Investigations into Colour Constancy by Bridging Human and Computer Colour Vision

A thesis submitted for the degree of

Doctor of Philosophy

Stuart Owen John Crichton MEng

School of Electrical and Electronic Engineering

Newcastle University

October 2014

Abstract

The mechanism of colour constancy within the human visual system has long been of great interest to researchers within the psychophysical and image processing communities. With the maturation of colour imaging techniques for both scientific and artistic applications the importance of colour capture accuracy has consistently increased. Colour offers a great deal more information for the viewer than grayscale imagery, ranging from object detection to food ripeness and health estimation amongst many others.

However these tasks rely upon the colour constancy process in order to discount scene illumination to allow these tasks to be carried out. Psychophysical studies have attempted to uncover the inner workings of this mechanism, which would allow it to be reproduced algorithmically. This would allow the development of devices which can eventually capture and perceive colour in the same manner as a human viewer.

These two communities have approached this challenge from opposite ends, and as such very different and largely unconnected approaches. This thesis investigates the development of studies and algorithms which bridge the two communities. Utilising findings from psychophysical studies as inspiration to firstly improve an existing image enhancement algorithm. Results are then compared to state of the art methods. Then, using further knowledge, and inspiration, of the human visual system to develop a novel colour constancy approach. This approach attempts to mimic and replicate the mechanism of colour constancy by investigating the use of a physiological colour space and specific scene contents to estimate illumination. Performance of the colour constancy mechanism within the visual system is then also investigated. The performance of the mechanism across different scenes and commonly and uncommonly encountered illuminations is tested.

The importance of being able to bridge these two communities, with a successful colour constancy method, is then further illustrated with a case study investigating the human visual perception of the agricultural produce of tomatoes.

“Cross the streams…”

Ray Stantz

Acknowledgement

I wish to express my gratitude and thanks towards both of my supervisors; Professor Anya Hurlbert and Professor Gui-Yun Tian for giving me the opportunity to pursue my PhD studies. This thanks also extends towards their countless encouragement, advice, and help over the years leading to this point. I also wish to recognise and state thanks for an EPSRC DTA from the School of Electrical and Electronic Engineering which funded this work, alongside funding from the Institute of Neuroscience.

I would also like to give thanks to Dr Michal Mackiewicz, Dr Angela Owen and Matthew Cranwell for their help, training and guidance on different aspects of my PhD. Furthermore, I would like to thank Dr Martin Johnston for his friendship and guidance over this period, without it the journey would not have been as enjoyable. My thanks also go to countless others I have met on a similar path.

To Saya goes my thanks and eternal appreciation for your love, support and unique humour. These have both kept me going, and kept things in perspective throughout these four years.

Finally I wish to thank my family, for their support and efforts throughout everything up to this point in my life; for spurring me on to succeed and to take paths to destinations unknown. As such this thesis is another step on the journey forwards.

Table of Contents

List of Tables

List of Figures

Abbreviations

C.I.E. – Commission Intenationale de l'Eclairage/ International Commission on Illumination

C.C.T. – Correlated Colour Temperature

C.D.F. – Cumulative Distribution Function

C.F.A. – Colour Filter Array

G.S. – Greyscale

H.E. – Histogram Equalisation

H.S.C. – Hyperspectral data cube

H.S.I. – Hyperspectral Image

H.V.S. – Human Visual System

J.N.D. – Just Noticeable Difference

P.D.F. – Probability Density Function

RGB – Red, Green and Blue colour space

Chapter 1. Introduction

The inability to fully replicate the phenomenon of colour constancy exhibited by the human visual system is a major hurdle to the use of computer vision in uncontrolled illumination scenarios, in tasks ranging from object recognition to consumer photography. Colour constancy is the ability of the human visual system to keep object colours roughly constant across the constant changes in illumination spectrum experienced on a day-to-day basis. Previous methods to solve the challenge posed by colour constancy have been focused on stringent statistical assumptions regarding images and their contents. These statistical assumptions have generally had no connection to findings from psychophysical research into colour constancy or colour vision, aside from the manner in which they scale each colour channel. Colour information is also encoded and processed differently in computer vision compared to the human visual system.

