces that contain specular reflections. The idea of finding specular reflections in order to infer the colour of the illuminant is first mentioned by Judd.[12] Other authors[60, 61] have proposed algorithms for colour constancy based on specular highlights.

Integrative theories

There is surprisingly limited work in combining several different approaches to computational colour constancy: A recent method for illuminant estimation based on multiple cues in the scene has been proposed by Maloney.[62] The basic assumption of an integrative method to colour constancy is that the human visual system may be using

different cues in determining the illuminant, with different weights associated to each of these cues that reflect their usefulness in actually determining the illuminant colour in a given scene. For instance, in a scene without specular highlights, the visual system would set the weight to such corresponding cue to zero, trying to extract useful information from other cues. A mechanism for cue promotion and dynamic re-weighting of cues is provided in the model. Although some experimental studies are presented on synthetic images, the results presented by the author are not very convincing that this is in fact the way the human visual system performs illuminant estimation.

CHAPTER THREE*
CHROMATIC ADAPTATION TRANSFORMS WITH TUNED SHARPENING

Traditionally, the chromatic adaptation transforms used in colour appearance models are calculated from the tristimulus values of the reference and test illuminants. However, with modern spectroradiometers it is just as easy to measure the spectral power distributions of the illuminants as their tristimulus values so there is no reason to restrict the input parameters to tristimulus coordinates. In this chapter we propose a new method of calculating chromatic adaptation transforms based on using the extra information available in the illuminants' spectral power distributions. The new method gives comparable results to the current tristimulus-based chromatic adaptation transforms in most cases, and better results in some specific situations such as the McCann-McKee-Taylor experiment.[1]

Introduction

While existing chromatic adaptation transforms[29,63,64] start with the CIE XYZ tristimulus values of the test and reference illuminants as input, the first processing step is a change of basis to a new coordinate system. One such change of basis is the Bradford transform empirically derived by Lam[29] and another is the spectral sharpening transform derived from Lam's data[29] by Finlayson and Süsstrunk[65] using white-point preserving sharpening. In either case, the same change of basis is applied no matter what the illuminants happen to be. If, instead of restricting the description of the illuminants to the tristimulus values, we describe them in terms of their spectral power distributions, we then can derive an illuminant-specific sharpening transformation. The hypothesis is that tuning the change of basis for each particular illuminant pair will lead to smaller errors in the final chromatic adaptation transform.

---

* This chapter also appears as a published paper: Brian Funt and Florian Ciurea, "Chromatic Adaptation Transforms with Tuned Sharpening", in Proc. First European Conf. on Color in Graphics, Imaging and Vision, 148-152, 2002, Poitiers, France.

Description of the method

Using the 336 Kodak[33] database of spectral surface reflectances which includes 102 DuPont paint chips, 64 Munsell chips (matte collection) and 170 natural and man-made objects, we compute the tristimulus values for each reflectance in the database, under the test and reference illuminants in the standard manner:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = k\sum_{i=1}^{101} E(\lambda_i)\cdot S(\lambda_i)\cdot \begin{bmatrix} \bar{x}(\lambda_i) \\ \bar{y}(\lambda_i) \\ \bar{z}(\lambda_i) \end{bmatrix}, \text{ where } k = \frac{100}{\sum_{i=1}^{101} E(\lambda_i)\cdot \bar{y}(\lambda_i)} \tag{9}$$

The sampling has been performed over 101 wavelengths λ_i in the range from 380 nm to 780 nm in 4 nm intervals. $E(\lambda_i)$ denotes the spectral power distribution of the illuminant and $S(\lambda_i)$ represents the surface reflectance at wavelength λ_i.

We find the best transformation T mapping the tristimulus values obtained under the test illuminant to the corresponding tristimulus values under the reference illuminant using the white-point preserving algorithm described by Finlayson and Süsstrunk.[65] This transformation can be used to predict the corresponding tristimulus values X^{ref}, Y^{ref}, Z^{ref} under the reference illuminant from the tristimulus values X^{test}, Y^{test}, Z^{test} under the test illuminant using the chromatic adaptation transformation model:

$$\begin{bmatrix} X^{ref} \\ Y^{ref} \\ Z^{ref} \end{bmatrix} = T^{-1}\begin{bmatrix} R_w^{ref}/R_w^{test} & 0 & 0 \\ 0 & G_w^{ref}/G_w^{test} & 0 \\ 0 & 0 & B_w^{ref}/B_w^{test} \end{bmatrix} T \begin{bmatrix} X^{test} \\ Y^{test} \\ Z^{test} \end{bmatrix} \tag{10}$$

where the RGB ordinates of the reference and test illuminants are computed respectively as:

$$\begin{bmatrix} R_w^{test} \\ G_w^{test} \\ B_w^{test} \end{bmatrix} = T\cdot \begin{bmatrix} X_w^{test}/Y_w^{test} \\ Y_w^{test}/Y_w^{test} \\ Z_w^{test}/Y_w^{test} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} R_w^{ref} \\ G_w^{ref} \\ B_w^{ref} \end{bmatrix} = T\cdot \begin{bmatrix} X_w^{ref}/Y_w^{ref} \\ Y_w^{ref}/Y_w^{ref} \\ Z_w^{ref}/Y_w^{ref} \end{bmatrix} \tag{11}$$

The white-point preserving least-squares regression used in[65] was introduced by Finlayson and Drew[66] and is summarized below:

Let us denote S the 336×3 matrix of 336 XYZ values under the reference illuminant and P the 336×3 matrix of 336 XYZ values under the test illuminant. The transformation T is obtained through eigenvector decomposition of the general matrix M that best maps P to S.

To preserve achromatic colours, M is derived using:

$$M = D + Z \cdot N \tag{12}$$

where Z is a 3×2 matrix of any two vectors orthogonal to XYZ_{test}, D is XYZ_{ref}/XYZ_{test} (diagonal matrix of the ratio of the two white point vectors) and N is given by:

$$N = \left[Z^T \cdot P^T \cdot P \cdot Z\right]^{-1} \cdot \left[Z^T \cdot P^T \cdot S - Z^T \cdot P^T \cdot P \cdot D\right] \tag{13}$$

The main difference between the approach proposed here and that of Finlayson and Süsstrunk[65] is that we tune the sharpening transform T for each illuminant pair. We compute XYZ values from illumination spectral information and the database of 336 natural reflectances as opposed to computing a generic transformation T using the XYZ values of the 58 samples reported by Lam.

Testing the model

To the our knowledge, most existing corresponding colour appearance experiments[29,67-71] do not document the actual spectral power distribution of the test and reference illuminants. One exception is the McCann-McKee-Taylor experiment[1] in which 5 illumination conditions were provided by three projectors with narrow-band filters (630 nm, 530 nm and 450 nm, respectively) having a bandwidth of 10 nm at half-height. Given that the rest of our computations were based on a 4-nm sampling interval, we modeled each peak in the McCann-McKee-Taylor spectra by a simple rectangular signal with a width of 8 nm (2 intervals). The relative power distribution of each illuminant was then computed to accord with the ratios of the reported triplet of radiances from the Munsell white paper for that illuminant. This presumes the Munsell white paper had a uniform surface spectral reflectance.

In most of the other experiments the only information available about the light is the illuminant type (i.e. CIE A, CIE D65, Philips TL84) and its chromaticity. Our proposed method depends on having the actual spectra of the two illuminants, not just their chromaticities. In the present circumstances, the best we can do is to estimate of the actual spectra of the illuminants used in each experiment. We started with the spectral power distributions of the standard sources used to simulate the ideal illuminant spectra A and D65, and measured the Philips TL84 with a PhotoResearch PR650 spectrophotometer. We then modified these spectra to match the chromaticities reported for the illuminants actually used in the experiments, as described below.

To obtain the least distortion in the illuminant spectral power distribution when matching the chromaticity, we solve the following under-determined system with (E^0–E^1) as unknown:

$$\begin{cases} \sum_{i=1}^{101} \bar{x}(\lambda_i) \cdot (E_i^0 - E_i^1) = X_w^0 - X_w^1 \\ \sum_{i=1}^{101} \bar{y}(\lambda_i) \cdot (E_i^0 - E_i^1) = Y_w^0 - Y_w^1 \\ \sum_{i=1}^{101} \bar{z}(\lambda_i) \cdot (E_i^0 - E_i^1) = Z_w^0 - Z_w^1 \end{cases} \tag{14}$$

E^0_i is the spectral power distribution at wavelength λ_i of the measured illuminant with tristimulus values X^0_w Y^0_w, Z^0_w, and E^1_i is the corresponding spectral power distribution of the similar illuminant that has the tristimulus values X^1_w Y^1_w, Z^1_w, as reported in the corresponding experiment. When we solve the under-determined system above using the pseudoinverse method we obtain the solution having the smallest norm, thus having the least deviation from the measured spectral power distribution E^0.

Results on corresponding colour datasets

We have used the corresponding colour datasets accumulated by Luo and Hunt, available online from the University of Derby.[72] Table 1 summarizes the experimental conditions.

Data set	Number of specimens	Test illuminant type	Reference illuminant type	Experimental method
Helson	59	A	D65	Memory
Lam & Rigg	58	A	D65	Memory
Lutchi (A)	43	A	D65	Magnitude
Lutchi (D50)	44	D50	D65	Magnitude
Lutchi (WF)	41	WF	D65	Magnitude
Kuo & Luo (A)	40	A	D65	Magnitude
Kuo & Luo (TL84)	41	TL84	D65	Magnitude
Breneman 1	12	A	D65	Magnitude
Breneman 2	12	PROJECTOR	D55	Magnitude
Breneman 3	12	PROJECTOR	D55	Magnitude
Breneman 4	12	A	D65	Magnitude
Breneman 6	11	A	D55	Magnitude
Breneman 8	12	A	D65	Magnitude
Breneman 9	12	A	D65	Magnitude
Breneman 11	12	D55	GREEN-B	Magnitude
Breneman 12	12	D55	GREEN-B	Magnitude
Braun & Fairchild 1	17	D65	D65	Matching
Braun & Fairchild 2	16	D65	D65	Matching
Braun & Fairchild 3	17	D93	D65	Matching
Braun & Fairchild 4	16	A	D65	Matching
McCann "blue"	17	BLUE	GREY	Haploscopic
McCann "green"	17	GREEN	GREY	Haploscopic
McCann "grey"	17	GREY	GREY	Haploscopic
McCann "red"	17	RED	GREY	Haploscopic
McCann "yellow"	17	YELLOW	GREY	Haploscopic

Table 1. Experimental conditions for each corresponding colour experiment data set

The results are presented in the following tables:

Data set	Mean ΔE_{Lab} Lam sharp	Mean ΔE_{Lab} Spectra sharp	RMS ΔE_{Lab} Lam sharp	RMS ΔE_{Lab} Spectra sharp
Helson	5.3	6.4	6.1	7.6
Lam & Rigg	4.4	5.8	5.1	6.8
Lutchi (A)	6.8	8.5	7.6	9.6
Lutchi (D50)	6.3	6.6	6.8	7.1
Lutchi (WF)	7.8	5	8.7	5.8
Kuo & Luo (A)	6.9	7.7	7.7	8.5
Kuo & Luo (TL84)	4.3	4.4	4.7	4.9
Breneman 1	10.5	11.2	10.8	12.1
Breneman 2	7.1	7.8	7.4	8.1
Breneman 3	12	13.8	14.2	16
Breneman 4	12.3	15.7	14.9	17.6
Breneman 6	7.9	9.6	8.3	10.4
Breneman 8	12	15.5	14	16.8
Breneman 9	17.9	22.1	20.7	24.7
Breneman 11	7.4	6.8	8.2	7
Breneman 12	8.9	9.6	9.1	10.2
Braun & Fairchild 1	3.8	3.4	4	3.8
Braun & Fairchild 2	5.9	6	6.6	6.6
Braun & Fairchild 3	7.1	7.2	7.2	7.3
Braun & Fairchild 4	5.9	6.3	6	6.5
McCann "blue"	21.4	18.7	22.2	19.8
McCann "green"	27.7	23.2	29.7	24.4
McCann "grey"	10.1	10.1	11.2	11.2
McCann "red"	16.9	20.2	17.8	21.6
McCann "yellow"	26.2	19.4	29.6	21
average value	**10.5**	**10.8**	**11.5**	**11.8**

Table 2. Mean and RMS ΔE_{Lab} error of the full spectra method compared to the sharpening method based on Lam data. Student's t-test significance is 0.85 for the average of mean ΔE_{Lab} and 0.89 for the average RMS ΔE_{Lab}, so the differences in the results are not statistically significant.

Data set	Mean $\Delta E_{CMC(1:1)}$	Mean $\Delta E_{CMC(1:1)}$	RMS $\Delta E_{CMC(1:1)}$	RMS $\Delta E_{CMC(1:1)}$
Helson	4.4	4.9	5.5	6.2
Lam & Rigg	3.7	4.7	4.6	5.5
Lutchi (A)	4.7	5.5	5.2	6.1
Lutchi (D50)	4.2	4.5	4.7	4.9
Lutchi (WF)	5.2	3.2	5.7	3.6
Kuo & Luo (A)	4.9	5.8	5.4	6.3
Kuo & Luo (TL84)	3.2	3.5	3.7	4
Breneman 1	7.2	8	7.8	9.1
Breneman 2	4.9	5.1	5.5	5.8
Breneman 3	7.5	8.4	9.4	10.1
Breneman 4	8.6	10.7	10.7	12.2
Breneman 6	5.8	6.7	6.2	7.3
Breneman 8	8.2	10.1	9.7	11.3
Breneman 9	12.3	14.6	14.1	16.1
Breneman 11	5.1	5	5.6	5.4
Breneman 12	6	6.6	6.7	7.6
Braun & Fairchild 1	3.5	3.2	3.9	3.8
Braun & Fairchild 2	6.4	6.4	7.9	7.7
Braun & Fairchild 3	5.7	5.6	6.1	6
Braun & Fairchild 4	5.2	5.3	5.4	5.6
McCann "blue"	12.7	11.6	13.2	12.3
McCann "green"	18	16	20.2	18.6
McCann "grey"	7.1	7.1	8.6	8.6
McCann "red"	13	13.9	14.5	15.2
McCann "yellow"	15.8	14.7	17.4	17.7
average value	**7.3**	**7.6**	**8.3**	**8.7**

Table 3. Mean and RMS $\Delta E_{CMC(1:1)}$ error of the full spectra method compared to the sharpening method based on Lam data. Student's t-test significance is 0.78 for the average of mean $\Delta E_{CMC(1:1)}$ and 0.77 for the average RMS $\Delta E_{CMC(1:1)}$, so the differences in the results are not statistically significant.

Data set	Mean ΔE_{CIE94} Lam sharp	Mean ΔE_{CIE94} Spectra sharp	RMS ΔE_{CIE94} Lam sharp	RMS ΔE_{CIE94} Spectra sharp
Helson	3.5	3.8	4.1	4.5
Lam & Rigg	2.9	3.8	3.4	4.3
Lutchi (A)	3.9	5.2	4.4	5.9
Lutchi (D50)	3.5	3.7	3.9	4.1
Lutchi (WF)	4.1	3.1	4.4	3.5
Kuo & Luo (A)	4.2	5.1	4.5	5.5
Kuo & Luo (TL84)	2.7	3	2.9	3.3
Breneman 1	5.8	6.8	6.2	7.3
Breneman 2	3.9	4.2	4.3	4.6
Breneman 3	5.8	6.5	7.3	7.8
Breneman 4	7	8.6	8.4	9.7
Breneman 6	4.7	5.8	4.9	6.2
Breneman 8	6.7	8.2	7.6	9
Breneman 9	9.5	11.3	10.8	12.3
Breneman 11	4.2	4.1	4.6	4.4
Breneman 12	4.6	5.2	5.1	5.7
Braun & Fairchild 1	2.8	2.7	3.1	3
Braun & Fairchild 2	4.5	4.5	5.1	5.2
Braun & Fairchild 3	4.6	4.6	4.8	4.8
Braun & Fairchild 4	4.2	4.2	4.3	4.3
McCann "blue"	11.3	10.5	12	11.3
McCann "green"	14.5	12.8	15.6	14.3
McCann "grey"	6.1	6.1	7.1	7.1
McCann "red"	10.6	11.3	11.4	12
McCann "yellow"	13.7	12.4	14.9	13.8
average value	**6.0**	**6.3**	**6.6**	**7.0**

Table 4. Mean and RMS ΔE_{CIE94} error of the full spectra method compared to the sharpening method based on Lam data. Student's t-test significance is 0.72 for the average of mean ΔE_{CIE94} and 0.73 for the average RMS ΔE_{CIE94}, so the differences in the results are not statistically significant.

For most of the data sets, our method, denoted as "spectra sharp", performs almost as well as the sharpening method by Finlayson and Süsstrunk[65] based on Lam data (noted here as "Lam sharp") which is considered to give among the best results on these data sets.[73] For the Lutchi (WF), Breneman 11, Braun and Fairchild 1 and McCann "blue", McCann "green" and McCann "yellow" data sets, our new method gives more accurate predictions. We have used three different error metrics: ΔE_{Lab} , $\Delta E_{CMC(1:1)}$ and ΔE_{CIE94} to evaluate the predictions. Student's t-tests for the samples conclude that the differences in the results of the two methods are not statistically significant.

We speculate that the somewhat better overall results obtained with the full spectra method on the McCann data sets are due to the fact that it contains the actual spectral power distribution of the 5 rather unusually narrow-banded illuminants which differ substantially from the illuminants in the Lam experiment. For the other cases, the full spectra method performs almost as well as the Lam-based sharpening. In fact, given the significant noise in experimental data, which is based on memory matches, magnitude estimation and halposcopic matching, the performance difference between the models is probably not significant. Furthermore, it must be remembered that we have only an imprecise specification of the required illuminant spectra.

We also considered the possible effect of incomplete adaptation[63] by solving for the optimal value of the incomplete adaptation factor, D. While the errors for both methods dropped slightly, the overall results found were qualitatively similar to those in Table 2, Table 3 and Table 4.

Conclusions

The main hypothesis of this chapter is that a better chromatic adaptation transform could be developed if it were to be computed from the spectra of the illuminants rather than simply from their tristimulus values. In the proposed method, the spectral power distribution of the illuminants is used to derive a sharpening transformation that is specific to the adapting illuminant pair. This contrasts with the fixed transformation approach inherent in either the Bradford transform or the Finlayson-Süsstrunk sharpening as it is applied in most current chromatic adaptation transforms. The new full spectra illumination-specific sharpening method performs better on the McCann-McKee-Taylor data, which is the main case where we should definitely expect some improvement. Many of the other experiments involve the same illuminant pairs (A and D65) as used in the Lam-based sharpening or ones similar to them.

Two exceptions are the Kuo & Luo TL84-D65 pair and the Lutchi D65-WF. In the former, the errors are both very small relative to the experimental noise; in the latter, the full spectra method does slightly better. These results indicate, but do not prove, the potential of using the full spectra of the illuminants as input parameters for a better chromatic adaptation

transform. Since chromatic adaptation transforms are a crucial part of all colour appearance models, this could also lead to improved predictions of colour appearance. Unfortunately, information about the actual spectral power distribution of the illuminants is lacking for most of the existing experiments. This means that at this point, it is difficult to evaluate conclusively the relative performance of the method. Still, future experiments to obtain corresponding colours under different illuminants should definitely include a record of the spectra of the illuminants involved.

CHAPTER FOUR*
RETINEX IN MATLAB™

Many different descriptions of Retinex methods of lightness computation exist. This chapter provides concise MATLAB™ implementations of two of the spatial techniques of making pixel comparisons. The code is presented along with test results on several images and a discussion of the results. We also discuss the calibration of input images and the post-Retinex processing required to display the output images.

Introduction

The Retinex model+ for the computation of lightness was introduced by Land and McCann.[37] Since that time Land and his colleagues have described several variants on the original method.[1,16,39,40,74] The variants on Retinex mainly aim to improve the computational efficiency of the model while preserving its basic underlying principles.

Retinex calculations aim to predict the sensory response of lightness. It is important to distinguish between physical reflectance, the sensation of lightness, and perceived reflectance, which are three distinct entities. A single model can attempt to calculate only one of the three – the Retinex goal is to calculate the sensation of lightness. Consider the case of two faces of a white cube, one in direct sunlight and the other in shadow. Physical reflectance is a measure of a property of the cube's surface relating its radiance to its irradiance.

The reflectances of the two faces are identical. Sensations, on the other hand, are the appearances of the faces of the cube in the sun and the shade. To create the same appearances in a painting, a fine arts painter would mix white with a little yellow to make the

---

* This chapter also appears as a published paper: Brian Funt, Florian Ciurea, John McCann, "Retinex in Matlab", Journal of the Electronic Imaging, Vol. 13, No. 1, 48-57, 2004.

+ McCann refers to these models as Ratio-Product-Reset-Average, but for simplicity here we call these operations the Retinex model. Frankle and McCann34 provide complete FORTRAN code for their algorithm with extensive discussion of image processing steps that follow spatial comparisons.

sunny face, but use white with blue and a little black to reproduce the appearance of the shadowed face. These samples of different coloured paints are measures of sensation. Here the two faces are different.[75] In comparison, the question of the perceived reflectances of the cube's surfaces involves cognition. It asks the observer to recognize the paint on the cube. Asked to repaint the cube, the observer is not confused by sun and shade, and would simply apply white paint. In terms of perception, the two faces of the cube are identical. In contrast, Retinex calculates lightness sensations – it cannot be used to calculate physical reflectances or perceived reflectances.

The first model designed to calculate lightness was described in Land's Ives Medal Address to the Optical Society of America in 1968 and later published.[37] This lecture included a working demonstration of a primitive electronic Retinex camera. This was followed by publications and patents with additional details and improvements.[39,76,77] McCann McKee and Taylor[1] described a study of human colour constancy that included colour-matching experiments, the details of the lightness model and successful results of modelling the experimental data. This result was further developed to show that there is no effect of cone pigment adaptation in colour constancy.[78] The Retinex operators were selected for simplicity to mimic biological operators that sum, difference and rectify input signals to obtain spatial interactions.