Whilst tri-chromatic vision may have inspired sensor design, the colour capture capabilities of the majority of current sensors cannot match that of the human visual system. Other capabilities of computer vision systems, connected to dynamic range capabilities and most importantly visual perception also differ greatly to other capabilities of the human visual system. Recent advances in imaging technology has led to the development of hyperspectral imaging devices, which have become faster and more affordable as time has progressed. The extra spectral information alongside the superior dynamic range capabilities of these devices allows the capture of ground truth chromatic information in comparison to standard 3 channel imaging devices. In short, the colour information captured by devices with a small number of discrete-spectrum channels are under sampling the visual spectrum in comparison to what is achieved by the retina. This most likely compounds the challenges faced by computer vision systems when attempting to replicate the performance of the visual system.

Hyperspectral imaging means that we can capture spectral information at a greater level than the human visual system, allowing us to match the information the visual system utilises and to design methods in more perceptually and physiologically correct colour spaces. In short it captures the spectral information within a scene in a continuous manner, unlike the discrete channel defined manner of RGB imaging devices. This

allows the capture of data at a higher level than the actual human visual system, allowing us to match the information utilised by the visual system.

1.1 Motivation

The vast majority of historic image processing research has tended to ignore processes within the human visual system which carry out similar tasks to the challenges posed by natural scenes to image capture and processing systems. These methods have tended towards finding novel methods to meet these challenges. If methods have utilised aspects of human colour and visual perception into algorithm investigation and design they have tended to be based upon data captured at a lower colour and luminance resolution that the human visual system utilises.

This thesis aims to show the development, and limit, of a perceptually based algorithm upon RGB images, and how hyperspectral imaging and the higher level of colour information captured can aid the development of more physiologically and perceptually based algorithms for image processing, including colour constancy algorithms and colour based food analysis.

1.2 Aims and objectives

The main aim of this thesis is the development of image processing algorithms based upon perceptual and/ or physiological features in the human visual system in natural images alongside the development of studies into uncovering the mechanisms that underlie colour constancy within the human visual system.

We first develop a contrast image enhancement technique for colour images captured by standard RGB imaging devices to show how this can be carried out with standard images. We then move on to show the development of a spectroradiometrically calibrated hyperspectral dataset of human skin under a large number of illuminations, and the use of this database to create a physiologically-inspired contents-based colour constancy algorithm. The aim of this study is to determine whether the information available in the colour signal from a polychromatic surface under changes in illumination is sufficient to drive a colour constancy algorithm. We develop such an algorithm based on skin chromaticity gamuts. It could then be inferred that changes in colour of known objects could be used as cues to estimate illumination within the human visual system.

We then show the results of a psychophysical study into the mechanisms of colour constancy within the human visual system itself. This study is novel in two ways, the first being the manner in which it utilises a reflective display rather than an emissive one, meaning that the display behaves more similarly to a natural chromatic surface under illumination. The second was the choice of illuminations that subjects were tested under, with a range of daylight and non-daylight illuminations. The first aim was to determine whether reflective displays could be utilised to test the mechanisms of colour constancy, and the second to test the performance of human colour constancy under illuminations which are commonly and uncommonly encountered. The latter aim is analogous to the use of known/ commonly encountered scene contents as a cue for illumination estimation in the previous section.

With the utilisation of hyperspectral imaging in the development of the skin chromaticity gamut algorithm, we then move on to show another use of this imaging technique to bridge human and computer visual perception applications. Here we aim to compare the spectral reflectance changes in tomatoes stored under different conditions to the change in human visual and other sensory perception of the same tomatoes. The aim of this study is to find key wavelength markers which can help to predict the human visual and other sensory responses to tomatoes of varying material properties, and also help to differentiate key wavelengths related to storage length.

1.3 Statement of originality

This thesis merges the two disciplines of human and computer vision studies and utilises findings in both to improve existing, and develop novel algorithms. It shows the applicability of this approach by utilising findings related to the perception of surface gloss to improve existing contrast enhancement algorithms. The thesis then moves on to show the novel development of a hyperspectral skin database under a wide array of illuminations to develop a novel polychromatic image-content based colour constancy algorithm with a good level of performance utilising a bespoke LED lighting system. This aims to test whether commonly encountered and known polychromatic objects can be used as the basis for illumination estimation within the human visual system. We then test the same hypothesis with regards to illumination, rather than scene contents, by testing for colour constancy within the human visual system incorporating the first use of a reflective display within this field of study. Finally we move on to test whether hyperspectral imaging techniques, within the visual wavelength range, can be utilised to

estimate the consumer's visual and other sensory perception of tomatoes stored under different conditions.

1.4 Publications arising from this thesis

Each chapter within this thesis has led to publications, either published or in preparation:

[1] S. Crichton, A. Hurlbert, and G. Y. Tian, "Preserving scene texture perception in colour image enhancement: Luminance-based monotonic normalisation," in *AIC 2013*, Newcastle-upon-Tyne, 2013, pp. 1069-1072.

[2] S. Crichton, J. Pichat, M. Mackiewicz, G.-Y. Tian, and A. Hurlbert, "Skin chromaticity gamuts for illumination recovery," *Conference on Colour in Graphics, Imaging, and Vision*, vol. 2012, pp. 266-271, 2012.

[3] M. Mackiewicz, S. Crichton, S. Newsome, R. Gazerro, G. D. Finlayson, and A. Hurlbert, "Spectrally tunable LED illuminator for vision research," *Conference on Colour in Graphics, Imaging, and Vision*, vol. 2012, pp. 372-377, 2012.

[4] S. Crichton, B. Pearce, M. Mackiewicz, G. Finlayson, and A. Hurlbert, "The illumination correction bias of the human visual system," *Journal of Vision*, vol. 12, pp. 64-64, 2012.

[5] M. Mackiewicz, B. Pearce, S. Crichton, G. Finlayson, and A. Hurlbert, "Achromatic adjustment outdoors using MEMS reflective display," in *Perception*, 2012, pp. 1522-1522.

[6] M. Mackiewicz, B. Pearce, S. Crichton, A. Hurlbert, and G. Finlayson, "Achromatic adjustment outdoors and indoors using the Mirasol reflective display," in *AIC 2013*, Newcastle-upon-Tyne, 2013, pp. 295-298.

1.5 Thesis organisation

This thesis is comprised of six further chapters. A literature review into the previous research into colour constancy in both the psychophysical and computer vision research fields is firstly presented. It explains the derivation of current state of the art methods from previous hypotheses regarding the mechanisms of colour constancy within the human visual system. It then goes on to discuss more recent advances in

understanding these mechanisms and how these can be translated into newer algorithmic approaches.

Using the knowledge gained from the literature review the luminance skewness transform is then introduced. Its derivation from psychophysical research regarding human visual texture perception and the performance of the existing monotonic transform under certain image conditions is explained alongside its basis in previous methods. Its performance on under and over exposed images is then visually compared across both colour and grayscale images to a wide number of existing methods. This chapter illustrates how the application of psychophysical research findings can aid the performance of image processing algorithms, even those which operate within a non-perceptual or physiological colour space.

The next chapter moves forward to introduce the concept of the application of hyperspectral imaging, in conjunction with a large number of illuminations, for the creation of an image dataset of human skin. It then utilises this data to investigate the validity of a physiologically inspired colour constancy algorithm based upon the use of a known surface, human skin, as a cue for illumination estimation. This chapter bridges certain aspects of human visual perception -- with the use of cone-contrast colour space and known polychromatic object colour signals-- with current computer vision approaches. The chapter introduces two novel approaches for the use of human skin as a cue to illumination estimation, and compares their performance.

We then investigate colour constancy within the human visual system, attempting to determine whether 1) illumination, rather than object, familiarity aids colour constancy, and 2) reflective display technologies can be utilised to accurately investigate colour constancy. This chapter then shows how the performance of colour constancy, in terms of achromatic adjustment, varies across different illuminations.

To achieve the aim of bridging human and computer colour vision in an application scenario the next chapter then builds upon the application of hyperspectral imaging to aid the development of algorithms and the psychophysical measurements of the perceptual mechanism of colour constancy. It does this through the combination of both with regards to measuring the visual and other sensory perception of tomatoes stored under different conditions by human subjects and relating these behavioural measures to the colour and spectral information gained through hyperspectral imaging. We then

show how hyperspectral imaging can be used as the basis for food processing algorithms with regards to consumer perception.

The achievements of this thesis, with regards to the aims and objectives, along with its contribution to knowledge are then summarised.