Dynamic range compression of real images was described in a patent by Frankle and McCann.[39] This implementation used specialized hardware (International Imaging Systems I²S image processor with scrollable 8-bit image planes) for efficient image calculation. It described the idea that information from 2^n pixels is accumulated after n steps of the process. This patent also described the multi-resolution approach to Retinex calculation used for computer applications.[16,79]

Appropriate Input Data

For quantitative testing of the Retinex model it is crucial that the data be calibrated in the sense that the image digital values must be a logarithmic function of scene radiance and they must be represented with sufficient precision. McCann[80] used a slope 1.0 photographic film to capture real images (Ektachrome 5071 slide duplicating film). He was able to measure an

in-camera dynamic range of 3.5 log units. The importance of the logarithmic function follows from Wallach's experiments on appearance.[81] He showed that equal radiance ratios generate equal lightness differences. A pair of papers, i.e., a 20% gray paper and a 100% white paper, have the same lightness difference in sun and shade. The pair also has a $\log_{10}$ edge difference of 0.7, regardless of illumination. If the input image data deviates from logarithmic, then the $\log_{10}$ edge difference for these papers will change with illumination, and the calculated lightness difference of the pair will change. For Retinex to work well, edge ratios, or $\log_{10}$ differences, within an object must be independent of illumination. Accurate logarithmic calibration guarantees this to be the case.

Illumination	Paper	Scene Radiance	Log Radiance	Quantized Log Digit	Quantized Linear Digit	Log Quantized Digit
Sun	White	3162	3.50	255	255	255
	Light Gray 1	1412	3.15	229	114	229
	Gray 2	631	2.80	204	51	204
	Mid Gray 3	282	2.45	178	23	179
	Dark Gray 4	126	2.10	153	10	152
	Black	56	1.75	127	5	131
Shade	White	56	1.75	127	5	131
	Light Gray 1	25	1.40	102	2	102
	Gray 2	11	1.05	76	1	80
	Mid Gray 3	5	0.70	51	0	0
	Dark Gray 4	2	0.35	25	0	0
	Black	1	0.00	0	0	0

Table 5. Preparing input images for Retinex processing

Table 5 describes the care one must take in preparing input images. The data comes from the image of two test targets: one in sun, the other in shade (See Figure 4). The shade reduced the illumination such that white paper in the shade sends the same radiance to the eye as the black paper in the sun. Columns four and five demonstrates the digitization of raw image data as equally spaced $\log_{10}$ increments. In other words, convert the scene into log

radiance and then quantize to 8 bit (0 to 255) digits (log then quantize). The first column specifies either sun or shade illumination. The second column describes the papers in the gray scale. The third column lists the Scene Radiance from the two identical gray scales in sun and in shade. Note that the radiance from the black in the sun is equal to that from the white in the shade. The fourth column lists Log Radiances of Scene Radiance values (column 3). The fifth lists the eight-bit Quantized Log Digits for the values in column four scaled in the 0-255 range. Column four now represents equal log increments with constant difference of 0.35 log units in radiance. Columns six and seven demonstrate the problems arising from quantizing before converting to log. The sixth column lists the eight-bit Quantized Linear Digit (from values in column three scaled in the 0-255 range). The seventh column lists the Log Quantized Digit. Here, values are computed by first rescaling values in the sixth column to the range of radiances in column three, taking the log and then rescaling to the 0-255 range, similarly to how the fifth column is computed. The consequence of the quantization prior to taking the logarithm is that, in column seven, all radiance values for Black, Dark Gray 4 and Mid Gray 3 are all represented by the same digit, 0. In other words, quantizing the log input image shows poor use of digits. Following quantization with a $\log_{10}$ transform does not improve the image. Representing radiances of the input image as Log Quantized Digit (log then quantize) makes a suitable input image for studying high dynamic range images. Using Log Quantized Digits (quantize then log) makes a highly undesirable input image. One cannot take an existing 8-bit image, apply a log to it and have meaningful input image data for quantitative testing the Retinex model.

Nevertheless, Retinex often enhances random images that have unknown and unknowable radiances for inputs.[39,82,83] The process improves the visibility of dark objects while maintaining the visual discrimination of the light areas. Unlike lookup tables, which improve one range of radiance at the expense of others, Retinex improves visual differentiation in all ranges of radiances. The danger is that artefacts such as noise create artificial edge information that is enhanced by Retinex processing. The ability to bring out shadow detail is limited by image noise.

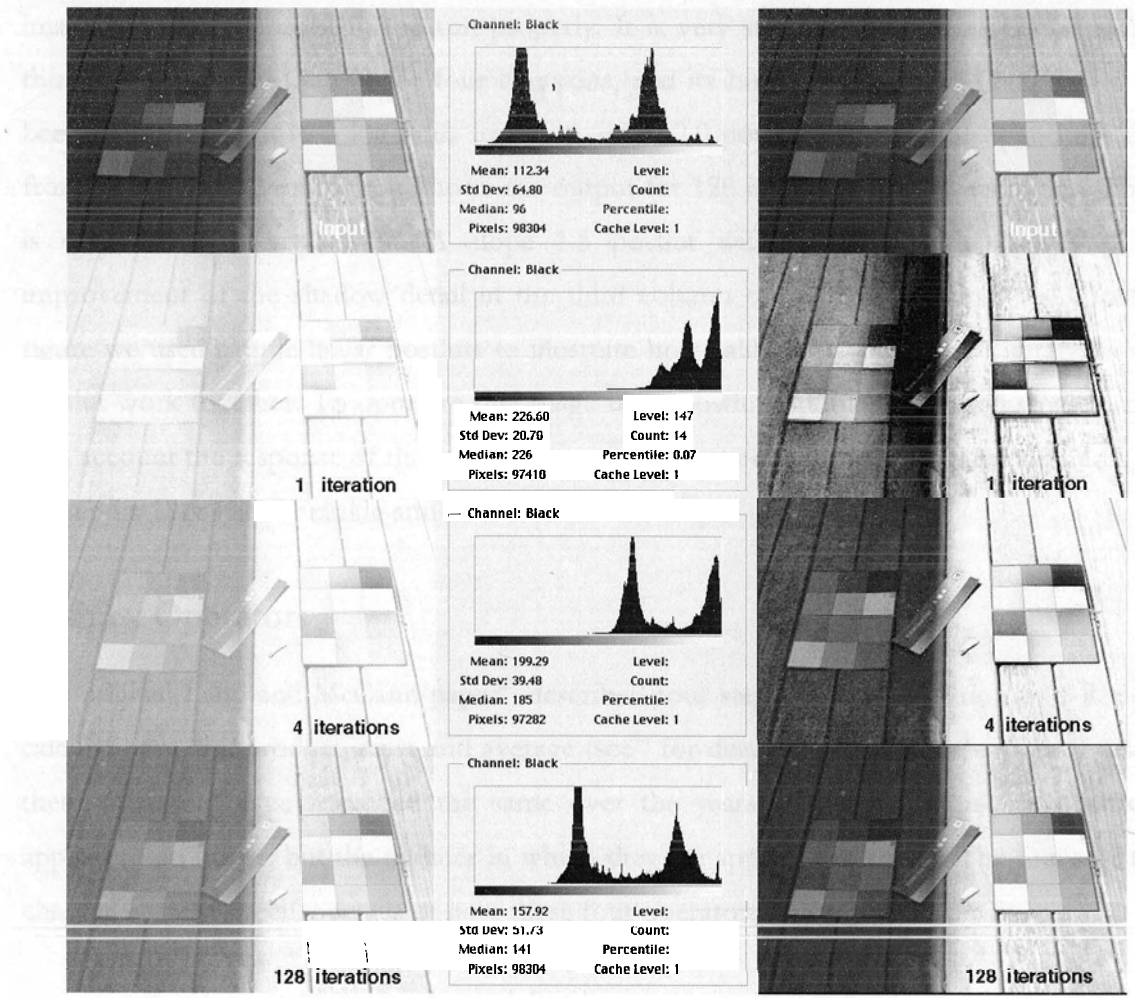


Figure 4. The role of number of iterations and post-LUT in Retinex processing

Figure 4 is due to John McCann and demonstrates the role of number of iterations and postluts. The first column shows the effect of spatial comparisons (ratio-product-reset-average). The second column is the histogram of the images in the first column. The third column shows images that have been stretched by a different postlut for each number of iterations. The first row shows the input $\log_{10}$ image scaled so that 3.5 $\log_{10}$ units covers 0-255. The sun half of the image is on the right and the shade half is on the left. The shade image is a lower radiance copy of the sun image. The histogram of this image is in the second column. The third column image is the same as first column, illustrating that it has a slope 1.0 postlut. Output equals input. The second row shows an output image using one iteration, with its histogram. Here the output dynamic range has been compressed into the top 25% of the 0-255 digit range. A slope 4.0 linear postlut will stretch the first column

image to render contrast in the sun properly. It is very steep and generates artefacts. The third row shows the output for four iterations, and its histogram. Here the range data has been compressed from 3 log units to 1.5. A slope 2.0 postlut has only to expand the data from 128 to 0. The fourth row shows the output for 128 iterations and its histogram. There is only a 25% compression. A slope 1.5 postlut will be very gentle; however, the improvement of the shadow detail in the third column output image is minimal. In this figure we used simple linear postluts to illustrate how calibration, number of iterations and postlut work together. To optimize the image these postluts should be shaped so as to take into account the response of the output device and the tone reproduction curve desired. (See Appendix II & III of Frankle and McCann for details[39]).

Retinex Operators

The original Land and McCann paper[37] described four steps for each iteration of a Retinex calculation: ratio, product, reset and average (see[84] for details). With the exception of reset[40] these operators have remained the same over the years. These operators are iteratively applied to an image, but the manner in which they are applied has varied. The focus of this chapter is to list specific details of how these four operators are applied to the image.

A fundamental concept behind Retinex computation of lightness at a given image pixel is the comparison of the pixel's value to that of other pixels. The main difference between the Retinex algorithms is the way in which the other comparison pixels are chosen, including the order in which they are chosen. They use the same calculations but have dramatically different computational efficiencies in dealing with large real images. The original way of defining comparisons is by following a path, or set of paths, from pixel to neighbouring pixel through the image.[37] Lightness estimates are accumulated along the path in a "sequential product" SP. SP starts as 1 and then is modified by multiplying it with the ratio of the next pair of pixels along the path. In the case of path following, path length affects the results substantially. Short paths mean the comparison is made only to others in a spatially localized group of pixels. Intermediate path lengths are to be used when modeling human vision. Infinite path lengths result in a degenerate case in which the output image is simply a scaled version of the input image. Infinite path lengths should not be used to model vision.[41,85]

A "reset" step is a second important feature of Retinex. Each time a comparison is made the SP is tested; if it exceeds 1.0, it is reset to 1.0. In this case, the value 1.0 becomes the current lightness estimate. A third aspect of Retinex is the way in which lightness estimates obtained from different paths to a pixel are combined. In earlier versions, Retinex also included a "thresholding" step; however, it is not included in later versions[16] and is not part of the MATLAB™ implementations presented. The fourth step averages present values of the Product with previous ones.

Implementations

We have chosen two versions of Retinex to implement. The first is a computer-based version described by McCann,[16] which we will refer to as McCann99 Retinex. The second is an older specialized-hardware version,[39] which we will call Frankle-McCann Retinex. The two versions both replace path following with more computationally efficient spatial comparisons. McCann99 Retinex creates a multi-resolution pyramid from the input by averaging image data. It begins the pixel comparisons at the most highly averaged or top level of the pyramid. After computing lightness on the image at a reduced resolution, the resulting lightness values are propagated down, by pixel replication, to the pyramid's next level as initial lightness estimates at that level. Further pixel comparisons refine the lightness estimates at the higher resolution level and then those new lightness estimates are again propagated down a level in the pyramid. This process continues until New Products have been computed for the pyramid's bottom level.

In comparison, Frankle-McCann Retinex uses single pixel comparisons with variable separations. An important difference between this method and that described in Land and McCann[37] is that there are no paths. A single pixel eventually averages different products from all other pixels. The advantage of this structure, and also for the multi-resolution approach, is that long distance interactions are propagated with fewer comparisons.

McCann99 Multi-Level Retinex Details

For this implementation the input images must be of dimension $w \cdot 2^n$ by $h \cdot 2^n$, where $w \geq h$ and w and h are integers in the range [1..5]. This constraint arises from the fact that each

level of the image pyramid differs from previous levels by a factor of 2 in each dimension. It is not a serious limitation in practice.

The algorithm assumes that input digits are proportional to the logarithm of scene radiance and are of meaningful precision. Using logarithms simplifies the computation of radiance ratios, which become simple differences. It also implies that when results from different spatial comparisons are averaged, the averaging is in log space and hence equivalent to a geometric mean.

In the first step, the log image is averaged down to the lowest resolution level, which depending on the input dimensions will be of the size 1×1, 1×2, 1×3, 2×3, 3×4, 3×5, 4×5 or 5×5. At each step the resolution level will be doubled. The number of layers in the pyramid depends on the size of the input image. The number of layers will be the greatest power of 2 dividing both the width and height of the input images as calculated by the function ComputeSteps.

When the results (called "New Products") at one level of dimension n-by-m have been computed, the values are then replicated to form an "Old Product" image of dimension 2n-by-2m. In our implementation, we pad the Old Product image with zeroes to simplify handling boundary conditions. These extra pixels are discarded at the end of the computation.

At all levels the New Product, a precursor of calculated lightness, for each pixel is computed by visiting each of its 8 immediately neighbouring pixels in clockwise order. Each visit involves a ratio-product-reset-average operation,[16] which is implemented by the function CompareWithNeighbor. It subtracts the neighbour's log luminance (the ratio step) and then adds the result to the old product (the product step). If the result exceeds the maximum defined by Maximum, it is reset to Maximum (the reset step). Finally, the new product for the pixel obtained by comparison to its neighbour is averaged with the previous old product.

A crucial parameter to the McCann99 algorithm is the number of times a pixel's neighbours are to be visited. In the code, this is set by nIterations. It controls the number of

times the neighbours are cycled through, which as a result affects the distance at which pixels influence one another. This occurs because the New Product values for all pixels are being computed in parallel, so that after one iteration all neighbouring pixels have had their New Products values updated. Hence, in the second iteration these new values involve information propagated from beyond a pixel's immediate neighbours. In the limiting case of an infinite number of iterations, the algorithm converges to produce an output image that is simply the input image scaled by the image's maximum value. A practical value for the number of iterations is 4. The final step is to scale the New Product values to make an estimated lightness (see Section "Scaling of Retinex output to desired media and purpose"). In the case of colour images, the function Retinex_mccann99 must be applied to each of the colour channels independently.

The code is based on MATLAB™ 5 (Version 5.1.0.421). For the reader unfamiliar with MATLAB™, the statement IP(IP > Maximum) = Maximum, which sets all values in matrix IP that are greater than Maximum to Maximum, demonstrates an important feature of the language; namely, that most of the functions and operators work on whole matrices applying the given function to all matrix elements.

Frankle-McCann Retinex

As in McCann99 Retinex, Frankle-McCann Retinex computes long-distance interactions between pixels first and then progressively moves to short-distance interactions. In Frankle-McCann, the spacing between the pixels being compared decreases with each step. The direction between pixels also changes at each step, in clockwise order. At each step, the comparison is implemented using the Ratio-Product-Reset-Average operation. The process continues until the spacing decreases to 1 pixel.

The original algorithm assumed the input image to be 512×512. This followed the hardware design of the I²S. As a result, the initial spacing between pixels started at 256. We have generalized the algorithm slightly so that our implementation will work on an image of arbitrary size. In this case, the initial spacing (as encoded by the variable 'shift') is computed as the largest power of 2 smaller than both of the input image dimensions.

The function CompareWith(s_row, s_col) updates the current lightness estimate for a pixel using the ratio-product-reset-average operation described above. In the case of Frankle-McCann, it is based on the pixel located at a distance of s_row, s_col. The square spiral path structure in this implementation means that when this function is called, one of the two parameters will always be zero. The original Frankle and McCann[39] implementation had the option of either square or 8-direction comparisons.

Retinex Parameters

All spatial operators use variable parameters to appropriately match their effects to input images. For example, this is true of unsharp masking, jpeg and Retinex spatial operators.

The purpose of unsharp masking is to change the spatial content in the image, particularly in the high-spatial-frequency components. When successfully used the image looks sharper and free of artefacts. With inappropriate parameters the process will generate artefacts that are visible to the observer. If we compare the effects of a particular unsharp mask on same-size prints of a 256-by-256 digital image with the effects on an 2k-by-2k image, we see that they act very differently. A sharpening filter that is appropriate for the small image will have no effect on large images, while an appropriate filter for the large images will introduce artefacts in small ones. Given a print size and a viewing distance, one can optimize the shape of the filter kernel. The choice of sharpening kernel is selected so as to keep artefacts below visual threshold, which is a function of both spatial frequency,[86] size of the display[87] and light intensity of the display.[88]

An analogous spatial dependence is found in jpeg compression where knowledge of human sensitivity to spatial information is used to reduce the number of bits for rendering a visually similar image.[89] When we select a quality factor, we are controlling an underlying array of coefficients that filter the data so as to reduce the data needed to recreate the image. To make two same-size prints from a 256-by-256 versus a 2k-by-2k image requires different jpeg coefficients. Any reduction in information will likely be visible in the small number-of-pixel image, while the larger image might well be compressed by factors of 10:1 or 20:1 without noticeable effect. The difference arises because the size and viewing distance control what information the observer can see in the final prints. Large digital files often contain

more information than can be seen in a small print. This is the information that jpeg discards. As with unsharp masking the user specifies the spatial parameters to optimize performance and avoid artefacts.

Retinex has parameters that are responsive to both spatial frequency and dynamic range of the input data. The number of iterations, as specified in the MATLAB™ code by 'nIterations', controls the amount of dynamic range compression and sets the stage for a different level of post-processing by a postlut. The term "postlut" derives from historical use of image processing hardware using a lookup table. Postlut processing simply refers to the application of a function f applied uniformly to every image pixel, I(x,y)=f(I(x,y), for all image locations (x,y). The effect of the number of iterations can be seen in Figure 4.

As we can see the effect of the number of iterations (nIterations) is to reduce the contrast of the images as demonstrated by the smaller range in the histograms. The process moves the entire image into a smaller dynamic range, with smaller digit differences representing edge ratios. With very few iterations, the range of output digits is small. The postlut expansion (stretching of the image intensities) must be large to regenerate edge ratios appropriate for a print. With more iterations the range of output digits is larger. The postlut expansion will be moderate to regenerate edge ratios. With a very large number of iterations the range of output digits is large, approaching that of the input image. The postlut expansion must be small to none to regenerate edge ratios. The amount of postlut expansion and its shape will vary with the amount of dynamic range compression.

The examples of unsharp masking and jpeg compression demonstrated the need for selecting the right parameters to match viewing size and viewing distance. Analogously, the viewing distance, the viewing size, the dynamic range and noise level of the input image, the number of iterations, and the postlut are all important to make artifact-free Retinex images.

Scaling of Retinex output to desired media and purpose

As shown in Figure 4, the contrast of the output is controlled by the number of iterations. This parameter can vary the output from radical to no dynamic range compression. The input data also plays a major role. The total dynamic range of input data determines the

magnitude of radiance ratio associated with each digit. The final parameter is the postlut that matches the final new product with the output media. That media can be a printer, a monitor, a LCD display, a system profile, a 3-D plot of output at each pixel (output equal height), a pseudocolour image. The essential idea is that the input calibration controls the correlation between digital differences and radiances in the world. The number of iterations controls the degree of compression. The postlut controls the rendition of New Product digital differences in the output media. All three parameters (input dynamic range, number of iteration and postlut) are crucial to the process. All three share the control of the output image. They can be used only as well designed sets. They are not randomly interchangeable.

Results on Test Images

Figure 5 through Figure 8 illustrate the behaviour of the two algorithms. Figure 5 shows the behaviour when the input is a simple square at the very centre of the image. A slight asymmetry can be seen in both the McCann99 (using 4 iterations comparing 8 nearest neighbours) and Frankle-McCann (using 4 iterations of 4 directions) outputs.

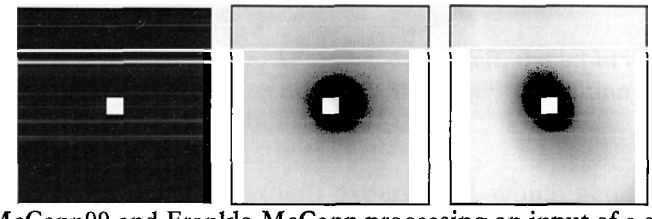


Figure 5. Effect of McCann99 and Frankle-McCann processing on input of a single bright square against a black background. No post-LUT has been applied to enhance the visibility effect

In the limiting case of the square being a single pixel, this is analogous to the point spread function for the algorithm. It must be noted that because of the reset step, the shape of this function varies depending on direction of individual comparisons the image content. Frankle-McCann used different patterns of spatial comparisons to minimize these effects. From left to right we have: input image, McCann99 4-iteration output, Frankle-McCann 4-iteration output.

The calculations in Figure 5 used the same pattern of spatial comparisons for each size of comparisons. The original Frankle and McCann calculation changed the order of the direction of comparisons in each size of spatial separation. This sequence of the spatial process was controlled by a lookup table of comparison directions. Such randomization of the comparison process minimizes the directional gradients shown in Figure 5. Alternatively, one can change the averaging process controlling the Old Product. If all the reports from different direction were averaged before changing the value of the Old Product, then these calculated spatial asymmetries are not observed. The use of postluts and more complex sequences of spatial comparisons all contribute to reducing the magnitude visibility of asymmetries.

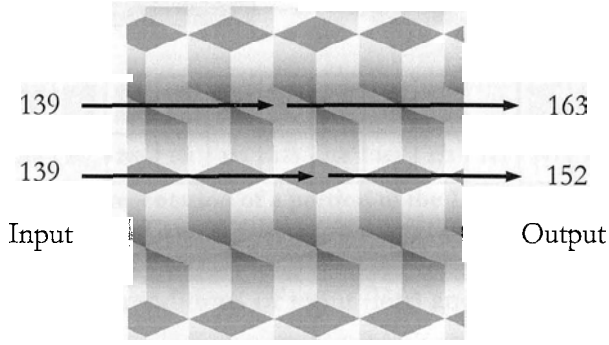


Figure 6. Logvinenko cubes pattern illusion. As shown on the left, the input values of the cube tops are equal despite the fact that we see them as unequal. McCann99 4-iteration Retinex output values are shown on the right.

Figure 6 shows Logvinenko's gradient experiment, which generates a large lightness change between the diamonds. A vertical sinusoidal gradient in non-diamond areas creates the illusion. The numbers on the left side of Figure 6 show that the input digits for the light and the dark diamond faces are both 139. The numbers on the right show the output from the corresponding faces to be 152 and 163 after McCann99 4-iteration processing. McCann[16] reports that "Retinex models can predict appearances that were previously attributed to cognitive behaviour." Despite the fact the upper cube faces on alternating rows appear to differ in intensity, the top faces of all the cubes are in fact both uniform and equal. In the output, however, the top faces of the cubes are no longer equal nor are they completely uniform.