Chapter 2. Literature Review

2.1 Introduction

Colour constancy has been defined as the ability of the human visual system to keep perceived object colours roughly constant across changes in illumination. Even after a great deal of research the question of whether colour constancy actually exists has still been asked [1, 2], given the differences between the day-to-day real world experience of constancy and empirical measurements of the phenomenon. The latter, from a number of different methodologies, conclude that although colour constancy exists, it is imperfect and highly variable depending on the experimental conditions and method of measurement, whereas everyday experience suggests a more robust phenomenon. Nonetheless, the question of how colour constancy operates, and to what degree other mechanisms such as memory and adaptation contribute to it, have still not been answered. This lack of understanding has then had a knock-on effect upon the development of computer vision systems.

The challenge of colour constancy within the imaging community was first seriously tackled by Land [3], with the Retinex theory. This attempted to bridge the gap between psychophysical attempts to uncover the mechanisms behind colour constancy and the potential for these findings to be implemented into image processing algorithms. Since this time however attempts to fully bridge these two regions of research have reduced in number. Psychophysical research has moved onward with new hypotheses, testing paradigms and mechanisms, with a constantly altering view of colour constancy. Research in the computer vision realm has moved in a different manner. Within computer vision research, colour constancy has the same basic definition; that object colours should be invariant to illumination changes. Historically however, this definition is where similarities between the two fields have ceased, due to the differences between their two capture and processing systems which make the definition of colour different between the two fields. Furthermore, in computational vision, colour constancy is often equated with spectral reflectance recovery, which is not the same as human colour constancy unless one assumes that perceived colour equals spectral reflectance, which it does not. Computer vision work has attempted to achieve colour constancy using a much more limited amount of low level information and simplistic processing capability These differences are linked to both the differing capabilities of the two types of ‘imaging devices’, alongside a degree of failure with regards to the dissemination of findings from knowledge regarding human vision to the computer

vision field. Initially [3] attempted to bridge both of these fields, however subsequent work in both fields has proceeded separately. To put it simply, colour constancy is a mechanism within the visual system, and as such if we hope to replicate it and match its performance, as such should we not attempt to use the findings of psychophysical experiments as a basis for perceptually accurate processing algorithms?

To do so, this chapter will firstly lay out the research carried out within the fields of psychophysics and computer vision. This will showcase the state of the art approaches that have been utilised to both attempt to understand, and provide solutions to the challenge of colour constancy within both fields. The importance and proven performance gains of utilising findings from psychophysical experiments in understanding human visual perception within perceptually based image processing algorithms will then be shown.

2.2 Psychophysical research into colour constancy

2.2.1 *The human perception of colour*

Colour cannot be quantified as a definitive measure, it is a perceived property of an object whose perception is heavily dependent upon the wavelength sensitivities of the viewer, the incident illumination and higher level colour processing.

As such the questions of what colour is, and how we perceive it in the way we do has fascinated people for millennia [4]. The spectral irradiance from a given point in the scene reaches a point within the retina, at which the three cone types translate this spectrum into a tri-stimulus representation, via the three broad, overlapping spectral sensitivity functions.

The way in which these L, M and S signals are utilised by the visual system is a question of debate however. Two systems, with regards to the processing mechanisms, the trichromatic and colour opponent colour models [5] are now known to be correct.

The idea of the tri-chromatic processing model was first fully proposed by von Kries [5], and is based on the fundamental theory that the LMS sensitivities are separate. Whilst this has been proven over time not to be fully accurate, the mechanism of trichromacy still persists and has been the basis for imaging sensor and processing algorithm design as will be discussed later within this chapter.

Within the 19th century Hering proposed the original opponent colour theory of human vision [5]. It is now known that the second stage of colour encoding the outputs of the

three cone types in the retina are combined in an opponent fashion much as Hering proposed. The cone signals are combined such that the opponent colour model has three axes: achromatic (lightness), red-green and blue-yellow. To calculate these from L, M and S responses:

$$Achromatic = L + M + S \quad (1)$$

$$RG = L - (M + S) \quad (2)$$

$$BY = L + (M - S) \quad (3)$$

Human colour perception and the mechanism of colour constancy are also not simply functions of the cone sensitivities and colour model. If the cone sensitivities and connected colour model were the sole factors upon colour constancy scaling them akin to the chromatic adaptation principle would always result in colour constancy. This would mean that colour constancy should perform well under any illumination.

2.2.2 *Psychophysical studies*

The generally unnoticed performance of colour constancy mechanisms across unconstrained scenarios in the real world has meant that testing for the individual influence of different scene characteristics is difficult. Before we can analyse previous testing methodologies we must first delve into the scenarios in which colour constancy operates subconsciously on a day-to-day basis. Consider that walking from a room lit by a standard incandescent bulb into another room illuminated by a fluorescent tube, both at differing colour temperatures: as we travel through these two rooms we do not notice any change in object colours even with the large change in illumination spectrum. This is the brief premise of colour constancy. If we change the location to an outdoor scene, with changes in illumination spectrum brought about by variable cloud cover we notice it still works well. In fact colour constancy operates without any general hint of existence to us, regardless of scene. What this means is that the underlying mechanisms can operate even across the vast number of variables within a scene. If we take the room scene as an example and consider a single object, say a green cup, the most basic of possible variables is that of the incident illumination incident. However a number of other variables also exist. The location of the illumination source may change, as well as the location of the human viewer as we travel through the room. The geometry of other objects in the room may also change and introduce changes in mutual illumination on the cup, changing the local colour surround to the cup and further to this the

chromatic information of the cup may also change from being monochromatic to polychromatic. Somehow the human visual system can account for all of these variables and keep the perceived colour of objects constant. However many other aspects of perception may or may not have in input into the mechanism of colour constancy, aspects such as the familiarity, or knowledge of an object amongst others [6, 7]. Psychophysical research has been undertaken into all of these factors in order to uncover their effect, if any.

Before reviewing investigations into specific scene variables we should give an overview of different testing methodologies. Ehrenstein [8] gives an overview of the main methods of psychophysical investigations. The purpose of this thesis is not to review different testing methods, but to review how they have been utilised in different individual investigations.

The first of these methodologies utilised in a number of works is that of the achromatic adjustment, where the subject is asked to adjust the colour of a patch to a neutral achromatic colour.

Foster [2] discusses the different testing methods in relation to what they actually test within the human visual system. Certain testing methods such as achromatic adjustment really test the ability of the subject to estimate the illumination chromaticity, whilst others such as surface matching test the ability of the visual system to perform a relative form of colour constancy [2], rather than absolute.

It can also be argued that all of these methods attempt to assess colour constancy to a level at which it may or may not need to operate [2]. Whilst with psychophysical testing connected to colour constancy we must be mindful of not only the manner of testing but also the environment in which stimuli are presented. How stimuli are presented, both in terms of objects and spatial and time variation, is the avenue which allows us to test the effect of different stimuli upon colour constancy performance.

This plethora of investigation methodologies is extremely useful for attempting to reproduce consistent results for a given experimental hypothesis. However due to their generally tightly controlled stimuli they cannot be expected to retrieve similar colour constancy results to the human visual system in real world situations. This variation in results has been noted in both [1, 2], and can be thought to relate the vast differences between a real world scene and a typical test scene within a psychophysical experiment and also differences in colour constancy performance in different scenarios. From this

style of testing it can be inferred that colour constancy is commonly considered by researchers as a singular process within the human visual system, such that no matter what test conditions are, in terms of scene cues, we expect that in a real world scenario, consisting of those cues amongst others, that performance will be matched.

Real world scenes generally contain a number of different cues which have been shown to aid colour constancy for a human observer. As Foster [2] stated these range from the idea of the combination of these cues aiding colour constancy in a human observer has been noted previously [9, 10]. These scene cues, under a single illumination, range from specular highlights [11], scene statistics [12], mutual illumination [13], 3D object cues [14] and polychromatic objects [15, 16] amongst others.

These scene cues are also supplemented by what Maloney et al. [9] call priors, such as prior colour knowledge of objects or colour memory [17, 18]. Standard stimuli for colour constancy, such as Mondrians or novel objects, traditionally use plain surfaces or shapes for adjustment or viewing to try to bypass the bias that may be introduced by object familiarity.

Further to methods which attempt to uncover human colour constancy performance to infer potential mechanisms behind the process, other methods have tried to uncover the same mechanisms by seeing where colour constancy fails [11].

From the above studies it can be noted that variables, including the incident illumination, scene contents and varied stimuli, have been investigated for their effect upon colour constancy performance.

Different testing scene scenarios have also been utilised in order to control possible environmental cues to illumination in the scene. This is specifically utilised to investigate the individual effectiveness of different cues to the human visual system. A number have proposed that, and attempted to investigate if overall scene contents or statistics have an effect upon colour constancy performance, often with varying degree of success. These investigations are connected to investigations into whether chromatic adaptation can fully explain colour constancy performance across all scenes [19].