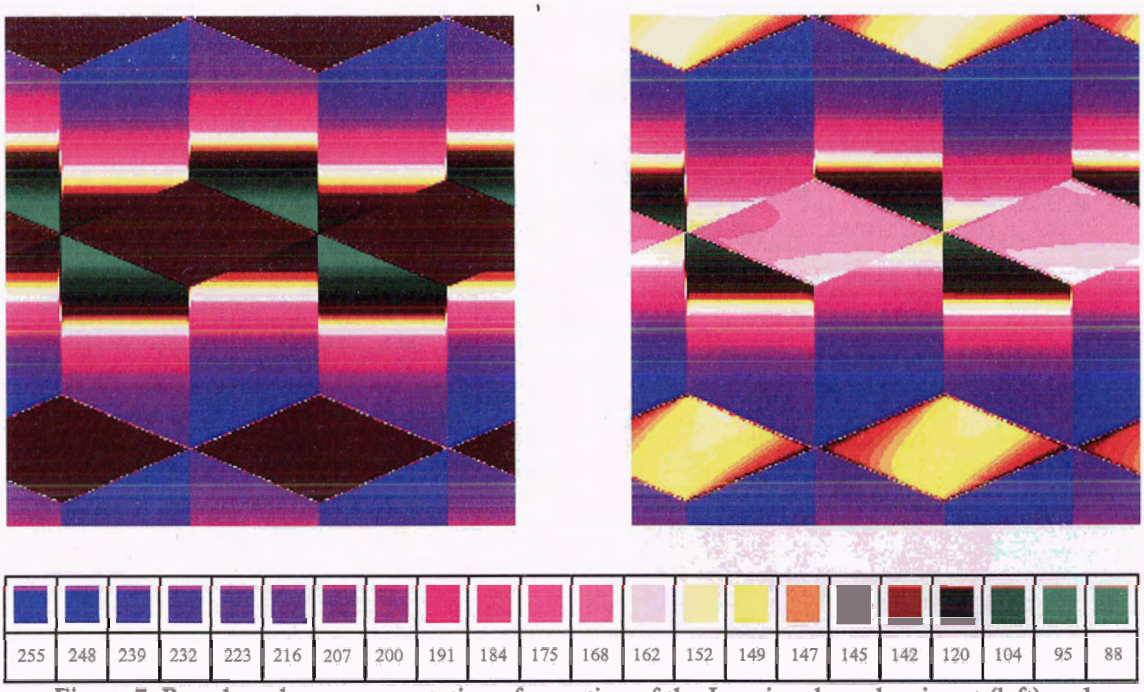


Figure 7. Pseudo-colour representation of a portion of the Logvinenko cubes input (left) and McCann99 4-iteration output (right).

Figure 7 shows pseudo-colour renditions of input (left) and output (right) of the Logvinenko illusion. The diamond shaped tops of the cubes are equal on the left and unequal on the right. Note that the upper faces of the output cubes are not uniform.

Figure 8 shows the effect of McCann99 applied to a colour image with a substantial blue colour cast. The algorithm has been applied to each of the colour channels independently. Clearly, in this case the colour cast has been removed. Retinex differs from many colour constancy methods in that it does not aim to find a single chromaticity for the scene illumination as is the case, for example, in the neural network[54] and colour by correlation[17] methods. Retinex instead adjusts the image colours in a non-global manner as is necessary since the model attempts to match human visual response. Some effects of this can be seen in the way that in Figure 8B some of the green bleeds into the white area surrounding the "C" in "Compiler," and the way the blue is darkened near the white lettering on the right-hand blue book.

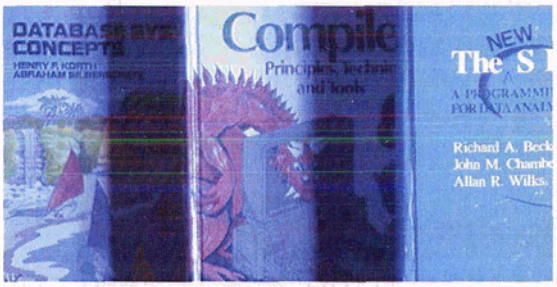


Figure 8A

Figure 8B

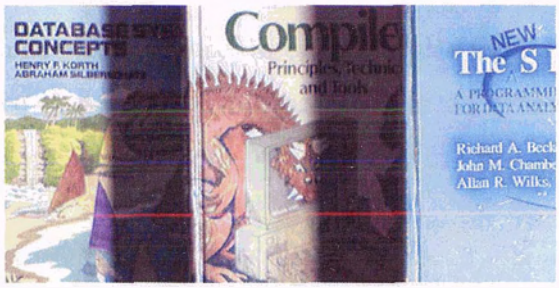


Figure 8C

Figure 8. (A) Input with blue colour cast created by scene illumination for which the camera was not balanced; (B) Output from the McCann99 4-iteration; (C) Output from the Frankle-McCann 4-iteration.

The results corresponding to Figure 8 can be compared with those of Barnard and Funt.[82] Note that both the input and output images have been adjusted with postluts for printing. The actual Retinex input image is in log space. The image also has extended dynamic range obtained by frame averaging.

Discussion

This work describes the basic Retinex algorithms in MATLAB™ code. It provides the starting point for many different implementations for many possible variations. This code is the basis of making spatial comparisons in a very efficient manner. In carefully calibrated situations it can be used as the basis for a model of human colour appearance. This requires accurate calibration in both the luminance and spatial frequency domains. Numbers of iterations for each pixel separation or level of pyramid processing must match human spatial frequency data.[1,16,90] Alternatively, it can be used to enhance images of unknown calibrations in digitization. In an uncalibrated mode it is more limited. The system works by enhancing edges. If poor calibration introduces edges from noise, the process will enhance the noise.

Nevertheless, uncalibrated input images generally appear better with Retinex processing than without it.

As in many image-processing operations,[91] there are three sequential steps :

1. Taking the raw input and transforming the information into an image space appropriate for the process.

2. Performing the process

3. Scaling the output result into a space appropriate for end use. In this particular case, the input converts the captured digits into a space in which constant scene edge ratios have constant differences in digits. With an appropriate output scaling, this property can be used by the process to model visual appearance.

The process assumes that the visual system uses edges to synthesize appearance. The Retinex algorithms provide an image processing engine that synthesizes sensation images from spatial comparisons of radiance inputs. The meaningful parameters in McCann99 are the pyramid level and the number of iterations. In Frankle and McCann it is separation and number of iterations. In McCann, McKee and Taylor it is path length and number of paths. A number of studies experimentally measured the appearance of a variety of achromatic and colour constancy experiments. Using this quantitative data it is possible to experimentally optimize the parameters of the model.[1,92] The details of this aspect are found in Chapter 6 and in McCann and Savoy.[93] All of these studies indicate that the human visual system is neither local, nor global with regard to spatial interactions. Neither local centre-surround operators, nor global gray world models can account for psychophysical results. The spatial frequency filter applied by human vision is image dependent.[84] The effect of maxima have an effect over large distances, but varies with distance and enclosure.[94]

In the examples described above, we used constant values for number of iterations for all levels of pyramid. Although efficient, this is not the best set of parameters for modeling human vision. An obvious variation is to have different numbers of iteration for each size. Franke and McCann used different number of iterations for each size of separation. They also changed the pattern of directions to remove the pattern found in

Figure 5.[39] The 1, 4, and 128 iteration images in Figure 4 could be described by their spatial-frequency content. The difference between the input and output images describes a spatial filter. That filter can be resolved into a two-dimensional spatial filter, or set of spatial filters. Since the work of Campbell and Robson, and Hubel and Wiesel, human visual processing has been regarded as sets of spatial channels.[95] As demonstrated in Figure 4, the number of iterations controls the strength of the filter. The greater the number of iterations the weaker the filter. The size of the separation or the pyramid level controls the spatial frequency of the response. The number of iterations at that level controls the strength of the filter at that frequency. Just as human vision requires models using multi-channel with different filter strengths, the Retinex models should have the same spatial frequency tuning.

Sobol[94] has described variations to the Retinex process that uses lookup tables to control the magnitude and shape of edges at different spatial separations. This algorithm produces dramatic images. The ability to control different spatial frequencies adds considerable power to the algorithm. In addition it makes the model more like human vision.

An important final variation is the use of the spatial comparison engine for gamut mapping problems. Examples are found in[84]. The principle is straightforward. If displays and printers had the same colour spaces then tristimulus matches would be able to successfully transform display/print images. However, they occupy only half of their combined physical colour space. Using strict colorimetric matches creates problems with extra-gamut colours. All of the variation between the gamut of the smaller space is represented by the gamut value. This clipping of local detail produces undesirable artefacts. Many algorithms systematically distort the colorimetric matches to achieve an image with a better appearance. All the transforms increase the colorimetric errors.[84,96]

The Retinex approach uses two different sets of RGB input images. One image (Goal) has digits representing the colour space values of the large gamut desired image. The other image (Best) has digits representing the colour space values of the best colorimetric reproduction possible in the smaller gamut media. The RGB Goal images are used to supply the ratios. The Best image is used to supply the reset values. The rest of the process is the same as described above. The colour gamut calculation provides an excellent example of using the Retinex spatial-comparison process to generate new sensation images that have

very similar appearance with different radiances at each pixel. Experiments have shown that human spatial processing is key to understanding colour constancy, high dynamic range sensations and transparency.[97] Further, spatial comparisons can be used to simplify gamut mapping algorithms. As long as spatial comparisons are constant, near constant appearances can be made from very different stimuli.

Conclusions

This chapter presents new, very concise MATLAB™ implementations** of two of the main practical Retinex algorithms. Our hope is that this will eliminate much of the variability in what is meant when different researchers refer to Retinex and thereby facilitate further rigorous testing and discussion of the method. For modelling human vision these MATLAB™ programs depend on calibrated input data. Although these MATLAB™ programs provide the details of how pixels are compared and processed during the ratio-product-reset-average steps of Retinex processing, they do not provide details on the selection of an appropriate postlut for a particular output device. The problem of adjusting the postlut is addressed in Chapters 5 and 6.

** The MATLAB™ code and figures are also available electronically at http://www.cs.sfu.ca/research/groups/Vision/

CHAPTER FIVE*
CONTROL PARAMETERS FOR RETINEX

In the previous chapter we provided Matlab implementations of the Retinex algorithm, which has free parameters for the user to specify. The parameters include the number of iterations to perform at each spatial scale, the viewing angle, image resolution, and the lookup table function (post-lut) to be applied upon completion of the main Retinex computation. These parameters were specifically left unspecified since the previous descriptions of Retinex upon which the new Matlab implementations were based do not define them. In this chapter we determine values for these parameters based on a best fit to the experimental data provided by McCann et. al.[1]

Introduction

The Retinex model for the computation of lightness was introduced by Land and McCann[37] in 1971. Land and his colleagues later described additional improvements to the original method.[1,16,39,40,74] These further refinements were mainly designed to improve computational efficiency while preserving the Retinex principle of comparing pixel values from different spatial locations. Matlab code for two of the main Retinex algorithms is provided in the Appendices.

Even though the Retinex algorithm is well documented, there are still many things which need to be specified before it can be used as a model of human colour or lightness perception. In particular, there are parameters which are sensitive to both the spatial frequency and the dynamic range of the input image. We estimate values for these parameters based on fitting the experimental data obtained by McCann, McKee and Taylor.[1] We will refer to their study as the MMT (McCann-McKee-Taylor) experiment. A central part of the MMT experiment involves haploscopic matching of Munsell papers arranged in a

---

* This chapter also appears as a published paper: Brian Funt and Florian Ciurea, "Control Parameters for Retinex", in Proc 9th Congress of the International Color Association, 287-290, 2001, Rochester, NY.

Mondrian display.[1] Our procedure is to reconstruct digital images corresponding to the cone responses of the standard observer, and then run the Retinex algorithm on that data while varying the algorithm's parameters to find those for which the program's output best matches the MMT corresponding colour data.

Preparing Retinex Input Data

Our goal is to determine the values of the free Retinex parameters which make Retinex work as accurately as possible as a model of colour appearance in complex visual scenes. In the MMT experiment,[1] subjects alternately viewed a Mondrian with one eye and a Munsell chip with the other eye. For each coloured area in the Mondrian, the subjects chose a matching Munsell chip. The experiment was repeated under 5 different combinations of three narrowband illuminants. The results in[1] are reported in terms of the designators of the matching Munsell chips. In this paper we rely on the corresponding CIE tristimulus values estimated by Nayatani et. al.[98]

Our first step was to construct an LMS image of the Mondrian used in the MMT experiment as it would be under each of the 5 illuminants. The layout of the colour patches in the Mondrian is given in MMT[1]. We convert the corresponding XYZ of each patch as estimated by Nayatani[98] to cone quanta catch values using the following transformation:[99]

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.38971 & 0.68898 & 0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{15}$$

Similarly, the XYZs of the matching Munsell chips are converted to LMS. The natural logarithm of each L, M, S value is then taken since the Matlab Retinex implementations require the logarithm of the image as input. Retinex is run on each of the L, M and S channels independently.

Post Retinex processing

The post-Retinex processing consists of four stages: exponentiation, scaling to white, conversion to Munsell Value scale, and compensation for differences in overall illumination

intensity. Exponentiation of the Retinex output simply compensates for the logarithm which was applied to the input data. Scaling to white is required because the Retinex algorithm normalizes each of the LMS channels to 1. After Retinex processing an ideal white patch will result in (1,1,1); however, the LMS value of the Munsell white (MMT area K) under the 'white' illumination in the MMT experiment is (92.55, 72.84, 49.23). Hence, we scaled the Retinex output values to make the Retinex white equal the MMT white. The three scaling factors, one for each channel, were then held constant across the 5 MMT illumination conditions.

The second post-Retinex stage is to convert to Munsell Value scale, which is required because McCann et. al. compare the colours in the Mondrian to the matching Munsell chips using it. They convert integrated reflectance ρ (e.g., L/Lwhite) to Munsell Value using the approximation:[100]

$$V = 2.539\,\rho^{1/3} - 1.838, \quad \text{for } \rho > 0.384\% \tag{16}$$

The third stage is to compensate for differences in overall illumination intensity between the test and match conditions based on the data in Figure 8 of MMT. McCann et. al. found that overall intensity affected subjects' matches. Hence, we incorporated their correction factor as a function of the ratio of overall illumination between the two scenes. By analyzing MMT Figure 8, we computed the correction to be added to the Retinex output converted to Munsell Value, based on the scene radiances E at 630-nm, 530-nm and 450-nm as:

$$\begin{aligned}
\text{Correction}^{630} &= 1.53 \times \log_{10}(E^{630}_{\text{Mondrian}}/E^{630}_{\text{Munsell}}) + 0.04 \\
\text{Correction}^{530} &= 1.19 \times \log_{10}(E^{530}_{\text{Mondrian}}/E^{530}_{\text{Munsell}}) + 0.11 \\
\text{Correction}^{450} &= 0.93 \times \log_{10}(E^{450}_{\text{Mondrian}}/E^{450}_{\text{Munsell}}) + 0.01
\end{aligned} \tag{17}$$

Results on choosing the parameters

To establish the optimum choice for the number of Retinex iterations, we ran Retinex with the number of iterations (parameter nIterations in the Matlab implementation) varying from 1 to 500. The post-Retinex processing described above was then applied in each case. For each iteration setting, we computed the difference between the final Retinex prediction and

the matching Munsell chip data found in the MMT experiment. The image difference measure is the RMS over all pixels of the following single pixel difference measure:

$$d_3(R^c(i,j) - M^c(i,j)) = \sqrt{\sum_{c=L,M,S} (R^c(i,j) - M^c(i,j))^2} \tag{18}$$

$R^c(i, j)$ denotes the pixel value at channel c for Retinex output including post-processing; $M^c(i, j)$ denotes the pixel value at channel c for an image of the Mondrian made up of Munsell matching chips.

We found that for each of the five different MMT experimental set-ups—"gray", "red", "blue", "green" and "yellow"— different numbers of iterations were required to give the best match to the matching Munsell data. Although the number of iterations varied across the cases, 33 iterations gave the best overall result.

Image resolution is another variable which must be considered. To determine how the optimum number of iterations might be affected by image resolution, we constructed images of otherwise identical Mondrians at resolutions of 128×128, 256×256, 512×512 and 1024×1024. Contrary to what might be expected, we found almost no change in the optimum number of Retinex iterations required as a function of image resolution.

The following graphs illustrate how the number of iterations affect the distance between the Retinex prediction and the actual image as seen by the observers, for each experiment. For these tests we ran Retinex with input images of 256×256 pixels.

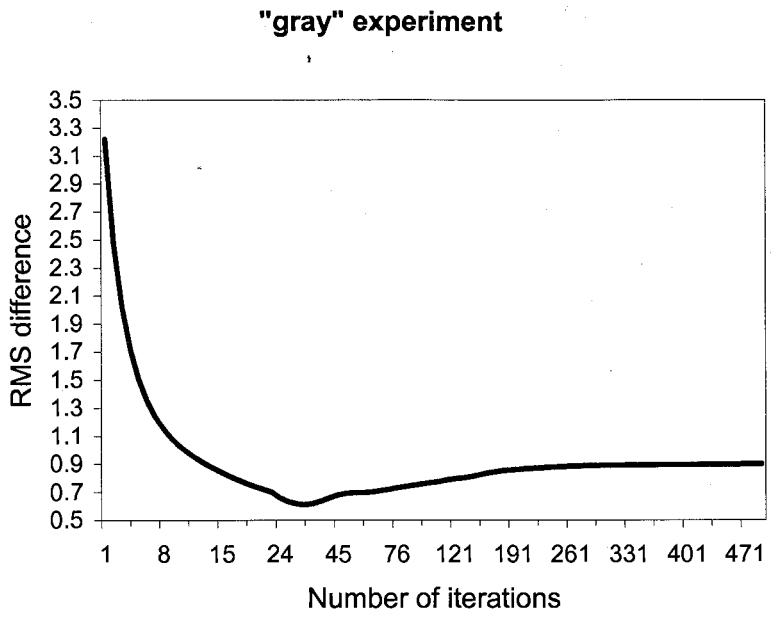


Figure 9. MMT "gray" experiment. The accuracy of Retinex as a model measured by the RMS difference between the Retinex output and the corresponding colour data as a function of the number of Retinex parameter nIterations.

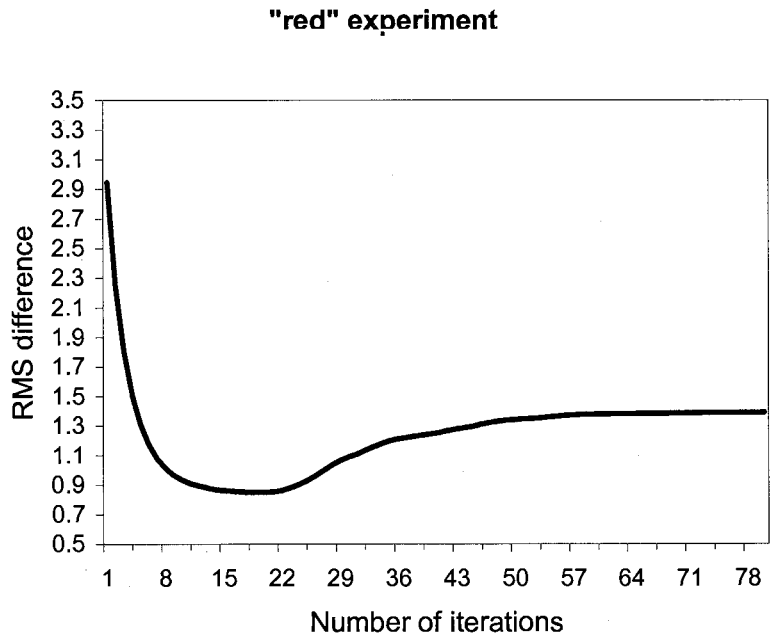


Figure 10. MMT "red" experiment. The accuracy of Retinex as a model measured by the RMS difference between the Retinex output and the corresponding colour data as a function of the number of Retinex parameter nIterations.

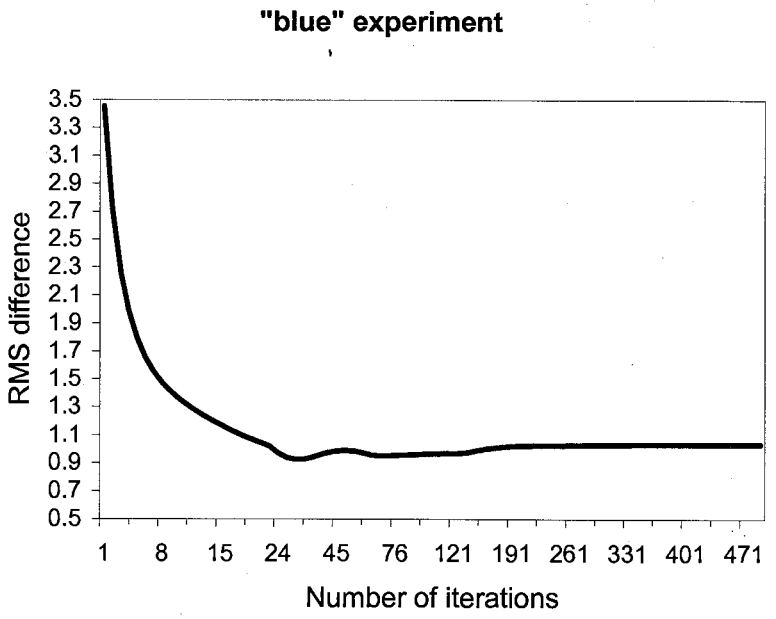


Figure 11. MMT "blue" experiment. The accuracy of Retinex as a model measured by the RMS difference between the Retinex output and the corresponding colour data as a function of the number of Retinex parameter nIterations.

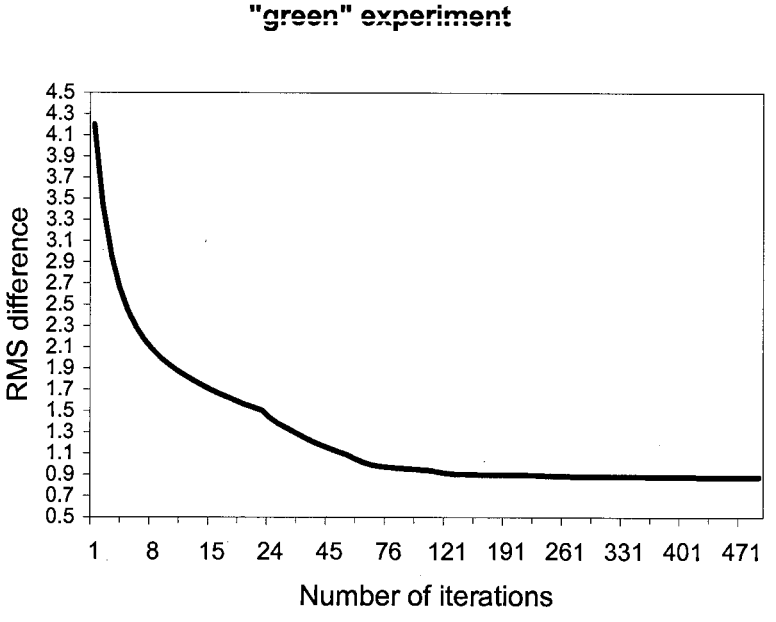


Figure 12. MMT "green" experiment. The accuracy of Retinex as a model measured by the RMS difference between the Retinex output and the corresponding colour data as a function of the number of Retinex parameter nIterations.

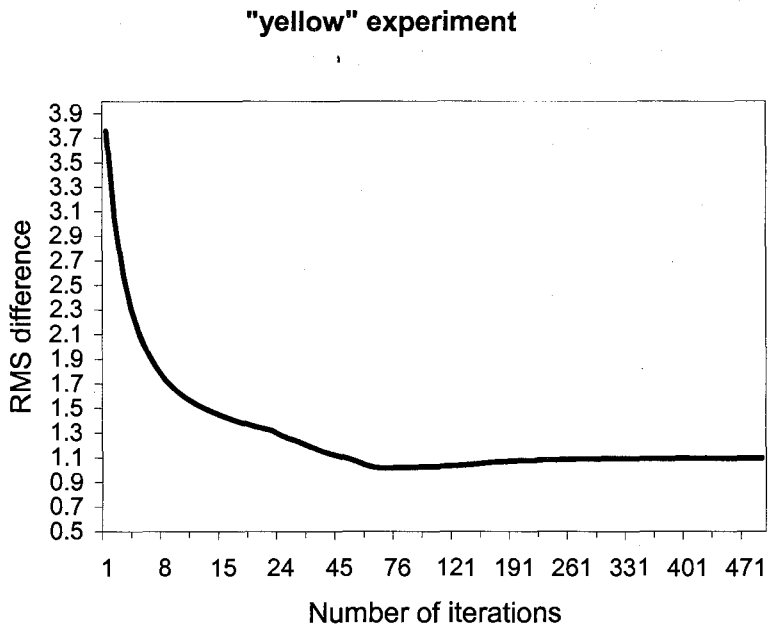


Figure 13. MMT "yellow" experiment. The accuracy of Retinex as a model measured by the RMS difference between the Retinex output and the corresponding colour data as a function of the number of Retinex parameter nIterations.

Conclusions

For the MMT experiments we have been able to solve for the parameters needed in the Retinex algorithm, based on a best fit with the experimental data. We have found that if we run either of the Matlab Retinex algorithms, the optimum results can be obtained by choosing the number of iterations to be 33. An interesting finding was that the image resolution has very little effect on the accuracy of prediction. Details of the post-Retinex processing step were also established. Although we have established parameters for the Retinex computation based on the available MMT data, it would be helpful to have more extensive experimental data to improve the reliability of the results.

CHAPTER SIX*

TUNING RETINEX PARAMETERS

The goal of this chapter is to understand how the Retinex parameters affect the predictions of the model. A simplified Retinex computation is specified in the MATLAB™ implementation provided in Chapter 4; however, there remain several free parameters which introduce significant variability into the model's predictions. Here we extend previous work started in Chapter 5 on specifying these parameters. In particular, instead of looking for fixed values for the parameters, we establish methods that automatically determine values for them based on the input image. These methods are tested on the McCann-McKee-Taylor asymmetric matching data[1] along with some previously unpublished data that include simultaneous contrast targets.

Introduction

The Retinex MATLAB™ implementations presented in Chapter 5 have three important input parameters: the number of iterations the algorithm performs at each level of its multi-level computation, the "post-lut" output lookup table function, and the input image size. The model's final output depends strongly on the values chosen for the parameters.

The Retinex model aims to predict the sensory response of lightness. In Chapter 5 we suggested values for the parameters based upon fitting the model's predictions to the data originally described over 35 years ago by McCann, McKee and Taylor.[1] This fit led to the conclusion that 33 iterations had the lowest global average of the differences between observer data and computed values, assuming that the number of iterations was constant for all levels of the multi-resolution computation. However, McCann+ felt that 33 was too high a number, and would not lead to a good model of simultaneous contrast. Hence, together we began the current series of experiments by including previously unpublished data from

---

* This chapter also appears as a published paper: Florian Ciurea and Brian Funt, "Tuning Retinex Parameters", Journal of the Electronic Imaging, Vol. 13, No. 1, 58-64, 2004.

+ John McCann, personal communication

lightness matching experiments with simultaneous contrast targets. We also added other unpublished data for targets containing a fixed set of patches of various shades of gray appear on a background that varied from black to gray to white.

For the simultaneous contrast data, we indeed did find that a much smaller value is required for the iteration parameter in order to make a good fit. However, we could no longer find a universal value for the number of iterations that simultaneously would minimize the error for the combined data from the MMT (McCann-McKee-Taylor), SB (fixed scale of grays on different backgrounds), SC (simultaneous contrast) and GW (gray on white). This led us to consider a method of automatically calculating how many iterations to use based on how the computation was proceeding. As described in the Chapter 5, the post-lut processing needs to change as a function of the number of iterations, so this led to a method of automatically calculating the appropriate post-lut.

Number of Iterations

The two MATLAB™ implementations (see Appendices) are referred to as McCann99 Retinex and Frankle-McCann Retinex. For brevity, we concentrate here only on McCann99 Retinex, but the results are similar for both versions. McCann99 Retinex creates a multi-resolution pyramid from the input by averaging image data. It begins the pixel comparisons at the most highly averaged, or top level of the pyramid. After computing so called New Products (precursors to the final lightness estimates) on the image at a reduced resolution, the resulting New Product values are propagated down, by pixel replication, to the pyramid's next level as initial estimates at that level. Further pixel comparisons refine the estimates at the higher resolution level and then those new estimates are again propagated down a level in the pyramid. This process continues until values have been computed for the pyramid's bottom level.

At each level, the basic step is the comparison of each pixel to each of its immediate neighbors. The number of iterations refers to the number of times all the immediate neighbors are cycled through before moving down to the next level in the pyramid. Since pixels are only directly compared to immediate neighbors, comparisons to more distant pixels at the current pyramid level are only made implicitly by propagation of information

from pixel to pixel during these iterations. Hence, increasing the number of iterations increases the spatial distance across which pixels are related during the computation. McCann99 Retinex uses the same number of iterations at all levels and so there is only a single iteration parameter to specify and have limited this paper to considering a single value for all levels.

Post-lut processing

Postlut processing refers to applying a function f uniformly to every image pixel, $I(x,y)=f(I(x,y))$, for all image locations (x,y) immediately after the main Retinex computation. The term "postlut" derives from historical use of image processing hardware using a lookup table (lut) as a final post-processing step. Post-lut processing is important in bringing the final result into the appropriate dynamic range, compensating for differences in overall illumination intensity between test targets, and in converting to the coordinates of Munsell Value scale used in recording the experimental data. Although all these factors can be thought of separately, they are all eventually combined into a single post-lut function.

The first post-lut step adjusts the dynamic range. Retinex output from the pyramidal spatial comparison stage, falls in the [0,1] range. Because the value 1 represents 'white' and Retinex assumes there is at least one white pixel in every image, the value 1 necessarily arises in the output. However, the lowest output value depends on the image content and varies with the number of iterations used. The fewer the iterations, the more local the spatial comparisons will be, and therefore, the less the likelihood of big intensity differences being found. As a result, the fewer the iterations, the higher the minimum Retinex output value (Figure 4 illustrates this effect). The first purpose of the post-lut is to stretch the Retinex output to a reasonable range. Since the amount of stretching needed depends on the number of iterations, and we vary the number of iterations in our experiments, we decided to always linearly scale the Retinex output to the full [0, 1] range. This stretch does not correct for the fact that the number of iterations performs a non-linear compression of the image. The post-lut is not fixed, but rather depends on the input image and number of iterations used. This decision effectively means that we are assuming that there is at least on black location in the test target. While this assumption need not be true for images in general and could lead to errors in Retinex predictions, it is true for all the test targets subjects viewed.

After scaling to the [0,1] range, the post-lut then converts the Retinex output values, ρ, to the lightness scale used for recording subject's matches. For the MMT data set, the conversion is to Munsell Value scale V^{100} described in Equation 16.

For the SB, SC and GW data sets, the conversion is to a lightness scale described by Stiehl et. al.[101] Based on a fit to the raw data, we use the following function to convert the log luminance to the lightness scale values, *L*:

$$L = 129.6 \cdot \rho^{1/100} - 132.45 \tag{19}$$

The final post-lut component compensates for differences in overall illumination intensity between the test and match conditions. Only the MMT experiments involved such intensity differences. The compensation is based on data from Figure 8 of McCann, Land and Tatnall.[92] Generally, the effect of this correction is slight. Details are provided in Chapter 5.

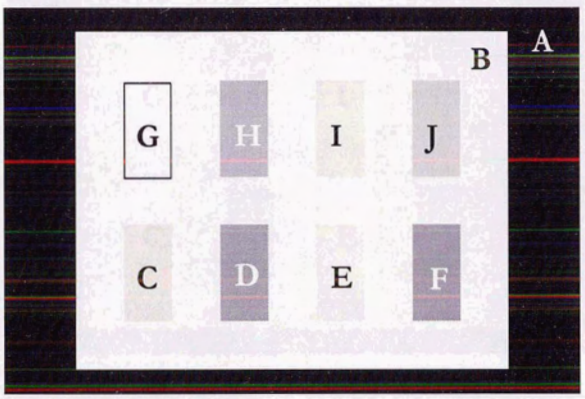
Lightness Matching Data

The experimental technique for the MMT matching experiments was reported a long time ago.[92] The 'new' data we report here is based on experiments by McCann, which were also conducted earlier, but not previously reported in the literature. These experiments involve transparent greyscale targets lit from behind with uniform illumination. Subjects were asked to report the lightness of each patch in the target display using a standard lightness transparency display as a reference. The standard lightness display consists of 25 squares of different lightness values against a white surround. The squares are arranged in a serpentine path such that the change in lightness from any of the 25 squares to the next is constant[101]. In the resulting lightness scale, 1.0 corresponds to an opaque area and 9.0 to the brightest area. The experiments were based on 4 to 7 subjects, which each subject repeating the matches on 3 different occasions.

The matching procedure was set up such that in the normal viewing position, the subject saw the test display as the only thing in the field of view. By turning 90 degrees to the right, the subject would see instead the standard lightness display as the only thing in the field of view. Subjects were allowed to look back and forth between viewing the test display

and the standard display as many times as desired without a time constraint.[92] The test display and the standard lightness display had the same level of luminance.

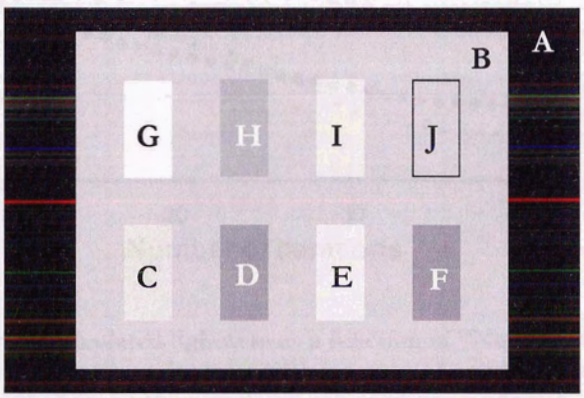
Figure 14, Figure 15, Figure 17, Figure 19, Figure 20 and Figure 22 illustrate the targets along with the corresponding luminance, pixel value for each patch as input to the Retinex algorithm, and average observed lightness reported for each patch. All the patches have uniform luminance. It should be noted that the figures are intended only to illustrate the corresponding targets. They are not accurate reproductions of the targets. Their printed appearance is not the same as under the controlled experimental conditions.



Patch	Lumi-nance	Pixel value	Match	Std dev	Calc.	Calc-Match
G	1001	255	8.75	0.15	9.01	-0.26
E	595	236	7.55	0.20	7.66	-0.11
I	439	225	6.25	0.25	6.97	-0.72
C	336	215	5.94	0.31	6.40	-0.46
J	228	200	5.19	0.19	5.63	-0.44
H	125	178	4.36	0.26	4.66	-0.30
D	63	153	3.37	0.50	3.76	-0.39
F	50	145	2.80	0.30	3.51	-0.71
B	1001	255	8.80	0.20		
A	1	0	1.0	0		

Figure 14. "Scale on White" target along with patch identification, the luminance values measured in the original display, the digit representing log luminance, the mean and standard deviation of observer matches in Munsell Value units. The sixth column lists the calculated Lightness for all iterations above 3. The seventh column lists the errors between Observed and Calculated Lightness.

The calculated lightness for the "Scale on White" display are nearly constant with changes in "Number of Iterations". In a solid white surround all gray patches have a constant value after the third iteration. As shown in the table in Figure 14 the calculated lightness values (sixth column) are close to the observer matches (fourth column). There are residual errors (seventh column) with an average value of 0.42 ± 0.2. Since the white surround is the control case that establishes the shape of the lookup table, the lack of perfect correlation is due to experimental and lookup table error. These errors have no effect on the analysis of number of cycles, but contribute to any global average.



Patch	Lumi-nance	Pixel value	Match	Std dev	Calc.	Itera-tions
G	1001	255	9.00	0.20	8.95	1
E	595	236	8.50	0.50	8.53	24
I	439	225	7.31	0.31	7.32	31
C	336	215	7.06	0.31	7.06	26
J	228	200	5.88	0.38	5.88	28
H	125	178	4.98	0.28	4.98	26
D	63	153	4.08	0.48	4.08	23
F	50	145	3.05	0.55	3.51	50
B	228	200	5.75	0.25		
A	1	0	1.0	0		

Figure 15. "Scale on Gray" target along with patch identification, the luminance values measured in the original display, the digit representing log luminance, the mean and standard deviation of observer matches in Munsell Value units and the calculated values for the best fit to observer match. The iterations column list the number of iterations for best fit for Calculated to Observed Lightness. The average number of iterations for best fit from Areas E, I, C, J, H, and D is 26.33 ± 2.88, while the average that included areas G and F is 26.13 ± 13.32. The best fit for "Scale on Gray" is 26 iterations.

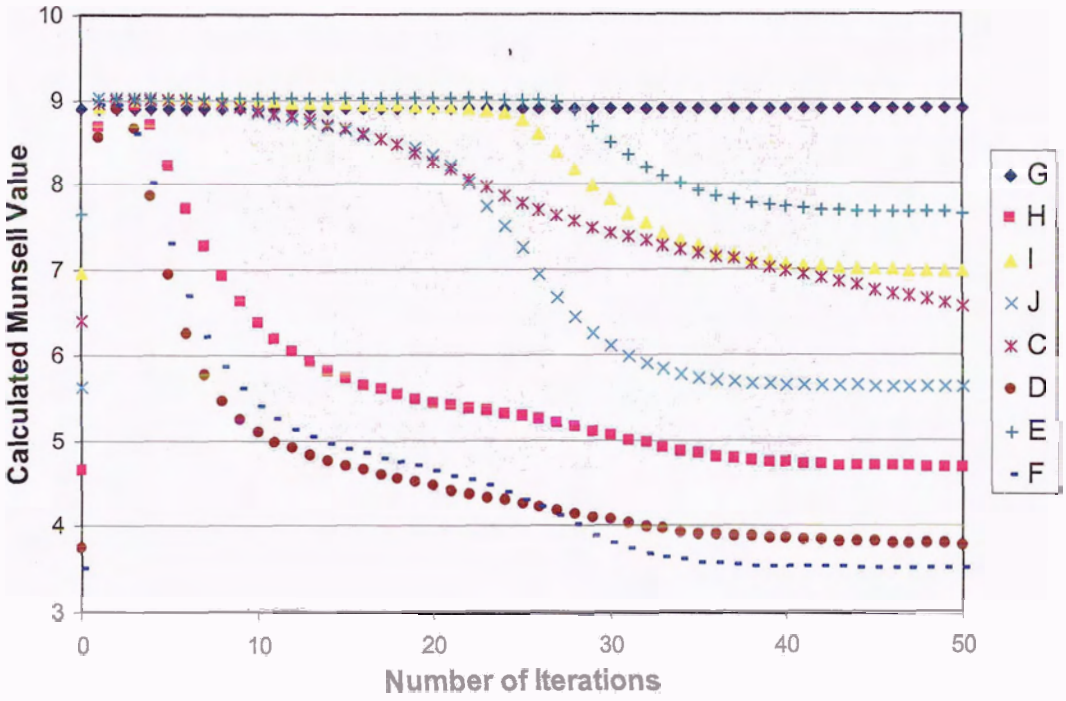
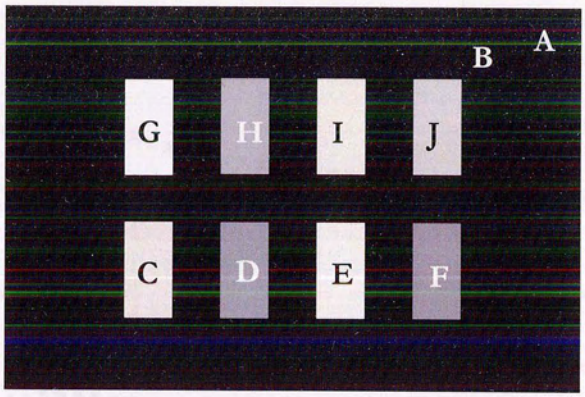


Figure 16. "Scale on Gray" calculated lightness as a function of "Number of Iterations". In a gray surround all gray patches except white decrease with increase of number of iterations. The number of iterations has significant effect on the calculated values of grays. Area E, the lightest gray has a calculated lightness equal to white up until 20 iterations. Areas E, I, C, H, D, and F show different degrees of non-monitonic decrease in calculated Munsell Lightness. The darkest gray, Area F, and mid-gray, Area J, both show second phase starting at 20 iterations. A slightly lighter gray, Area C, shows a similar change in slope at 35 iterations.



Patch	Lumi-nance	Pixel value	Match	Std dev	Calc.	Itera-tions
G	1001	255	9.00	0.05	8.88	1
E	595	236	8.55	0.45	8.60	3
I	439	225	7.53	0.78	7.54	26
C	336	215	7.29	0.54	7.28	33
J	228	200	6.80	0.50	6.67	27
H	125	178	5.65	0.35	5.65	16
D	63	153	5.20	0.50	5.25	9
F	50	145	4.68	0.38	4.69	19
B	1	0	0.90	0.10		
A	1	0	0.83	0.08		

Figure 17. "Scale on Black" target along with patch identification, the luminance values measured in the original display, the digit representing log luminance, the mean and standard deviation of observer matches in Munsell Value units, number of iterations and calculated lightness. In a black surround, the calculated lightness for all gray patches, except white, decreases with increase of number of iterations. The average for best fit from Areas G, E, I, C, J, H, and D is 16.8 ± 11.7

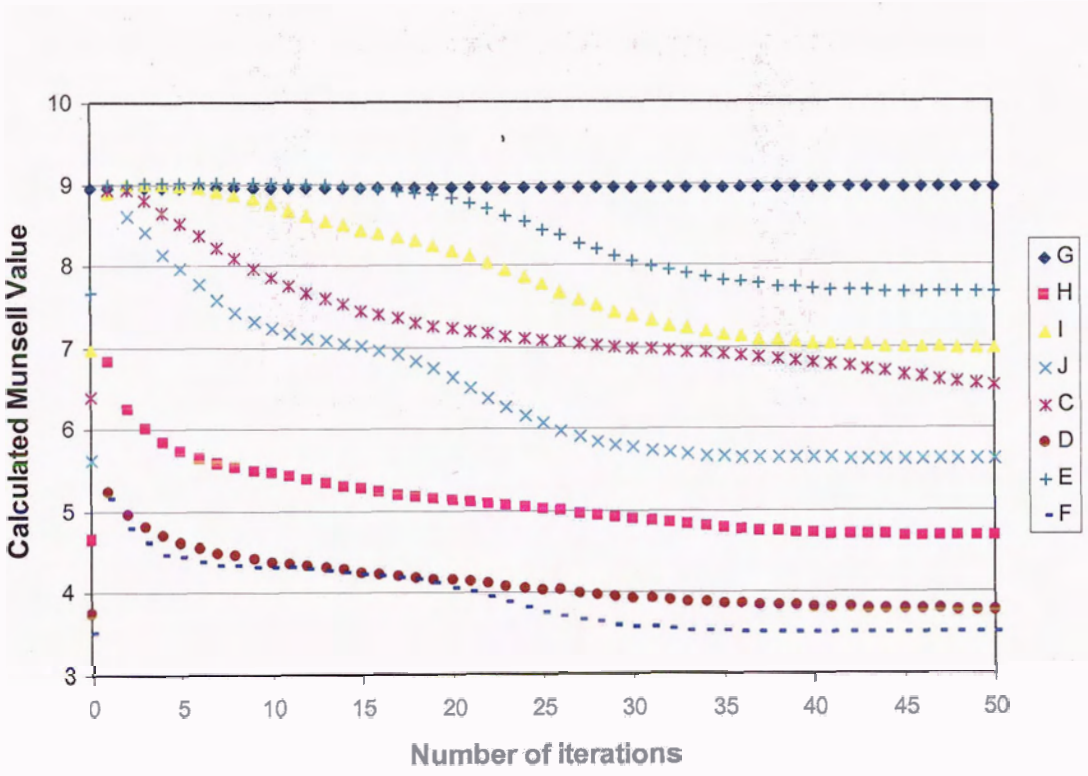
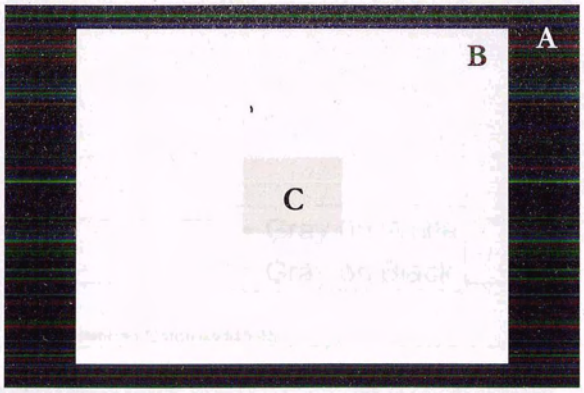
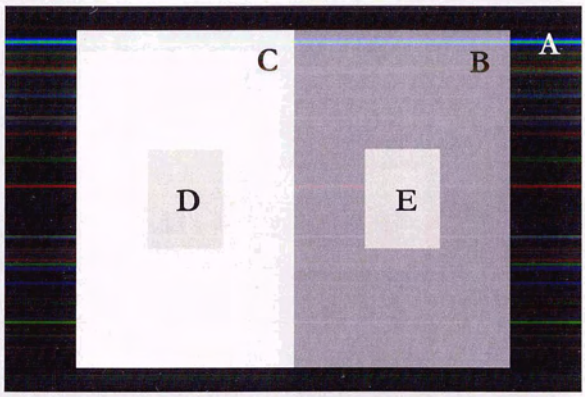


Figure 18. "Scale on Black" calculated lightness as a function of "Number of Iterations". In a black surround all gray patches except white decrease with increase of number of iterations. Area E, the lightest gray has a calculated lightness equal to white, up until 30 iterations. Areas I calculated lightness begins to fall at 25 iterations. C and J calculated lightness begin to fall at 12 iterations. The darkest grays begin to fall at 5 iterations.