These have ranged from the grey world theory (assumption of a spatial average cue) [20], to the white patch theory (maximum reflectance patch directly connected to illumination in each scene) [3] and local surround effects [19, 21]. Investigations into global scene cues have included grey scenes, Mondrian scenes and black scenes [22-25]

amongst others. These global scenes have also been utilised to investigate their effect upon performance whilst also incorporating other cues in the scene, such as specular highlights [11, 26].

Generally colour constancy experiments, as noted in [19], utilise stimuli which are simple flat, matte and uniformly coloured surfaces. These are gross simplifications of surfaces in the real world, but this does mean that cues such as specular highlights, shading and other 3D cues to illumination and objects are removed.

These methods assume colour constancy is performed through the adaptation of cone values relative to a single reference, whether it that is the global average or local average surround. However Foster et al. [12, 27] put forward the idea of what they call relational colour constancy, which noted that across changes in illumination in real world scenes the relationships between the colours of different objects in the field of view are generally invariant. Foster et al. undertook investigations to measure cone ratios in typical scenes (constructed from natural surfaces) and found that these cone ratios are largely invariant across Planckian/ daylight illuminations. Foster et al. [27] further build upon the basic theory laid out within [12] by discussing issues that surround colour constancy and relational colour constancy and the way these phenomena are investigated. One point of importance is the combined effect of subject instructions and scene choice on how subjects will approach a task. The same instruction set and the use of different stimuli, such as paper or an adjustable patch, will result in the subjects using different methodologies during the experiment.

Whilst above we have seen that numerous investigations into human colour constancy have been undertaken within the psychophysical field, colour constancy is also a major field of interest within the computer vision research field.

2.2.3 *Illumination specific investigations*

Whilst the previously mentioned studies has specifically targeted investigations towards scene contents and potential cues, investigations into performance across different illuminations in constant scenes have not been as prevalent. The testing of colour constancy performance under different illuminations is directly linked to the process of colour constancy, and as such is the most logical variable to test. However due to the technological availability of illumination sources and spectra initial investigations studied small numbers of very chromatically distinct illuminations [3] through the combined use of lights and filters. Initial investigations focused upon the chromaticities

of illuminations due to the broadband spectral nature of light sources. Illumination comparison methods such as the CRI and CQI [28] have been developed to quantitatively illustrate the gamut of colours that can be reproduced under a given light source. However with the advent of multiple channel solid state lighting, and its affordability [29], an increase in illumination metamerism has been seen across illuminations with very different spectra and CRI values. Metamerism here refers to the ability of multiple channel solid state light fixtures to output chromatic metamers with vastly different spectra. This has meant that although the range of possible illuminations and also spectral accuracy to standard CIE illuminations for testing have increased from a single light unit, investigators need to be more careful with respect to illumination choices by paying attention to illumination spectra and CRI values than has previously been done on an illumination by illumination basis.

As Foster et al. observe [12, 27], these cone ratios may be geared towards providing a greater invariance across daylight illuminations. Logically if we consider colour constancy to be a product of evolution we would expect the mechanism to perform better under more commonly encountered illuminations.

A potential bias of human colour constancy towards the Planckian locus or at least certain regions of the Planckian locus [25, 30] has been investigated by numerous studies. Some have shown a bias towards the daylight in terms of performance [25], whilst others have not shown any overall relationship [30]. Whilst they may disagree in terms of their overall findings, some similarities in results are seen across both studies. Delahunt [30] uncovered better performance in different directions from the location of CIE illumination D65 on the Planckian locus. Delahunt et al. found that constancy was best in the blue direction from this point, whilst constancy in the green direction was worst. These results, specifically the greater level of constancy in the green direction than the yellow, point towards there being no overall bias within the visual system for the daylight locus. This performance deficit was tentatively linked to the idea that human colour vision has evolved for use within forested green scenes.. The spectral sensitivities of the cones within the retina do not necessarily lend themselves towards being the most efficient for the separation of daylight spectral signals if we utilise the principle von Kries chromatic adaptation [31]. Pearce et al. [25] uncovered a similar better performance of colour constancy under blue and green illuminations in comparison to red and yellow. The similar finding was found utilising a different test method, a two alternate forced choice based method in a viewing box, in comparison to

an achromatic match using a stereo display. Illumination choices were similar in that they were centred around a central point on the daylight locus. Whilst the central point of [25] was D67, with [30] this was D65. Gaining similar results across these two methods is a good indicator that this performance differential across illuminations does indeed exist. The results pointing towards these findings form the basis of many of the chapters within this thesis.