Patch	Lumi-nance	Pixel value	Match	Std dev
B	1001	255		
C	321	213	6.15	0.52
A	1	1	1.13	0.13

Figure 19. "Gray on White" target along with patch identification, the luminance values measured in the original display, the digit representing log luminance, the mean and standard deviation of observer matches in Munsell Value units. There is no significant change in calculated values for white and gray. Black values vary for iterations of 1 to 7. The best fit is 3 iterations with a calculated value of 1.16, while observed value is 1.13. The calculated asymptotes are 1.00, 6.29 and 9.01.



Patch	Lumi-nance	Pixel value	Match	Std dev
C	1001	255	8.85	0.18
D	321	213	6.16	0.40
E	321	213	6.95	0.45
B	50	145	3.95	0.45
A	1	1	1.08	0.27

Figure 20. "Simultaneous Contrast" target along with patch identification, the luminance values measured in the original display, the digit representing log luminance, the mean and standard deviation of observer matches in Munsell Value units.

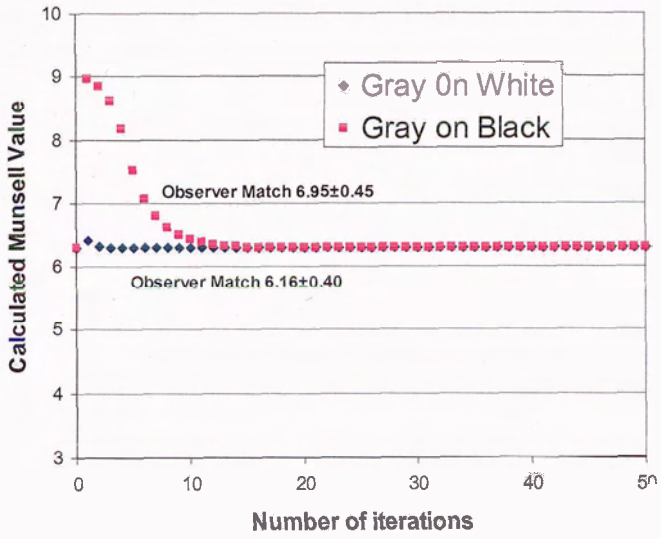
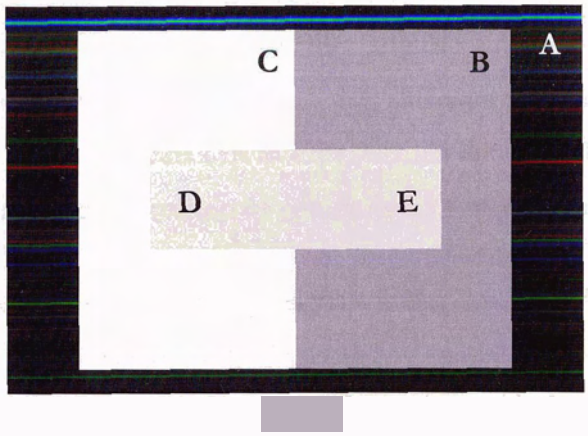


Figure 21. The best fit is 6 iterations for Gray on White and 8 for Gray on Black.



Patch	Lumi-nance	Pixel value	Match	Std dev
C	1001	255	9.09	0.29
D	321	213	6.2	0.50
E	321	213	6.7	0.58
B	50	145	4.04	0.54
A	1	1	1	0.25

Figure 22. "Simultaneous Contrast - Strip". Best fit is 6 iterations for Gray on White and 8 iterations for Gray on Black.

Discussion of Results

The principle effect of selecting the number of iterations is to establish the degree of local vs. global influence from spatial comparisons. As seen in the above data, it has no effect on grays in a white surround and significant effect on grays in a black surround. Using a very large number of iterations so as to have the lightness asymptote to the limit of the calculation makes the output approach the input.[85] That special case serves no purpose. Human observers make matches consistent with mechanisms that are between local and global. McCann, McKee and Taylor[1] reported good fit from their experimental data using a path algorithm of length 200 hops, a moderately global process. In Chapter 5, we reported 33 iterations for experiment that applied the same number of iterations for all spatial channels. In these experiments it is clear that an intermediate number gives the best results for the "Scale on Gray" target (Figure 15) and "Scale on Black" target (Figure 17). In addition, the best fit to observer data is with very few iterations with larger gray patches in the Simultaneous Contrast series. Seven iterations gave the best fit.

The displays that required the fewest iterations had large uniform surrounds. The scale displays had slightly smaller test patches and there were many more of them. The Mondrians had many more patches with smaller angular subtends. This, combined with results of other recent experiments,[94,102] suggest that different number of iterations in each spatial channel will give the best overall fit to experimental data. Frankle and McCann used a table to control the number and direction of comparisons for each spatial channel.

Larger simple displays generate large signals in the low spatial frequencies or highest levels of the image pyramid. These channels need few spatial comparisons. Scales displays generate signals with higher spatial-frequency information and these channels best fit the observer data with more iterations. The colour Mondrians have the most high spatial-frequency information and these channels need the highest degree of spatial comparisons.

Automatic Selection of the Number of Iterations

To investigate the advisability of automatic processes to measure the optimal number of iterations (i.e., cycles of comparing a pixel to its neighbours at each pyramid level), we plotted the RMS (root mean square) error between the mean lightness values reported by

human subjects and those predicted by Retinex as a function of the number of iterations. The variation in error is shown in Figure 20 ,for the case of the SC and GW data from Figure 17 through Figure 20.

Since subjects reported a single lightness value for each patch, we calculate the Retinex lightness of a patch as the mean of the Retinex lightness values for all pixels from the patch. The Retinex prediction error for a patch, therefore, reflects the difference between the Retinex lightness estimate and the mean across all subjects of the lightness of the matches made for that patch. The overall prediction error for a target is the simply the RMS of the errors for the individual patches it contains.

For the simultaneous contrast targets (SC), the minimum target prediction error occurs when the number of iterations is small, as can be seen from Figure 23. The line labelled "GW" shows the average RMS error of Retinex predictions in lightness units for the case of a target (Figure 19) in which there are three areas: the gray centre, the white surround and the black background. At one iteration, with a linear post-lut that expands the dynamic range of the raw Retinex output to [0..1], the RMS value is 0.9. That is much larger than the standard deviation of observer results of 0.52, 0.23 and 0.13. Increasing the number of iterations to 10 causes a drop in RMS values to 0.2 units. From 10 to 50 iterations the values drop from 0.2 to 0.1. For this target, any number of iterations over 5 does reasonably well at matching the observer data.

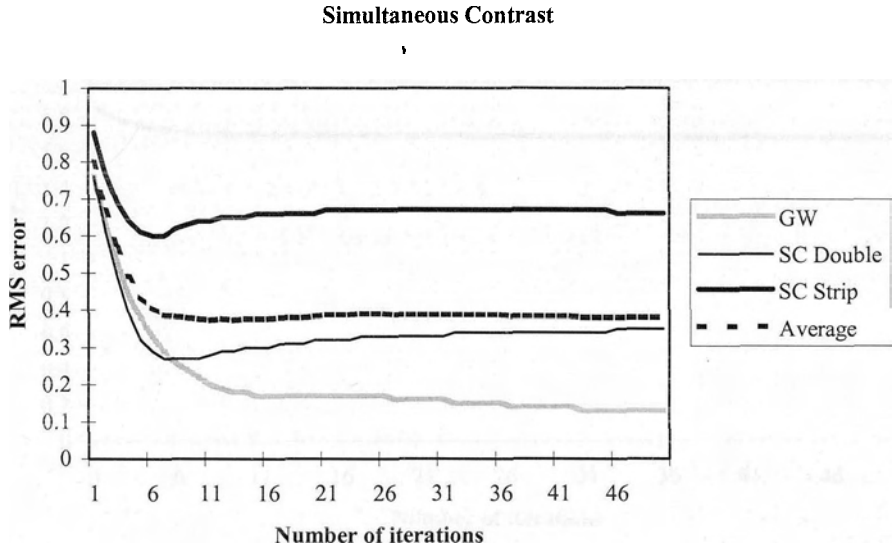


Figure 23. Simultaneous Contrast (SC) and Gray on White (GW) targets: RMS error measuring the difference between Retinex lightness predictions and subjects' reported matching lightness as a function of the number of iterations.

The thin line labelled "double" represents the data from Figure 20. In this simultaneous contrast target the prediction error (average error over all patches) is at a minimum around 6 or 7 iterations. This is because the dark gray surround and the gray area within the dark gray surround are very sensitive to the number of iterations. This target is of particular interest because the two central grays have different perceived lightness values although the patches have the same luminance. With too few iterations the calculated value for the gray in black is too high. At the point of minimum error, the calculation renders the gray-in-black one lightness unit higher than the gray-in-white. This actually conforms to the observer's predictions for this target. When the number of iterations is increased beyond 7, Retinex reports that the two grays are almost identical in lightness. This means that with too many iterations the simultaneous contrast effect is no longer predicted correctly.

Figure 24 shows the average error for the targets from the combined MMT, SB, SC and GW data sets versus the number of iterations. The minimum error now occurs when the number of iterations is quite large; although, the curve is quite flat so the minimum is also not very distinct.

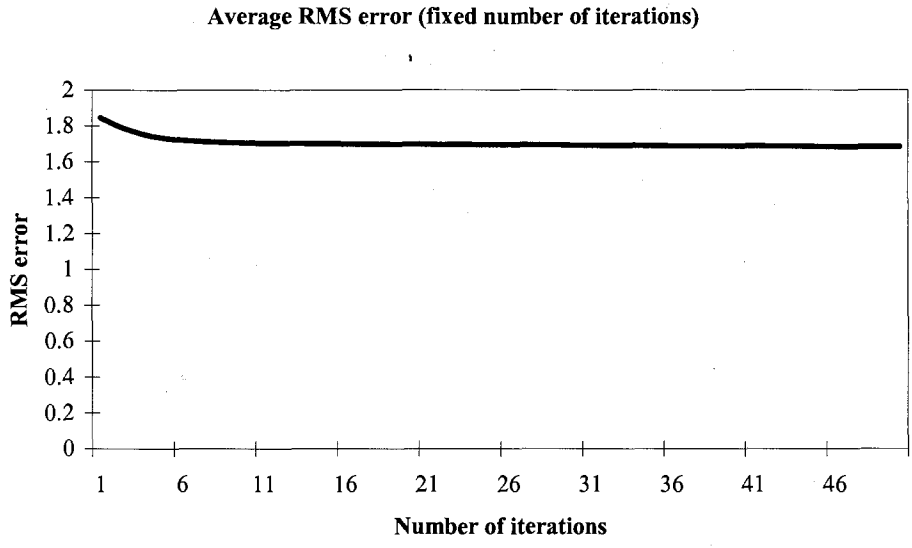


Figure 24. RMS error in Retinex lightness prediction averaged across MMT, SB, SC, GW experiments as a function of the number of iterations. For each choice of the number of iterations parameter, the same choice is then used for Retinex for all targets.

From Figure 15, Figure 17, Figure 21, Figure 23 and Figure 24 it is clear that there is no single optimal choice for the number of iterations based on minimizing the RMS error measurement alone. The number of iterations required to minimize the error for one target does not necessarily minimize the error for other targets. Therefore, a stopping condition providing a method of adjusting the number of iterations automatically on a case-by-case basis is required. Note that the stopping condition cannot be based on minimizing the RMS error directly, since the subjects' matches are not available to Retinex – the lightness matches are, after all, what Retinex is supposed to be predicting.

We introduce and test two possible stopping conditions: one based on the relative change in Retinex output,[103] the second based on the average brightness of the Retinex output. We will refer to them as the change-based and brightness-based stopping conditions. The change-based condition measures the change in Retinex output as the number of iterations is increased from n to $n+1$ and stops when the change becomes small. Although this is analogous to the situation of numerical solution of a typical optimization problem where the minimization process is iterated until the change becomes small enough, it is not precisely the same. The difference is in the meaning of the term 'iteration'. In the optimization case, the entire process is repeated until convergence; whereas, in the Retinex

case, the processing is not being repeated in its entirety. Here the number of iterations denotes the number of times the process of cycling through the neighbours is repeated at each level.

Let R_x^n be the Retinex output at location x when Retinex's iterations parameter has been set to n. The change-based Retinex stopping condition for an image of N pixels and threshold ε can be expressed as:

$$\sqrt{\frac{\sum_x (R_x^{n+1} - R_x^n)^2}{N}} \leq \epsilon \qquad (20)$$

Using this stopping condition, the number of Retinex iterations will vary with the input target. What is the optimal value of ε? We determined an optimal value for it by brute force search. In other words, we chose an initial high value for ε, ran Retinex on all the test targets and calculated the RMS prediction error, decreased ε by a small amount and repeated the process. A minimum occurs at $\varepsilon = 0.015$. The average prediction error drops to 0.62. In comparison, the minimum average error for any fixed choice of the number of iterations (as shown in Figure 24) was 1.71.

The second brightness-based stopping condition is based on the observation that Retinex reaches an optimal solution for bright targets (ones for which the average of all image pixel values is high) at fewer iterations than for dark ones. This effect can be seen in the "Scale on White", "Scale on Gray" and "Scale on Black" targets (Figure 14, Figure 15, Figure 17). The "Scale on White" target, a quite bright one, requires just 3 iterations. On the other hand, the darker "Scale on Gray" and "Scale on Black" targets require 28 iterations and 30 iterations, respectively. These are the individual number of iterations for each target that would give the best correlation with the observer matches. Intuitively, the correlation between average brightness and the optimal number of iterations is to be expected because Retinex proceeds by subtracting from white, which has the highest average brightness. At 0 iterations, the Retinex output consists of a white image (all pixels set to 1). After each successive iteration, the average brightness of the image goes down. At an infinite number of iterations, the Retinex output image would equal the input image scaled by the maximum value in each channel.

As with the change-based stopping condition, we run the Retinex algorithm at *1, 2, ...n* iterations until the stopping condition is reached. The brightness-based stopping condition is reached when the current average brightness of the Retinex output image exceeds 110% of the average brightness of the input scaled by its maximum value. The 110% value was determined empirically. The resulting slight increase in the overall image brightness can be compensated for in the Retinex post-lut. Since scaling the input by its maximum value is equivalent to the Retinex output in the limit as the number of iterations approaches infinity, the stopping condition in essence is comparing the average lightness estimate at *n* iterations to what it would converge to at an infinite number of iterations.

This new brightness-based stopping condition yields better results than the previous incremental-change-based stopping condition[103] in that the Retinex lightness estimates correlate better with the observer predictions. The average prediction error drops to 0.51 (brightness-based) from 0.62 (change-based). Either stopping condition error is substantially less than the 1.71 obtained in the optimal fixed-iteration case. If we look at each target individually and manually choose number of iterations yielding the best prediction, we get an average error of 0.39. This gives a lower bound on the error that we could obtain with a perfect stopping condition.

Conclusion

Our goal has been to study the effects of number of iterations in the special case where all spatial channels use the same number of iterations. Further this study uses the same pattern of spatial comparisons. However, Retinex requires the parameters 'postlut' and 'number of iterations' be set. In this paper, we introduce methods for setting these parameters automatically. Using these methods, Retinex yields an average RMS prediction error of only 0.51 units on a 1-to-9 lightness scale in predicting the available psychophysical data. By comparison, optimization for a fixed setting for the number of iterations resulted in an overall average RMS error of 1.71, so the new automatic-stopping-condition technique constitutes a significant improvement over a single choice for the number of iterations. Since the method changes only Retinex's input parameters, the Retinex model itself has not changed. However, the advantage of using the Retinex model in conjunction with automatic parameter selection is that it can be applied in a hands-off manner without requiring further

intervention. Future work will include modifying Retinex to employ different numbers of iterations automatically at each pyramid level.

CHAPTER SEVEN*

A LARGE DATABASE FOR COLOUR CONSTANCY RESEARCH

We present a study on various statistics relevant to research on colour constancy. Many of these analyses could not have been done before simply because a large database for colour constancy was not available. Our image database consists of approximately 11,000 images in which the (R, G, B) colour of the ambient illuminant in each scene is measured. To build such a large database we used a novel set-up consisting of a digital video camera with a neutral gray sphere attached to the camera so that the sphere always appears in the field of view. Using a gray sphere instead of the standard gray card facilitates measurement of the variation in illumination as a function of incident angle. The study focuses on the analysis of the distribution of various illuminants in the natural scenes and the correlation between the rg-chromaticity of colours recorded by the camera and the rg-chromaticity of the ambient illuminant. We also investigate the possibility of improving the performance of the naïve gray world algorithm by considering a sequence of consecutive frames instead of a single image. The set of images is publicly available and can be used as a database for testing colour constancy algorithms.

Introduction

Several image databases for colour constancy research exist [104-112]. While the complex calibration procedures involved in building such databases in which the ambient illumination is properly measured and controlled represents an asset, we are often faced by the immediate limitation of this set-up: the number of images in such a database is typically small. In some cases we would prefer to have a much larger image dataset at the expense of having a less rigorously controlled illumination. For practical considerations, it is often sufficient to measure the camera (R, G, B) of the dominant illuminants present in a scene.

---

* This chapter also appears as a published paper: Florian Ciurea and Brian Funt, "A Large Image Database for Color Constancy Research", in Proc. Eleventh Color Imaging Conference, 2003, 160-164, Scottsdale, AZ.

We construct the database using a digital video camera with a neutral gray sphere attached to the camera so that the sphere is always maintained in the field of view. Using a gray sphere instead of the standard gray card facilitates measurement of the variation in illumination as a function of incident angle. The simplicity of this set-up also facilitates recording images in locations where a conventional spectrometer would be impractical. Similarly, we can record a lot of images with relative ease. This large database with the illumination measured separately for every image allows us to study the statistics of illuminants and colours arising in a wide range of common scenes.

Setup of capture system

The database of approximately 11,000 images are of a variety of indoor and outdoor scenes, including many with people in them, shot using a Sony VX-2000 digital video camera. The outdoor scenes were taken in two locations that differ significantly in geography and climate: Vancouver, British Columbia and Scottsdale, Arizona. The images are still frames extracted from video clips captured in progressive scan mode to avoid video interlacing. In progressive scan mode, the camera generates 15 unique frames per second. When extracting the still images from the acquired video clips, we used at most 3 frames from any second of video in order to keep from having almost identical images in the database.

All the camera settings were fixed with the exception of automatic focus and automatic exposure. Although it would be preferable to hold the aperture and shutter speed constant, this is not practical because of the large range of average scene brightness relative to the camera's limited dynamic range. The camera's white balance setting was set to "outdoors" and held that way whether or not the scene was an outdoor scene or an indoor one. With this setup, the colour balance of the images taken outdoors is pretty good; those taken indoors generally have a yellow-orange colour cast.

The scene illumination was measured using a smooth, small sphere connected to the video camera by means of a monopod leg as shown in Figure 25. The sphere is 4.8 centimetres in diameter and spray painted with CIL Dulux 00NN20/000 matte paint. The paint is a spectrally neutral gray of Munsell Coordinates N4.75/ and has 18% reflectance (Figure 26). The sphere is positioned so that it appears in the video image at all times.

For each image, the scene illuminant is measured in terms of the RGB values of the pixels on the sphere. This means that illuminant is specified in units that depend upon the camera. Although it might be preferable to map these coordinates to a standard coordinate system such as CIE tristimulus values, this is not easily accomplished. Firstly, it would require complete calibration of the camera, which especially for digital cameras cannot be done reliably without knowledge of the camera's internal processing. Secondly, the R, G and B sensitivity functions are probably not perfectly colorimetric. In other words, they likely are not within a linear transformation of the human cone sensitivities, which means that there would be no precise mapping from camera RGB to XYZ in any case. Given these difficulties, we decided to settle for using the camera's native coordinates for all measurements.

Figure 25. Camera with gray sphere attached

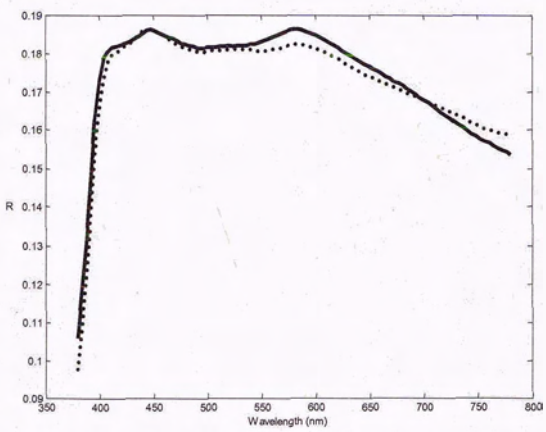


Figure 26. Reflectance spectra of gray sphere (solid line) and Munsell chip N 4.75/ (dotted line)

In order to use this camera setup effectively, we need to make sure that during video capture the illuminant cast on the gray sphere is indeed representative of the illuminant that is lighting the scene generally. In addition to being aware of the illumination conditions during video capture, we developed a post-capture validation process to discard the frames in which this condition is not met. Some video frames we discard are ones in which the dominant illuminant was coming towards the camera lens as opposed to coming from behind. Others, for example, are like those of Figure 27: an outdoor scene in which the gray sphere is temporarily in shade but the main scene is in direct sunlight.

Figure 27. A discarded image: the sphere is in the shadow, while the scene is in direct sunlight

The image database

We have recorded approximately 2 hours of digital video from which we have validated more than 11,000 images. The following images are examples of validated frames.

Figure 28. Database image (Vancouver, BC)

Figure 29. Database image (Apache Trail, Arizona)

Figure 30. Database image (Scottsdale, Arizona)

Statistics of illuminants

In general, the gray ball can be seen to have two main illuminants falling on it. For example, in Figure 29 there is a bright sunlit region and a shadow region. Hence, for each image we computed the rg-chromaticity of the two illuminants that appear on the gray sphere. The first, the dominant illuminant, is the one having the biggest impact on the scene. We determine this illuminant from the brightest region of the gray sphere. The secondary illuminant is present in the form of shading on the gray sphere. For outdoor scenes, the first illuminant is sunlight, while the second illuminant is skylight. For indoor scenes, the coordinates of the first and second illuminants are usually very similar. The following two

figures show the distribution of the first and second illuminants, respectively, in rg-chromaticity coordinates.