2.3 Colour constancy in computer vision

2.3.1 *Colour constancy and colour normalisation*

Whereas colour constancy is well defined within the domain of human vision science, its definition is slightly more diffuse within the computer vision field. For computer vision systems the difference in definition relates to the variation of the resulting algorithms in comparison to those produced by the human visual system. The methods can be grouped into the following;

- Colour constancy methods – attempting to replicate the performance of the human visual system across illuminations, recovering colours that correspond to human perception under varying illuminations.
- Colour invariance methods – To record colours invariant to illumination, which do not necessarily correspond to colours perceived by humans.

Colour constancy methods, or those wishing to result in perceptually accurate colour [32], are generally used in photographic applications.

2.3.2 *The challenge of colour constancy: Image complexity*

Photographs are simplified 2D representations of complex 3D scenes. Scene complexity relates not only to object contents, scene geometric information and illuminations which can all change wildly both across images, but also within them. Photographs can be considered to be a reduced dimension representation of a real world scene in terms of colour, depth and brightness information.

This scene complexity is greater than that experienced by the human visual system which has a number of other tools such as previous object and colour knowledge to hand.

Figure 1 - The complex 3D nature of a scene

Let us consider the 3D nature of a scene and how illumination can vary across this using Figure 1 and Figure 2. We can see that objects which are neighbours in pixel terms can be separated by a large physical distance in the scene. This can result in the surfaces corresponding to these pixels being under different illuminations. Even objects located within the same depth plane can be under two different illuminations, or a mixture of multiple illuminations, as shown in Figure 2.

Figure 2 - Illumination sources

On top of these complex scenes it can be accepted that comparing human colour constancy, which can possibly utilise the constant stream of image information from the retina, to attempts to replicate it using the information present within a single image, is another challenge. If we think of colour constancy as a continuous process of which a human observer is generally unaware, any video-based method that replicates it must not result in rapid illumination estimation changes.

2.3.3 *The challenge of colour constancy: The Challenge of Illumination*

As previously mentioned the possible range of illuminations under which objects can exist is constantly increasing with the advent of solid state lighting. However if we only consider outdoor daylight illuminations the range of illuminations and changes in their spectra are well known and documented [33]. The potential colour constancy bias towards these illuminations has previously been mentioned. Metamerism and the effect upon human colour constancy, with respect to light sources, has also been previously noted. It follows that these sources also introduce further challenges for any colour constancy algorithm.

2.4 Image capture technologies

Prior to discussion of colour constancy techniques current image capturing device technology must be discussed. The vast majority of image devices capture pixel colour information through the use of 3 channels, but research into the usefulness of adding channels has recently become of great interest.

2.4.1 *3 Channel RGB devices*

Each 3 channel RGB device has its own set of spectral sensitivity functions, this, as stated previously, determines the irradiance spectra that can be recorded by the device and are determined by the colour filter array (CFA) placed on top of the imaging sensor. The form of the sensitivity functions also determine the accuracy of the sensor to the von Kries model of chromatic adaptation, which led to the development of spectral sharpening methods [34, 35]. These sensitivity functions can also differ from the LMS sensitivities of the three cone types within the human retina, and so can be expected to have an effect upon the comparable performance between the two systems. What this means in reality is that certain stimuli will invoke different responses at the sensor/ retina level of the two systems; therefore, we must ask whether we may expect to be able to develop a comparable colour constancy process when the systems differ in such a manner?

2.4.2 *A question of increasing the channels*

This difference in sensitivities has long been known and discussed, and has led to the question of how many unique channels, and the spectral characteristics of these channels , are enough to be able to reproduce/ capture the entire observable range of colours in both imaging and printing [36]. The characteristics of these channels are important for both colour capture and reproduction.

The importance of this for colour constancy within colour reproduction in printing is also another important research area [37]. The fact that most imaging sensors, excluding for example Fovean sensors [38] or 3 sensor systems, are given their spectral sensitivities through the use of a CFA means that the spectral sensitivities can be altered, or extra channels added through the use of a new filter. However another method of image capture is starting to become more commonplace: hyperspectral imaging.