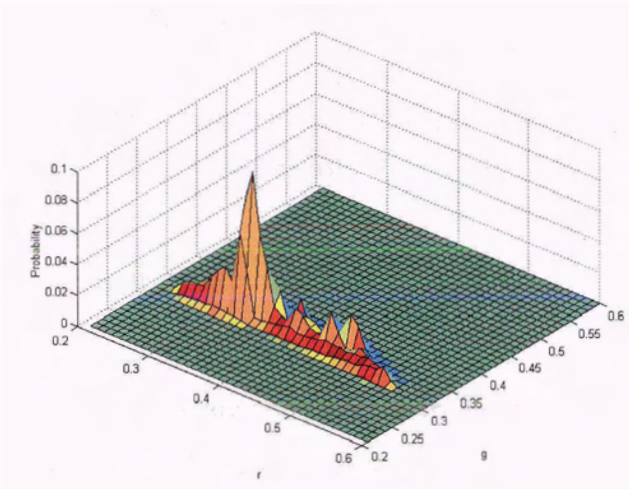


Figure 31. Probability distribution of dominant illuminant

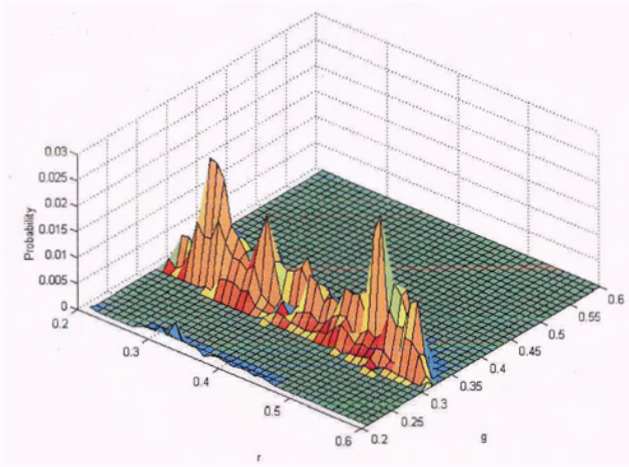


Figure 32. Probability distribution of secondary illuminant

Most discriminatory colours

We conducted an analysis to find the colours whose presence in a scene reveals the most information about the illuminant. We call these the most discriminatory colours. For this purpose we built a correlation matrix[17] relating the rg-chromaticity of image colours to the dominant scene illuminant. The rg-chromaticity space was divided into 30x30 bins. In this

discrete chromaticity space we represent both the colours measured by the camera and the colour of the illuminants. The colours measured by the camera are represented by rg-chromaticities of scenes and the illuminant colour is given by the rg-chromaticity of the dominant illuminant in a particular scene.

The probability that the chromaticity of the dominant illuminant of a scene was illum given that we observe a certain chromaticity colour is computed from Bayes' rule:

$$P(\text{illum} \mid \text{colour}) = \frac{P(\text{colour} \mid \text{illum}) \cdot P(\text{illum})}{P(\text{colour})} \tag{21}$$

All the terms in the right-hand side are computed from the image database. P(colour|illum) is the probability of observing the given colour under the particular illuminant illum. P(colour) is the probability of observing the particular colour in any scene and P(illum) is the probability that the scene was observed under illuminant illum. For a given colour we then compute the degree of discriminability as the maximum probability of any given illuminant relative to the average probability of all illuminants that correlate with the colour:

$$\text{Discriminability}(\text{colour}) = \frac{\max(P(\text{illum} \mid \text{colour})}{\text{avg}(P(\text{illum} \mid \text{colour})}, \text{ for all illum} \tag{22}$$

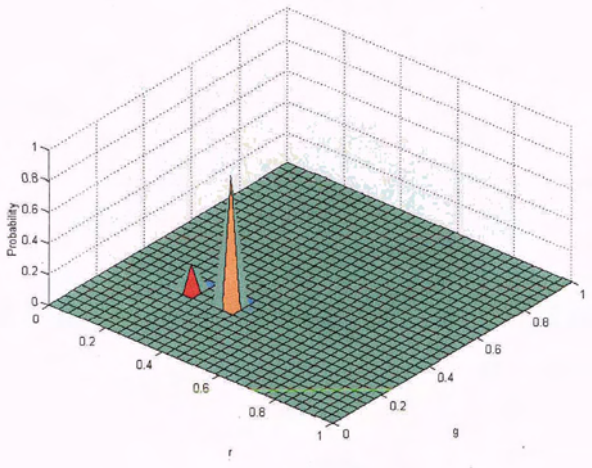


Figure 33. Probability distribution of the illuminants for a discriminatory colour

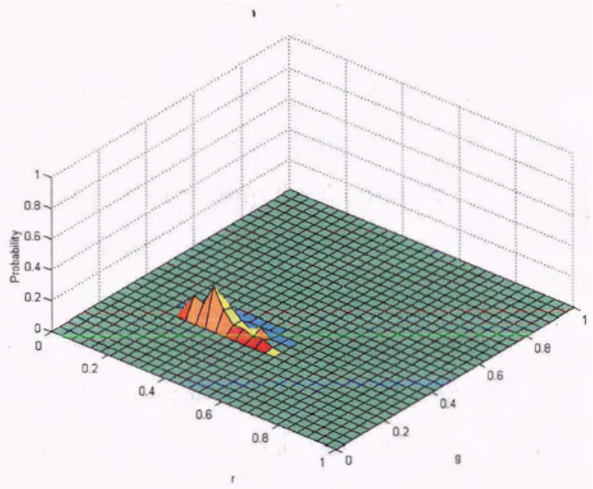


Figure 34. Probability distribution of the illuminants for a less discriminatory colour

Figure 33 and Figure 34 show the probability distribution of the illuminants for a colour with a high degree of discriminability and for a colour with a low degree of discriminability respectively. The following figure shows the discriminatory colours and the less discriminatory colours in the rg-chromaticity space. It is interesting to see that the colours that are far from white are more discriminatory than the neutrals. This means that when trying to infer the colour of the scene illuminant, the colours farther from white are the ones that contain more information about the illuminant.

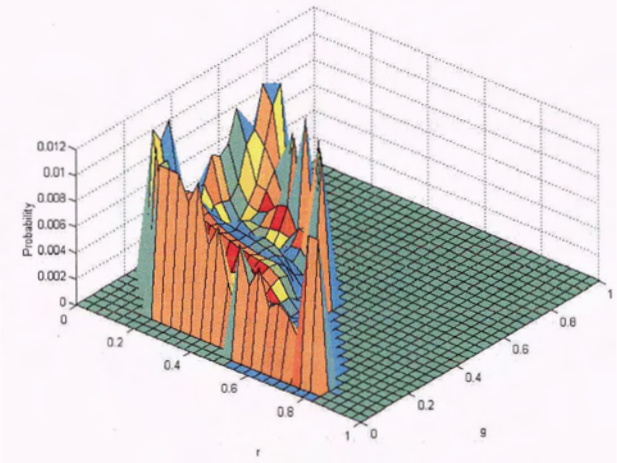


Figure 35. Distribution of discriminatory colours in the rg-chromaticity space

Gray world over time

Since the image database is derived from video sequences it means we can study colour constancy over time. Groups of frames that are close temporally in a video sequence may reveal more information about the ambient illumination than a single frame. One possibility is that the gray world method of estimating the illuminant may converge to a better answer if the average is computed over more frames.

From the database, we extracted clusters of sequential frames for which the scene illuminant, as measured from the gray calibration sphere, remains unchanged. The intent in clustering frames in this way is to find automatically clips in which the camera points at different parts of the same basic scene.

Gray world assumes that the average world is gray[35] and estimates the illuminant in the image by computing the departure of the current average RGB in the scene from that of the defined RGB of gray in the world. We simply extend the notion of a scene to a frame cluster and compute the average (R, G, B) over this set of images. Our defined gray in the world is the average (R, G, B) over the whole dataset. To measure how gray world performance might vary with time, we extract frame clusters based on video clips of 1.5, 3, 6 and 15 seconds. We use the following error measure:

$$\text{Error}(r_{est}, g_{est}, r_{act}, g_{act}) = \sqrt{(r_{est} - r_{act})^2 + (g_{est} - g_{act})^2} \qquad (23)$$

Where (r_{est}, g_{est}), (r_{act}, g_{act}) represent the rg-chromaticity of the estimated and actual (measured) illuminant respectively. The results are summarized in the following table:

	1 frame	1.5 sec.	3 sec.	6 sec.	15 sec.
Gray world avg. error	0.049	0.048	0.047	0.045	0.044
Gray world max. error	0.331	0.333	0.309	0.309	0.309

Table 6. Gray world over time

Our experience with this error measure indicates that an error of 0.02 - 0.03 in the illuminant estimation is hardly noticeable. The above results confirm that while gray world performs

reasonably well on average; however, it also fails badly in certain cases. The results also show that for the gray world, little is gained from an extended exposure to the scene. This concurs with previous experiments on adaptation to the scene average.[16]

Conclusion

We have assembled a database of more than 11,000 images in which the illuminants are measured in terms of camera RGB coordinates. This large database allows several studies on the distribution of colours in everyday scenes, the distribution of illuminants colours in everyday scenes and their mutual correlation. The database is publicly available and can also be used as a colour constancy database (with the images un-corrected for different illuminations) or as a properly white-balanced database (with the images properly colour corrected based on the illumination data).

CHAPTER EIGHT
COLOUR BY CORRELATION WITH REAL IMAGES

In this chapter we present an extension to the colour by correlation method inspired from the database introduced in Chapter 7.

A limitation of the colour by correlation method for colour constancy as described in[17] is that it assumes that, in the Bayesian framework, all illuminants are essentially equally likely and all colours (chromaticities) are equally probable. This assumption was primarily based on the fact that the actual statistics of colours and illuminants in the world is either not known, or not easily measurable. The large colour constancy database presented in the previous chapter allows gathering of such statistics that could be used to improve the effectiveness of the method on real images.

Introduction

In the Bayesian framework on which the colour by correlation method is based:

$$P(\text{illum} \mid \text{colour}) = \frac{P(\text{colour} \mid \text{illum}) \cdot P(\text{illum})}{P(\text{colour})} \cong P(\text{colour} \mid \text{illum}) \qquad (24)$$

These assumptions further lead to the conclusion that the probability that the scene illuminant was illum given that we observe the chromaticity colour P(illum | colour) is the same as the probability of observing the chromaticity colour given that the scene illuminant is illum: P(colour | illum). In reality, these probabilities are not equal, as not all illuminants are equally likely. With additional information derived from our large database of images and the corresponding illuminants, we expect to obtain better results using colour by correlation, even if only for the camera used to capture the image database.

We used the "gamut" program for colour constancy developed at Simon Fraser University, which is publicly available.* The rg-chromaticity space is divided in 50 × 50 bins for both the colours in the image and the illuminant colour. We count each colour appearing in any given image in the database only once and we make a note of all colours "observed" under a certain illuminant. For example, a given chromaticity will be counted five times if it appears in five distinct images of the database.

Results using real images

Table 7 illustrates the results for using the actual statistics derived from the large colour constancy database, as well as the results of the colour by correlation method trained on the same image data but without the additional statistics. The columns "C-by-C synthetic data" and "C-by-C synthetic data w. statistics" represent the variant in which the colour by correlation look-up table is constructed from a set of reflectances and illuminants, rather than actual images. For comparison, we also include the performance of the gray world algorithm.

	Gray world	C by C real data	***C-by-C real data. w. statistics***	C-by-C synthetic data	C-by-C synthetic data w. statistics
Avg. RMS error	0.054	0.055	***0.050***	0.109	0.059
Median RMS error	0.039	0.052	***0.035***	0.096	0.047

Table 7. Illuminant estimation (RMS error in rg-coordinates) on the large database for colour by correlation with real data statistics. For comparison, we include results for "colour by correlation without statistics" and the corresponding colour by correlation variants using synthetic data.

It should be noted that, from subjective experience, an RMS error (see Equation 22) in rg-illuminant estimation of around 0.03 is hardly noticeable and so an improvement from 0.05 to 0.035 is arguably significant.

* http://vision.cs.arizona.edu/kobus/research/programs/colour_constancy/index.html

Conclusion

These results indicate that using actual statistics derived from a large colour constancy database improves the performance of the algorithm. The limitation of the proposed setup resides in the fact that the actual statistics are only valid for the camera used for image capture. Future work includes computing joint probabilities for observable chromaticities under certain lights. Under the current framework, these probabilities are assumed to be independent and we know that this is hardly a realistic assumption.

CHAPTER NINE
HIGHER ORDER STATISTICS ON REAL IMAGES

This section continues the applications to the large database of images proposed in Chapter 7. We found this idea interesting but not very thoroughly tested. In particular, tests on real images were limited to scenes made up artificially from hyperspectral data,[113] spectral power distributions of various daylight illuminants, and the human cone sensitivity functions. The Ruderman database[113] of hyperspectral images is also quite peculiar because it consists of a small number of images of mostly foliage. Our experiments show that for scenes composed from a more diversified hyperspectral database combined with real illuminant spectra, the predicted correlation turns out to be very weak. For digital camera images, the redness-luminance correlation becomes unidentifiable.

Introduction

We want to test whether such a method is robust enough so that it can work effectively on real images. We have devised several sets of experiments in which to measure the luminance-redness correlation on real images. Experiments 1-4 are done on images synthetically generated from real hyperspectral data and full spectral information of illuminants. For experiment 5 we use real images from the SFU dataset[107] and for experiment 6 we used real images from the large database described in Chapter 7. In all the cases, we compute the redness-luminance correlation for each image of a scene under a certain illumination condition and then the correlation of the whole cluster of images under any given illumination (see figures below).

Experiment 1

For this experiment, images were generated using the hyperspectral database collected by Ruderman et al.[113] under standard illuminants D40, D55, D85 and D200. This is a replication of the experiment on real data reported by Macleod and Golz.[55]

Experiment 2

In this experiment, images were constructed from the hyperspectral database[114] collected by Nascimento et al. and standard illuminants D40, D55, D85 and D200.

Experiment 3

For this experiment we used the Ruderman hyperspectral database and the following three illuminants from the SFU calibrated database[107]: Philips Ultralume Tube (PH-ULM), Sylvania Cool White Tube (SYL-CWF), Sylvania Warm White Tube (SYL-WWF) and Solux 4700K with full Blue 3202 filter (Solux).

Experiment 4

In this case the images were constructed from the hyperspectral database by Nascimento et al.[114] and the following three illuminants from the SFU calibrated database[107]: Philips Ultralume Tube (PH-ULM), Sylvania Cool White Tube (SYL-CWF), Sylvania Warm White Tube (SYL-WWF) and Solux 4700K with full Blue 3202 filter (Solux).

Experiment 5

We used real images of 8 scenes from the "Mondrian" set of the calibrated database of images from the SFU dataset[107] acquired with the Sony DXC-930. These indoor scenes were taken under controlled laboratory conditions in which the spectral power distribution of the illuminant is measured. The "Mondrian" set consists of images with minimal specular reflections. We used the same set of three illuminants as in the previous experiments using the SFU dataset.

Experiment 6

In this experiment we tested the redness-luminance correlation on the real images in the large database acquired with the Sony VX-2000 video camera. These images represent mostly natural scenes, both indoor and outdoor. As described in Chapter 7, this database contains a description of the chromaticity of the illuminant in each of the scenes, given in camera RGB coordinates. Unlike in the first five experiments, here we do not have the same

scene available under multiple illuminants. We simply make two large classes of illuminants labelled "red" and "neutral" based on the information about the illuminant chromaticity available from the gray ball present in every scene.

Results on real images

The first experiment consists of reproducing the results on the Ruderman hyperspectral database under illuminants D40, D55, D85 and D200 using the human cone sensitivities. We see a strong correlation between redness-luminance correlation and mean redness of the scene. The values for the correlations are respectively: -0.56, -0.65, -0.65, -0.55 and are in accordance to the values reported by Macleod and Golz.[55] These values are significantly different from zero and so they indicate a strong correlation. Figure 37 illustrates the results of experiment 1, while Figure 36 shows the Ruderman database of hyperspectral images.

We believe that the dataset chosen for the original experiments by Macleod and Golz was somewhat limited and biased towards scenes with mostly foliage. In the second experiment, we used the 8 images from the hyperspectral database by Nascimento et al.,[114] which is more balanced in the types of scenes used. In this case, we obtain lower correlation scores for each cluster of images. The values are -0.40, -0.25, -0.09 and 0 respectively for the illuminants D40, D55, D85 and D200. The results are illustrated in Figure 39. Experiments 3 and 4 are similar with experiments 1 and 2 respectively, with the exception that four real illuminants have been used instead of the ideal daylight sources proposed by Macleod and Golz. For experiment 3, the results are illustrated in Figure 40. The values for the correlations are -0.62, -0.47, -0.31, -0.20. While the pattern of correlation is still visible in experiment 3, this pattern is hardly visible in experiment 4 (see Figure 41). In this case, the correlation values are much lower: 0.28, 0.18, -0.01, 0.12 and in this case the redness-luminance correlation is virtually unidentifiable.

Experiments 5 and 6 are based on real images. Eight scenes from the SFU database[107] have been used in experiment 5, under each of the four chosen illuminants. In this case, the correlation values are even lower: 0.11, 0.19, -0.03, 0.21 (see Figure 42). For this experiment we also note that the redness-luminance correlation is unidentifiable. The final experiment is based on the large database of images introduced in Chapter 7.

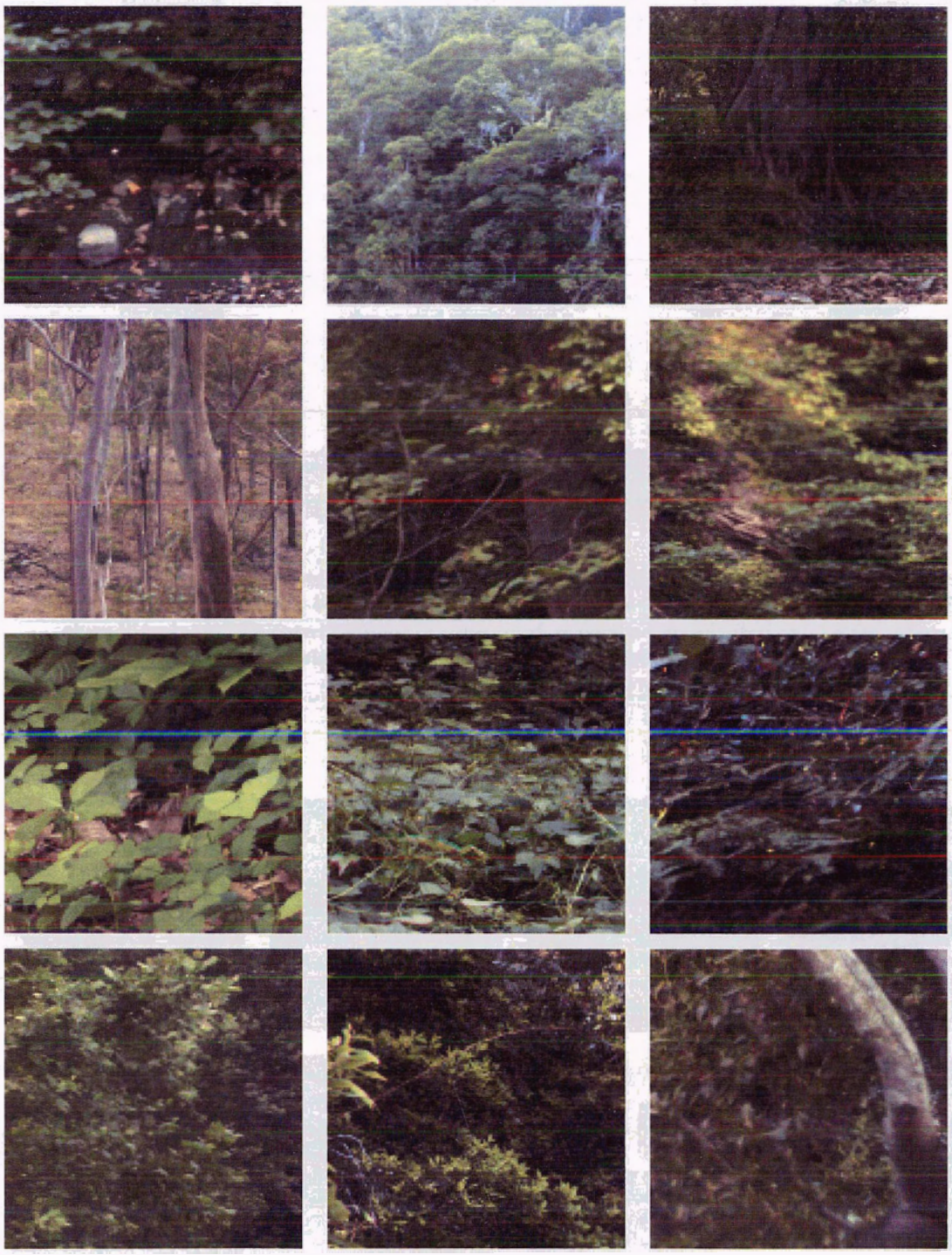


Figure 36. Hyperspectral database of 12 images by Ruderman et al.

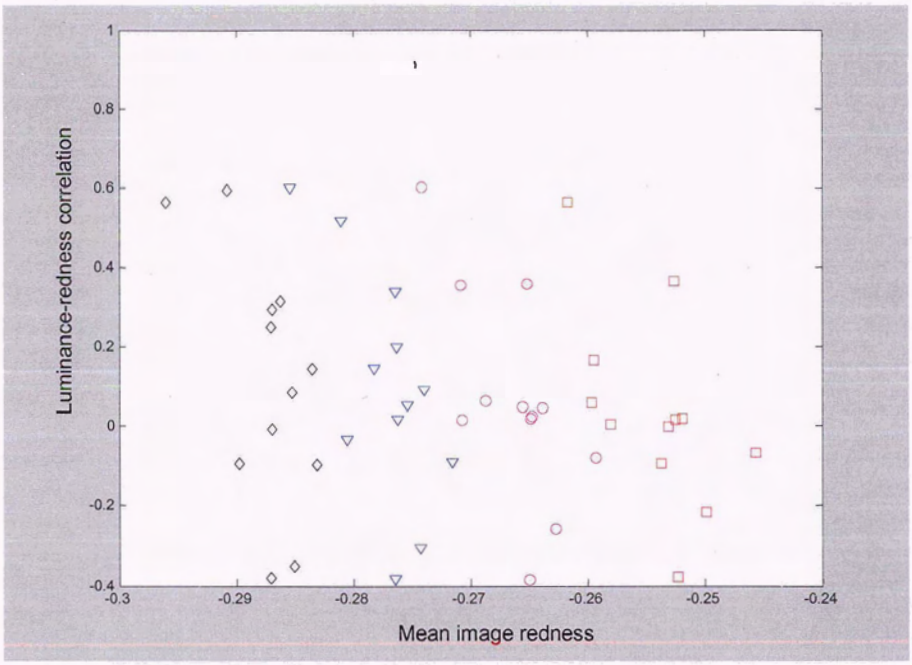


Figure 37. Experiment 1 in redness-luminance correlation. The 12 images are represented by black diamonds (illuminant D40), blue triangles (D55), pink circles (D85) and red squares (D200). For a given illuminant, the correlation for the cluster of images is around –0.55 and so the redness-luminance correlation (the Y axis) is able to explain some of the variation.

Figure 38. Hyperspectral database of 8 images by Nascimento et al.

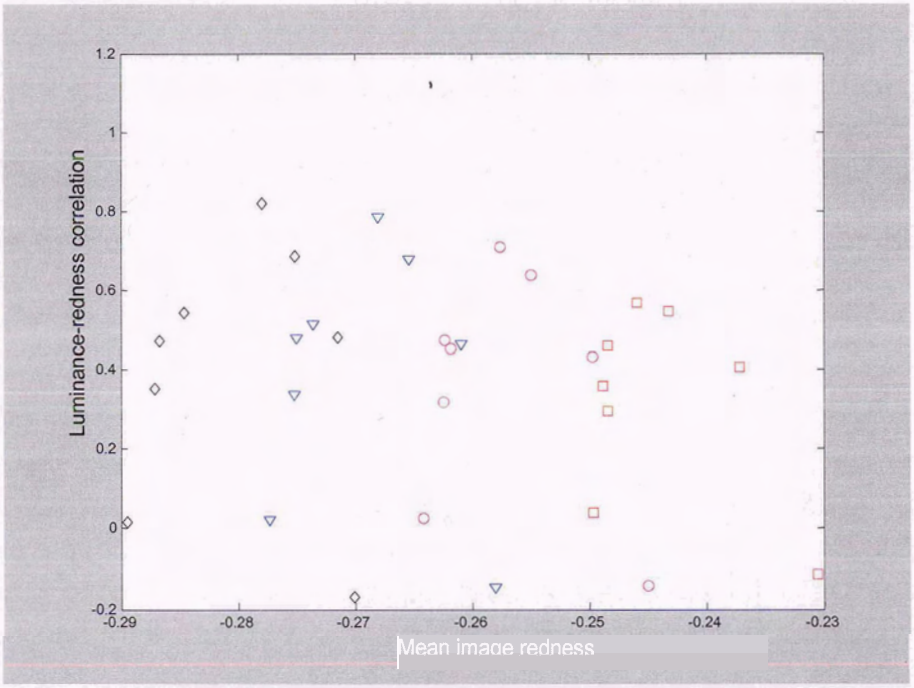


Figure 39. Experiment 2 in redness-luminance correlation. The 8 images are represented by black diamonds (D40), blue triangles (D55), pink circles (D85) and red squares (D200). For a given illuminant, the correlation now varies from –0.40 to 0.00

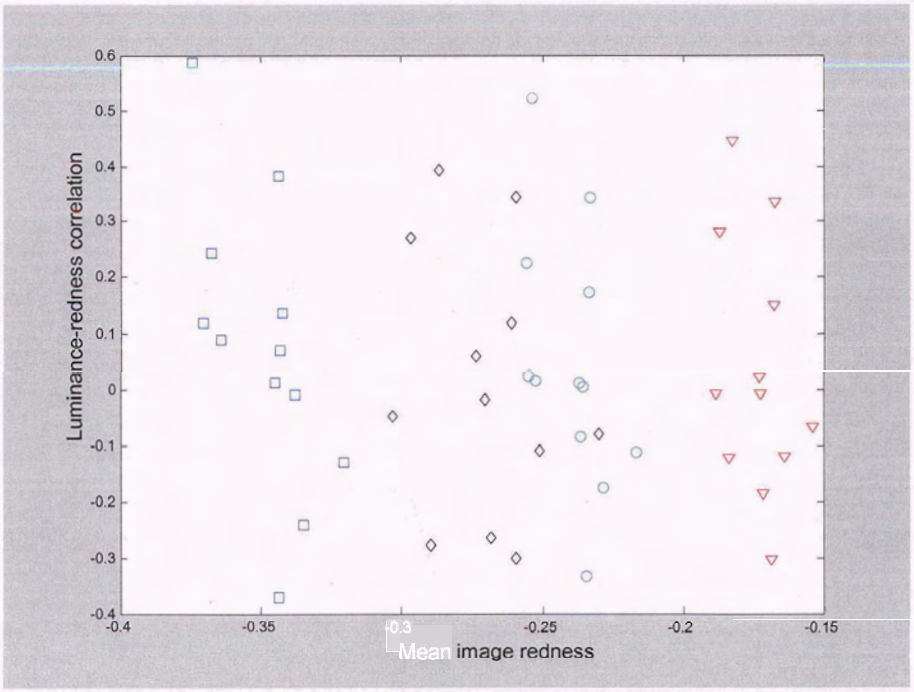


Figure 40. Experiment 3 in luminance-redness correlation. The 12 images are represented by black diamonds (PH-ULM), blue triangles (Solux), pink circles (SYL-CWF) and red squares (SYL-WWF). The correlations now range from –0.62 to –0.20.

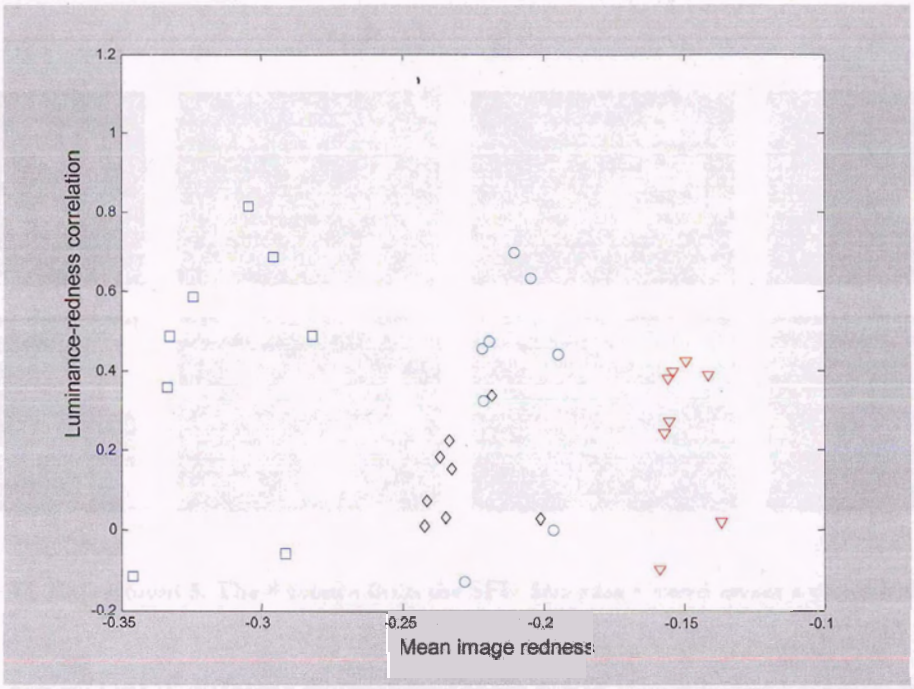


Figure 41. Experiment 4 in luminance-redness correlation. The 8 images are represented by black diamonds (PH-ULM), blue triangles (Solux), pink circles (SYL-CWF) and red squares (SYL-WWF). The correlations are poor, ranging from –0.01 to 0.28

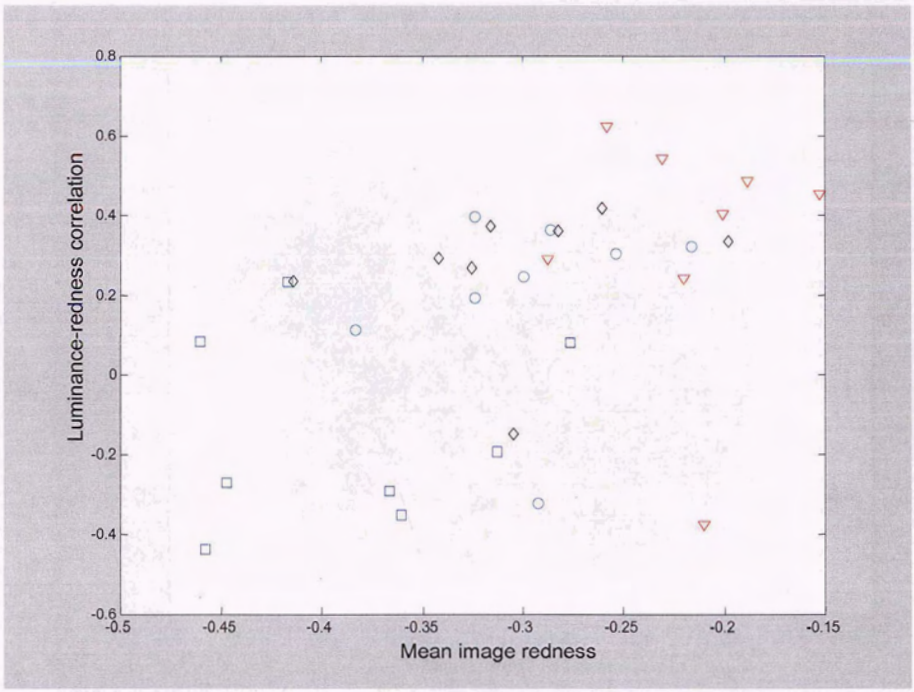


Figure 42. Experiment 5 in redness-luminance correlation. These are real images of the same scene under different illuminants. The 8 images are represented by black diamonds (PH-ULM), blue triangles (Solux), pink circles (SYL-CWF) and red squares (SYL-WWF).There is no pattern in this figure, with the values for the correlations ranging from –0.03 to 0.21.

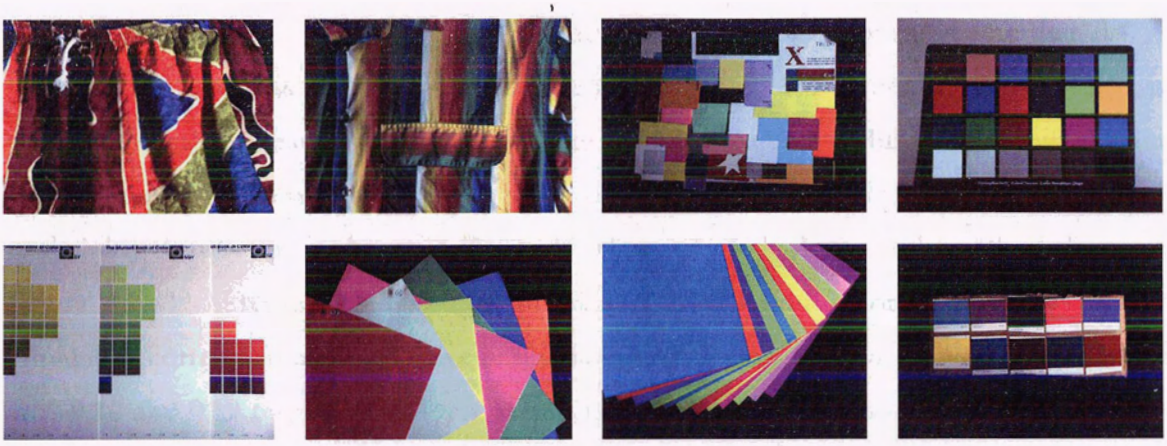


Figure 43. Experiment 5. The 8 scenes from the SFU database viewed under a canonical white illuminant.

The sixth experiment is different from experiments 1-5 in that it is not based on having the same scene viewed under different illuminants. Here, we are simply interested in the probability distribution of scenes under "reddish" or "neutral" illuminant.

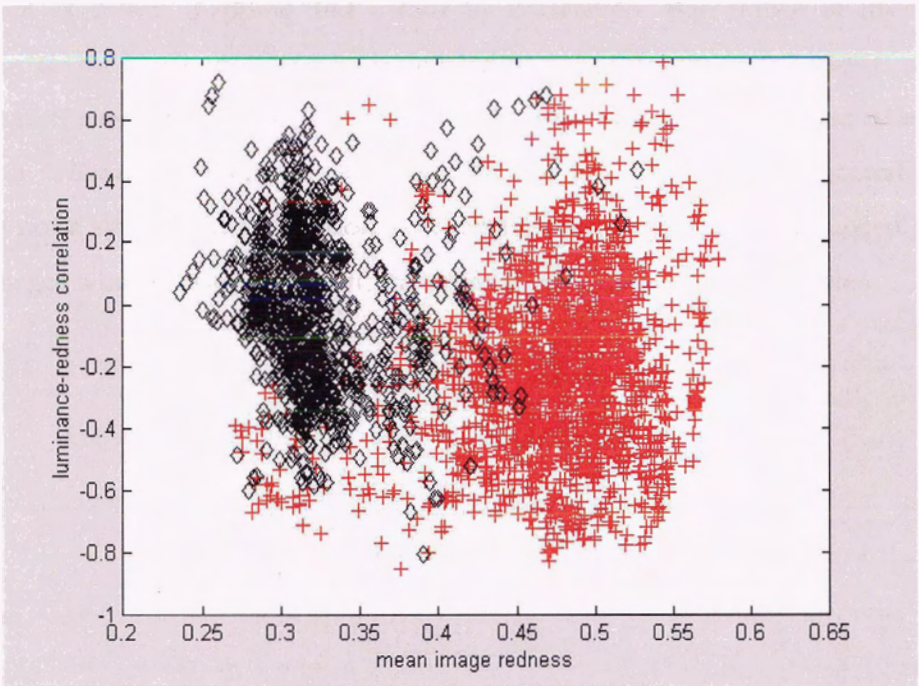


Figure 44. Experiment 6. The probability distribution of scenes under neutral illuminant (black crosses) and the probability distribution of scenes under reddish illuminant (red stars) in a two-dimensional space of mean image redness and luminance-redness correlation.

A "reddish" illuminant is defined as having the r chromaticity component greater than 0.4. A "neutral" illuminant is defined as having the r chromaticity component less than 0.3. It is worth mentioning that, as shown in Figure 31, the major variability in the illuminant chromaticity is expressed by the red-blue component (r chromaticity for instance). The probability distribution is shown in Figure 44. Here, we are looking for the pattern shown in Figure 3, which represents the ideal probability distribution in order for the redness-luminance correlation to be effective in illuminant estimation. Instead, the correlation values are –0.10 and 0.19 respectively. These results indicate that the redness-luminance correlation is again unidentifiable.

Conclusion

We have devised a set of experiments to test the redness-luminance correlation hypothesis on real data. This theory is interesting in that it relates the gray world approach which is based on the scene average alone with the more advanced algorithms for illuminant estimation such as gamut mapping and colour by correlation. But, unlike in the results reported by Golz and Macleod[56, 55] on hyperspectral images, this correlation proves to be a weak measure in the case of real images. Even on a different dataset of hyperspectral images, this correlation is lower. Furthermore, when adding real illuminants whose spectral power distribution is not a smooth, ideal function - this correlation is even less pronounced. In the case of real images where noise is inherent, the correlation becomes unidentifiable.

CONCLUSION

The colour constancy problem is under-constrained by nature. The most successful and complex methods try to attack the problem by reducing its complexity. Linear models like those introduced by Maloney and Wandell assume a reduced dimensionality of the surface reflectances and of the illuminants. Statistical based methods such as the gamut mapping approach, colour-by-correlation and neural-network colour constancy have proven to be the most successful implementations to date in scenes where specular reflections are either not present or difficult to identify. The success of these algorithms seems to suggest that statistical approaches to colour constancy are desirable, given that this is usually the way the colour constancy problem is regarded. A common denominator of these approaches is that they do not make use of the spatial information present in the scene, and rather rely solely on colour image histograms as a good descriptor for the scene illuminant. These methods also assume a single illuminant in the scene, and usually set out to find the chromaticity of this dominant illuminant. The Retinex model is a somewhat different approach, in that it exploits the spatial interactions in a scene and in that it computes the illuminant and reflectance component at each location (pixel). An interesting aspect of Retinex is that it goes beyond colour constancy – being proposed as a model of human colour vision – and trying to explain vision perception phenomena such as simultaneous contrast.

The work presented here introduces a series of advances in the area of colour constancy and colour appearance based on experimental data and real images. The first contribution is in the form of a series of experiments on chromatic adaptation transforms. Traditionally, at the core of chromatic adaptation transforms there is the same fixed transformation matrix that is used regardless of the two illuminants involved in adaptation. Based on the belief that this approach has been historically motivated by the use of colorimeters that could only measure tristimulus values for the two illuminants involved, a new method is proposed that could be used in the case where the full spectral power distribution of the two illuminants is available or can be measured. In this case, the transformation is tuned for the particular illuminant pair, thus possibly improving the

performance. Indeed, under some particular cases in which the spectral power distribution of the illuminants is atypical, the method performs better.

The second important contribution is the work towards characterizing the Retinex model such that it can be used, without intervention, on real images as a model for the human visual system. The Retinex model has been quite controversial over the years mainly because its authors have not provided a very clear implementation and an explanation of the influence of the models' parameters and a discussion of how these parameters should be chosen for a specific application. Even though the model is essentially computational, its way of application has been historically somewhat mysterious. One would often hear in the colour community that the Retinex model has failed under certain circumstances just because of the users' lack of knowledge in choosing or tweaking the parameters. The study presented here provides a clear implementation of two of the most successful Retinex variants, along with a discussion on choosing the parameters. A second part of the study has been devoted to finding values for these parameters in a practical manner for a specific application. Finally, a method for automatic choosing the parameters is also included, which finally provides a hands-off approach to the Retinex model. In this context, new unpublished psychophysical experiments contributed by McCann are also presented.

Another important contribution is represented by a novel setup that permits the acquisition of a large database for colour constancy research. Using a spectrally neutral gray sphere attached to a video camera, we can capture a large number of images for which the illumination conditions are documented within the camera coordinates. Traditionally, databases for colour constancy research have been limited in the number of images available because of the complex calibration procedures involved in measuring the full spectral power distribution of the illumination sources. Also, these scenes are usually restricted to indoor composition. With the proposed mobile setup – at the expense of a less rigorous measurement of the illumination – we can capture a series of scenes and their illumination conditions under many natural conditions. Additional statistics on the illuminants have been obtained from such a large database of over 11,000 images. Furthermore, using such a large database of real images allows the testing and improvement of several existing methods for colour constancy.

The fourth contribution is in the form of application of the large colour constancy database of images, an improvement in, the estimation of one of the currently most successful algorithms for colour constancy, "colour by correlation". In the lack of real prior probabilities about the distribution of illuminants and colours in the world, the current implementation assumes that all illuminants and surfaces are equally probable. Another aspect is the fact that in the original version, the correlation matrix, which is at the core of the algorithm, has been computed based on synthetically generated images, as opposed to using real data. Using real images and therefore providing real probabilities improves the performance of the method.

Finally, the fifth contribution is represented by the study of the application of a yet another method for colour constancy based on higher-order statistics present in images. This method has not been studied on real images and, unless tested in this manner, cannot be used effectively in estimating the scene illuminant. Several sets of experiments for testing the effectiveness of the method on real data have been devised, consisting of a) a large database of real images in which the illuminant is known; b) a smaller calibrated database of images that contains multiple views of the same scene under different illuminants and c) a more diversified database of hyperspectral images. Even though the method involving the redness-luminance correlation in scenes is appealing both conceptually and for its simplicity, it proves to be of little use in the context of real images.

APPENDICES

Matlab implementation of McCann99 Retinex

L: logarithmic input image

nIterations: number of iterations for each pixel

nLayers: number of pyramid layers

OP: matrix of Old Products for all pixels

RR: input radiance

NP: matrix of New Products for all pixels

OPE: OP padded with zeros for doing all computations at once

RRE: RR padded with two additional columns and two additional rows

IP: Intermediate Product computed with the Ratio-Product

```
function Retinex = Retinex_mccann99(L, nIterations)
global OPE RRE Maximum
[nrows ncols] = size(L);                          % get size of the input image
nLayers = ComputeLayers(nrows, ncols);            % compute the number of pyramid layers
nrows = nrows/(2^nLayers);                        % size of image to process for layer 0
ncols = ncols/(2^nLayers);
if (nrows*ncols > 25)                             % not processing images of area > 25
  error('invalid image size.')                    % at first layer
end
Maximum = 1;                                      % maximum colour value in the image
OP = Maximum*ones([nrows ncols]);                 % initialize Old Product
for layer = 0:nLayers
   RR = ImageDownResolution(L, 2^(nLayers-layer)); % reduce input to required layer size
   OPE = [zeros(nrows,1) OP zeros(nrows,1)];      % pad OP with additional columns
   OPE = [zeros(1,ncols+2); OPE; zeros(1,ncols+2)]; % and rows
   RRE = [RR(:,1) RR RR(:,end)];                  % pad RR with additional columns
   RRE = [RRE(1,:); RRE; RRE(end,:)];             % and rows

   for iter = 1:nIterations
     CompareWithNeighbor(-1, 0);                  % North
```

```
      CompareWithNeighbor(-1, 1);                    % North-East
      CompareWithNeighbor(0, 1);                     % East
      CompareWithNeighbor(1, 1);                     % South-East
      CompareWithNeighbor(1, 0);                     % South
      CompareWithNeighbor(1, -1);                    % South-West
      CompareWithNeighbor(0, -1);                    % West
      CompareWithNeighbor(-1, -1);                   % North-West
   end
   NP = OPE(2:(end-1), 2:(end-1));
   OP = NP(:, [fix(1:0.5:ncols) ncols]);             %%% these two lines are equivalent with
   OP = OP([fix(1:0.5:nrows) nrows], :);             %%% OP = imresize(NP, 2) if using Image
   nrows = 2*nrows; ncols = 2*ncols;                 % Processing Toolbox in MATLAB
end
Retinex = NP;

function CompareWithNeighbor(dif_row, dif_col)
global OPE RRE Maximum

% Ratio-Product operation
IP = OPE(2+dif_row:(end-1+dif_row), 2+dif_col:(end-1+dif_col)) + ...
     RRE(2:(end-1),2:(end-1)) - RRE(2+dif_row:(end-1+dif_row), 2+dif_col:(end-1+dif_col));

IP(IP > Maximum) = Maximum;                          % The Reset step

% ignore the results obtained in the rows or columns for which the neighbors are undefined
if (dif_col == -1) IP(:,1) = OPE(2:(end-1),2); end
if (dif_col == +1) IP(:,end) = OPE(2:(end-1),end-1); end
if (dif_row == -1) IP(1,:) = OPE(2, 2:(end-1)); end
if (dif_row == +1) IP(end,:) = OPE(end-1, 2:(end-1)); end
NP = (OPE(2:(end-1),2:(end-1)) + IP)/2;              % The Averaging operation
OPE(2:(end-1), 2:(end-1)) = NP;

function Layers = ComputeLayers(nrows, ncols)
power = 2^fix(log2(gcd(nrows, ncols)));              % start from the Greatest Common Divisor
while(power > 1 & ((rem(nrows, power) ~= 0) | (rem(ncols, power) ~= 0)))
   power = power/2;                                  % and find the greatest common divisor
end                                                  % that is a power of 2
Layers = log2(power);

function Result = ImageDownResolution(A, blocksize)
[rows, cols] = size(A);                              % the input matrix A is viewed as
result_rows = rows/blocksize;                        % a series of square blocks
result_cols = cols/blocksize;                        % of size = blocksize
Result = zeros([result_rows result_cols]);
for crt_row = 1:result_rows                          % then each pixel is computed as
   for crt_col = 1:result_cols                       % the average of each such block
      Result(crt_row, crt_col) = mean2(A(1+(crt_row-1)*blocksize:crt_row*blocksize, ...
                                       1+(crt_col-1)*blocksize:crt_col*blocksize));
```

```
   end
end
```

Matlab implementation of Frankle-McCann Retinex

```
function Retinex = Retinex_frankle_mccann(L, nIterations)
global RR IP OP NP Maximum
RR = L;
Maximum = 1;                                    % maximum colour value in the image
[nrows, ncols] = size(L);

shift = 2^(fix(log2(min(nrows, ncols)))-1);     % initial shift
OP = Maximum*ones(nrows, ncols)                 % initialize Old Product

while (abs(shift) >= 1)
   for i = 1:nIterations
      CompareWith(0, shift);                    % horizontal step
      CompareWith(shift, 0);                    % vertical step
   end
   shift = -shift/2;                            % update the shift
end
Retinex = NP;

function CompareWith(s_row, s_col)
global RR IP OP NP Maximum
IP = OP;
if (s_row + s_col > 0)
    IP((s_row+1):end, (s_col+1):end) = OP(1:(end-s_row), 1:(end-s_col)) + ...
    RR((s_row+1):end, (s_col+1):end) - RR(1:(end-s_row), 1:(end-s_col));
else
    IP(1:(end+s_row), 1:(end+s_col)) = OP((1-s_row):end, (1-s_col):end) + ...
    RR(1:(end+s_row),1:(end+s_col)) - RR((1-s_row):end, (1-s_col):end);
end
IP(IP > Maximum) = Maximum;                     % The Reset operation
NP = (IP + OP)/2;                               % average with the previous Old Product
OP = NP;                                        % get ready for the next comparison
```

REFERENCE LIST

1. J.J. McCann, S.P. McKee, and T.H. Taylor, *Quantitative Studies in Retinex Theory.* Vision Research, 1976. **16**: p. 445-458.

2. D.H. Brainard, W.A. Brunt, and J.M. Speigle, *Color constancy in the nearly natural image. I. Asymmetric matches.* Journal of the Optical Society of America A, 1997. **14**(9): p. 2091-2110.

3. D.H. Brainard, W.A. Brunt, and J.M. Speigle, *Color constancy in the nearly natural image. 2. Achromatic loci.* Journal of the Optical Society of America A, 1998. **15**(2): p. 307-325.

4. M.J. Swain and D.H. Ballard, *Color Indexing.* International Journal of Computer Vision, 1991. **7**(1): p. 11-32.

5. D.H. Hubel and T.N. Wiesel, *Brain Mechanisms of Vision.* Scientific American, 1979(241): p. 150-162.

6. S. Zeki, *A Vision of the Brain.* 1993: Oxford: Blackwell Scientific Publications.

7. D.J. Ingle, *The goldfish as a retinex animal.* Science, 1985. **227**(4687): p. 651-654.

8. S. Dörr and C. Neumayer, *The Goldfish - A Colour-Constant Animal.* Perception, 1996. **25**: p. 243-250.

9. A. Werner, R. Menzel, and C. Wehrhahn, *Color constancy in the honeybee.* Journal of Neuroscience, 1998. **8**: p. 156-159.

10. M. Kinoshita and K. Arikawa, *Colour constancy of the swallowtail butterfly Papilio Xuthus.* Journal of Experimental Biology, 2000. **203**: p. 3521-3530.

11. H. Helson, *Fundamental problems in color vision. I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination.* Journal of Experimental Psychology, 1938. **23**(5): p. 439-476.

12. D.B. Judd, *Hue saturation and lightness of surface colors with chromatic illumination.* Journal of the Optical Society of America, 1940. **30**: p. 2-32.

13. D.H. Brainard and B.A. Wandell, *A bilinear model of the illuminant's effect on color appearance.* Computational Models of Visual Processing. 1991, Cambridge, MA: MIT Press. 171-187.

14. D.H. Brainard and B.A. Wandell, *Asymmetric color matching: how color appearance depends on the illuminant.* Journal of the Optical Society of America A, 1992. **9**(9): p. 1433-1448.

15. P.K. Kaiser and R.M. Boynton, *Human Color Vision.* Second edition ed. 1996: Optical Society of America.

16. J.J. McCann. *Lessons Learned from Mondrians Applied to Real Images and Color Gamuts.* in *Proceedins IS&T/SID Seventh Color Imaging Conference*, 1-8, 1999. Scottsdale, AZ.

17. G.D. Finlayson, S. Hordley, and P.H. Hubel, *Color by correlation: a simple unifying framework for color constancy.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001. **23**(11): p. 1209-1221.

18. L.T. Maloney and B.A. Wandell, *Color constancy: a method for recovering surface spectral reflectance.* Journal of the Optical Society of America A, 1986. **3**: p. 29-33.

19. S.A. Shafer, *Using color to separate reflection components.* COLOR Research and Application, 1985. **10**: p. 210-218.

20. B.A. Wandell, *The synthesis and analysis of color images.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987. **9**: p. 2-13.

21. G.D. Finlayson, M.S. Drew, and B.V. Funt. *Diagonal Transforms Suffice for Color Constancy.* in *Fourth International Conference on Computer Vision*1993: IEEE Computer Society & European Vision Society.

22. J.v. Kries, *Beitrag zur Physiologie der Gesichtsempfinding.* Arch. Anat. Physiol., 1878. **2**: p. 5050-524.

23. G.D. Finlayson, *Color in perspective.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996. **18**(10): p. 1034-1038.

24. J.A. Worthey, *Limitations of color constancy.* Journal of the Optical Society of America [Suppl.], 1985. **2**: p. 1014-1026.

25. J.A. Worthey and M.H. Brill, *Heuristic analysis of von Kries color constancy.* Journal of the Optical Society of America A, 1986. **3**(10): p. 1708-1712.

26. G.D. Finlayson, M.S. Drew, and B.V. Funt, *Spectral Sharpening: Sensor Transformations for Improved Color Constancy.* Journal of the Optical Society of America A, 1994. **11**(5): p. 1553-1563.

27. M.D. Fairchild, *Color Appearance Models.* 1998: Addison-Wesley.

28. R.W.G. Hunt, *The Reproduction of Color.* 2 ed. 1987, England: Fountain Press.

29. K.M. Lam, *Metamerism and colour constancy (PhD thesis).* 1985, University of Bradford.

30. J. Cohen, *Dependency of The Spectral Reflectance Curves of The Munsell Color Chips.* Psychonomic Science, 1964. **1**: p. 369-370.

31. E.L. Krinov, *Spectral Reflectance Properties of Natural Formations.* Technical Translation TT-439. 1947: National Research Council of Canada.

32. J.L. Dannemiller, *Spectral reflectance of natural objects: how many basis functions are necessary.* Journal of the Optical Society of America A, 1992. **9**: p. 507-515.

33. M.J. Vrhel, R. Gershon, and L.S. Iwan, *Measurement and Analysis of Object Reflectance Spectra.* COLOR Research and Application, 1994. **19**(1): p. 4-9.

34. L.T. Maloney, *Evaluation of linear models of surface spectral reflectance with small numbers of parameters.* Journal of the Optical Society of America A, 1986. **3**(10): p. 1673-1683.

35. G. Buchsbaum, *A spatial processor model for object colour perception.* Journal of the Franklin Institute, 1980. **310**: p. 1-26.

36. R. Gershon, A.D. Jepson, and J.K. Tsotsos, *From [R, G, B] to Surface Reflectance: Computing Color Constant Descriptors in Images.* Perception, 1988: p. 755-758.

37. E.H. Land and J.J. McCann, *Lightness and Retinex theory.* Journal of the Optical Society of America, 1971. **61**: p. 1-11.

38. B.A. Wandell, *Foundations of Vision.* 1995: Sinauer Associates Inc.

39. J. Frankle and J. McCann, *Method and Apparatus for Lightness Imaging,* in *United States Patent #4,384,336.* 1983.

40. E.H. Land, *An alternative technique for the computation of the designator in the Retinex theory of color vision.* Proceedings of the National Academy of Science, 1986. **83**: p. 3078-3080.

41. D.A. Brainard and B.A. Wandell, *Analysis of the Retinex theory of Color Vision.* Journal of the Optical Society of America A, 1986. **3**(10): p. 1651-1661.

42. B.K.P. Horn, *Determining lightness from an image.* Computer Vision, Graphics, and Image Processing, 1974. **3**: p. 277-299.

43. A. Blake, *Boundary conditions for lightness computation in Mondrian world.* Computer Vision, Graphics, and Image Processing, 1985. **32**: p. 314-327.

44. A. Hurlbert, *Formal connections between lightness algorithms.* Journal of the Optical Society of America A, 1986. **3**: p. 1684-1692.

45. R. Kimmel, M. Elad, D. Shaked, R. Keshet, and S. Irwin. *A Variational Framework for Retinex.* in *Electronic Imaging*2002. San Jose, CA.

46. D. Forsyth and J. Ponce, *Computer vision: a modern approach.* 2002: Prentice Hall.

47. D.A. Forsyth, *A Novel Algorithm for Color Constancy.* International Journal of Computer Vision, 1990. **5**(1): p. 5-36.

48. G.D. Finlayson. *Color Constancy in Diagonal Chromaticity Space.* in *Fifth International Conference on Computer Vision*, 218-223, 1995: IEEE Computer Society Press.

49. K. Barnard, G. Finlayson, and B. Funt, *Colour constancy for scenes with varying illumination.* Computer Vision and Image Understanding, 1997. **65**(2): p. 311-321.

50. D.H. Brainard and W.T. Freeman, *Bayesian color constancy.* Journal of the Optical Society of America A, 1997. **14**(7): p. 1393-1411.

51. K. Barnard, L. Martin, and B. Funt. *Colour by Correlation in a Three Dimensional Colour Space.* in *6th European Conference on Computer Vision*, 275-289, 2000.

52. B. Funt, V. Cardei, and K. Barnard. *Learning Color Constancy.* in *IS&T/SID Fourth Color Imaging Conference: Color Science, Systems and Applications*, 58-60, 1996. Scottsdale, Arizona.

53. V.C. Cardei, *A neural network approach to color constancy.* 2000, Simon Fraser Univerisity.

54. V.C. Cardei, B. Funt, and K. Barnard, *Estimating the scene illumination chromaticity by using a neural network.* Journal of the Optical Society of America A, 2002. **19**(12): p. 2374-2385.

55. D.I.A. MacLeod and J. Golz, *A computational analysis of colour constancy*, in *Colour Perception: Mind and the Physical World*, R. Mausfeld, D. Heyer, and H. Dieter, Editors. 2004, Oxford University Press.

56. J. Golz and D.I.A. MacLeod, *Influence of scene statistics on colour constancy.* Nature, 2002. **415**: p. 637-640.

57. R. Mausfeld and J. Andres, *Second-order statistics of colour codes modulate transformations that effectuate varying degrees of scene invariance and illumination invariance.* Perception, 2002. **31**: p. 209-224.

58. D. Forsyth, *A novel algorithm for color constancy.* International Journal of Computer Vision, 1990. **5**: p. 5-36.

59. H.-C. Lee, *Method for computing the scene-illuminant chromaticity from specular highlights.* Journal of the Optical Society of America A, 1986. **3**(10).

60. M. D'Zmura and P. Lennie, *Mechanisms of color constancy.* Journal of the Optical Society of America A, 1986. **3**: p. 1662-1672.

61. G.J. Klinker, S.A. Shafer, and T. Kanade. *Using a color reflection model to separate highlights from object color.* in *First International Conference on Computer Vision*, 145-150, 1987: IEEE Computer Society Press.

62. L.T. Maloney, *Illuminant estimation as cue combination.* Journal of Vision, 2002(2): p. 493-504.

63. M.R. Luo and R.W.G. Hunt, *A chromatic adaptation transform and a colour inconstancy index.* COLOR Research and Application, 1998. **23**: p. 154-158.

64. C. Li, M.R. Luo, and B. Rigg. *Simplification of the CMCCAT97.* in *Proc. IS&T/SID Eighth Color Imaging Conference*, 56-60, 2000. Scottsdale, AZ.

65. G.D. Finlayson and S. Süsstrunk. *Performance of a Chromatic Adaptation Transform based on Spectral Sharpening.* in *Proc. IS&T/SID Eighth Color Imaging Conference*, 49-55, 2000. Scottsdale, AZ.

66. G.D. Finlayson and M.S. Drew, *Constrained least-squares regression in color spaces.* Journal of Electronic Imaging, 1997. **6**(4): p. 484-493.

67. K.M. Braun and M.D. Fairchild. *Psychophysical generation of matching images for cross-media color reproduction.* in *IS&T/SID Fourth Color Imaging Conference: Color Science, Systems and Applications*, 214-220, 1996. Scottsdale, AZ.

68. E.J. Breneman, *Corresponding chromaticities for different states of adaptation to complex visual fields.* Journal of the Optical Society of America A, 1987. **4**(6): p. 1115-1129.

69. H. Helson, J.D. B., and M.H. Warren, *Object-color changes from daylight to incandescent filament illumination.* Illuminating Engineering, 1952. **47**: p. 221-233.

70. W.G. Kuo, M.R. Luo, and H.E. Bez, *Various chromatic-adaptation transforms tested using new colour appearance data in textiles.* COLOR Research and Application, 1995. **21**: p. 313-327.

71. M.R. Luo, A.A. Clarke, P.A. Rhodes, S.A.R. Scrivener, S. A., and T. C.J., *Quantifying Colour Appearance. Part I. LUTCHI Colour Appearance Data.* COLOR Research and Application, 1991. **16**: p. 166-180.

72. M.R. Luo and P.A. Rhodes, *Corresponding Colour Data Sets.* 2001, University of Derby, http://colour.derby.ac.uk/colour/info/catweb/.

73. S. Süsstrunk, J. Holm, and G.D. Finlayson. *Chromatic Adaptation Performance of different RGB Sensors.* in *Proceedings IS&T/SPIE Electronic Imaging 2001: Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts VI*, 172-183, 2001.

74. E.H. Land, *The Retinex theory of Color Vision.* Scientific American, 1977. **237**: p. 108-129.

75. J.J. McCann and K.L. Houston, *Color Sensation, Color Perception and Methematical Models of Color Vision,* in *Colour Vision,* J.D. Mollon and L.T. Sharpe, Editors. 1983, Academic Press: London. p. 891-894.

76. E.H. Land and J.J. McCann, *Method and System for Image Reproduction Based on Significant Visual Boundaries of Original Object.* 1971: US Patent #3553360.

77. E.H. Land, L. Ferrari, S. Kagan, and J.J. McCann, *Image Reproduction System which Detects Subject by Sensing Intensity Ratios.* 1972: US Patent #3651252.

78. J.J. McCann. *Color Mondrian Experiments Without Adaptation.* in *Proceedings AIC,* 159-162, 1997. Kyoto.

79. W.R. Wray, *Method and apparatus for image processing with field portions.* 1988: US Patent #4750211.

80. J.J. McCann. *Calculated Color Sensations applied to Color Image Reproduction.* in *Proceedings SPIE Vol. 901 Image Processing, Analysis, Measurement and Quality,* 205-214, 1988. Bellingham, WA.

81. H. Wallach, *Brightness constancy and the nature of achromatic colors.* Journal of Experimental Psychology, 1948. **38**: p. 310-324.

82. K. Barnard and B. Funt, *Investigations into multi-scale retinex (MSR),* in *Colour Imaging: Vision and Technology,* L.W. Macdonald and M.R. Luo, Editors. 1999, John Wiley. p. 17-36.

83. D. Marini and A. Rizzi. *Color Appearance Approach to Image Database Visual Retrieval.* in *Proceedings SPIE 3964,* 186-195, 2000. San Jose.

84. J.J. McCann, *Capturing a Black Cat in Shade: The Past and Present of Retinex Color Appearance Models.* Journal of Electronic Imaging, 2004. **13**(1): p. 36-47.

85. J.J. McCann. *The role of nonlinear operations in modeling human color sensations.* in *Procedings SPIE: Human Vision, Visual Processing and Digital Display,* 355-363, 1989. Bellingham, WA.

86. F.W. Campbell and J.G. Robson, *Application of Fourier Analysis to the visibility of Gratings.* Journal of Physiology, 1968. **197**: p. 551-566.

87. J.J. McCann, R.L. Savoy, J. Hall, and J. Scarpetti, *Visibility of Continuous Luminance Gradients.* Vision Research, 1974. **14**: p. 917-927.

88. R.L. Savoy, *Low Spatial Frequencies and Low Number of Cycles at Low Luminances.* Journal of Photographic Science & Engineering, 1978. **22**(2): p. 76-79.

89. M. Rabbani and P.W. Jones, *Digital Image Compression Techniques.* 1991, Bellingham: SPIE Press.

90. J.J. McCann, *Calculating lightnesses in a single plane.* Journal of the Electronic Imaging, 2001(10): p. 112-122.

91. M. Abdulwahab, J.L. Burkhardt, and J.J. McCann, *Method and apparatus for transforming color image data on the basis of an isotropic and uniform colorimetric space.* 1984: US Patent #4839721.

92. J.J. McCann, E.H. Land, and S.M.V. Tatnall, *A Technique for Comparing Human Visual Responses with a Mathematical Model for Lightness.* Amer. J. Of Optometry and Archives of Amer. Academy of Optometry, 1970. **47**: p. 845-855.

93. J.J. McCann and R.L. Savoy. *Measurement of Lightness: Dependence on the position of a white in the field of view.* in *Proceedings SPIE: Human Vision Visual Processing and Digital Display II*, 402-411, 1991. Bellingham, WA.

94. R. Sobol. *Improving the Retinex algorithm for rendering wide dynamic range photographs.* in *Proceedings SPIE: Human Vision and Electronic Imaging VII*, 341-348, 2002. San Jose, CA.

95. J. Cowan. *Visual cortex and the Retinex algorithm.* in *Proceedings SPIE: Human Vision and Electronic Imaging VII*, 278-285, 2002. San Jose, CA.

96. J.J. McCann, *A Spatial Color Gamut Calculation to Optimize Color Appearance*, in *Colour Imaging: Vision and Technology*, L.W. Macdonald and M.R. Luo, Editors. 2002, Wiley & Son Ltd.: Chichester. p. 213-233.

97. S. Westland, P. O Da Pos, and C. Ripamonti. *Conditions for perceptual transparency.* in *Proceedings SPIE: Human Vision and Electronic Imaging*, 315-323, 2002. San Jose, CA.

98. Y. Nayatani, K. Takahama, and H. Sobagaki, *Prediction of Color Appearance of Object Colors in a Complex Visual Field.* Journal of Light and Visual Environment, 1995. **19**(1).

99. R.W.G. Hunt, *Measuring Colour.* 3rd edition ed. 1998: Fountain Press.

100. L.G. Glasser, A.H. McKinney, C.D. Reilly, and P.D. Schnelle, *Cube-root color coordinate system.* Journal of the Optical Society of America A, 1958. **48**: p. 736-740.

101. W.A. Stiehl, J.J. McCann, and R.L. Savoy, *Influence of Intraocular Scattered Light on Lightness-Scaling Experiments.* Journal of the Optical Society of America, 1983. **73**: p. 1143-1148.

102. J.J. McCann and A. Rizzi. *The Spatial Properties of Contrast.* in *Proc. IS&T/SID Eleventh Color Imaging Conference*, 51-58, 2003. Scottsdale, AZ.

103. B. Funt, F. Ciurea, and J.J. McCann. *Tuning Retinex Parameters.* in *Proceedings SPIE: Human Vision and Electronic Imaging VII*, 358-366, 2002. San Jose, CA.

104. *Natural Image Database at the Tominaga Laboratory*, www.osakac.ac.jp/labs/shoji/English/Database/index_e.html.

105. *Vision Texture database at MIT Media Lab*, www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html.

106. D.A. Brainard, *Hyperspectral Image Data*, http://color.psych.upenn.edu/hyperspectral/.

107. K. Barnard, L. Martin, B. Funt, and A. Coath, *A Data Set for Color Research.* Color Research and Application, 2002. **27**(3): p. 148-152.

108. E. Marszalec, B. Martinkauppi, M. Soriano, and M. Pietikäinen, *A physics-based face database for color research.* J. of Electronic Imaging, 2000. **9**(1): p. 32-38.

109. D.L. Ruderman, *Database of hyperspectral images*, ftp://ftp.sloan.salk.edu/pub/ruderman/hyperspectral/.

110. J. Watlington, *Image database*, http://web.media.mit.edu/~wad/color/exp1/new/.

111. F. Xiao, J. DiCarlo, P. Catrysse, and B.A. Wandell. *High Dynamic Range Imaging of Natural Scenes.* in *Proc. IS&T/SID Tenth Color Imaging Conference*, 338-342, 2002. Scottsdale, AZ.

112. A. Rizzi, C. Gatta, and D. Marini. *YACCD: Yet Another Color Constancy Database.* in *Proceedings of the Electronic Imaging*2003. Santa Clara, CA.

113. D.L. Ruderman, T.W. Cronin, and C.C. Chiao, *Statistics of cone responses to natural images: implications for visual coding.* Journal of the Optical Society of America, 1988. **15**: p. 2036-2045.

114. S.M. Nascimento, F. Ferreira, and D.H. Foster, *Statistics of spatial cone-excitation ratios in natural scenes.* Journal of the Optical Society of America A, 2002. **19**(1): p. 1484-1490.