2.4.3 *Hyperspectral Imaging*

Whilst sensors with distinct spectral sensitivities convert the irradiance at a pixel into a discrete number of broadband channel intensities, hyperspectral imaging records the irradiance spectrum at the pixel at close to a wavelength by wavelength level. Both the spatial and spectral resolution of the camera can be altered at time of capture using camera software control. Below is an example of a real world image and data captured using our calibrated Specim V10E camera [39] are below:

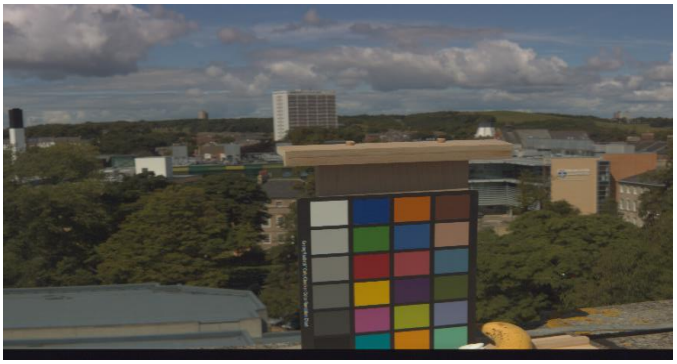
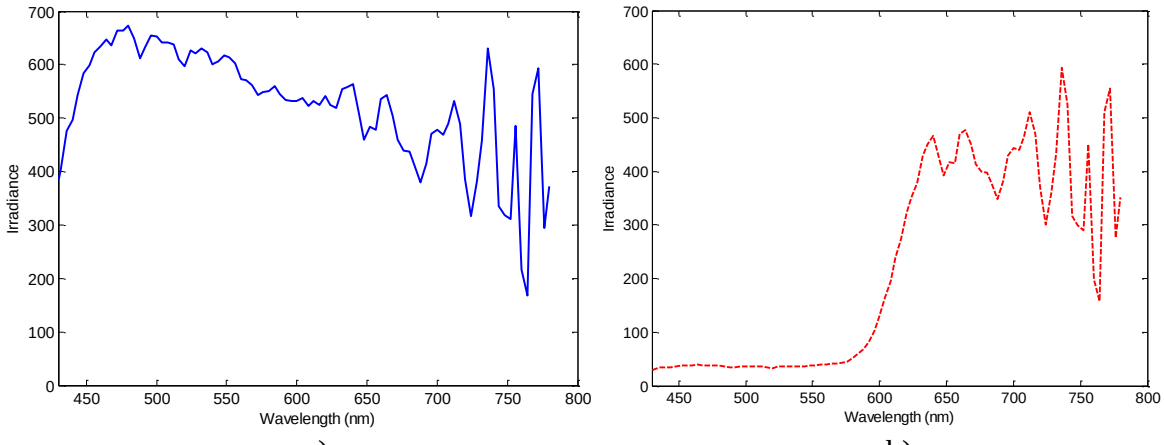


Figure 3 - An example of an outdoor hyperspectral image in sRGB space



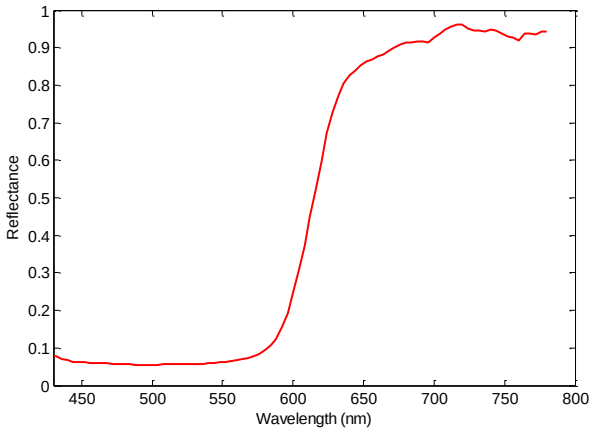


Figure 4 - Example spectral data from the HSC; a) Illumination spectrum from white patch, b) Irradiance spectrum from red patch and c) Recovered reflectance spectrum of red patch.

This form of imaging allows the capture of ground truth spectral information for a scene, and also allows for the removal and application of various illumination spectra to allow in depth colour investigations with high levels of control.

2.5 Colour spaces

Colour space choice in algorithms has generally been dictated by the colour space of the original captured device. These machine colour spaces are very different to the colour models for human observers. The perceivable range colours for human subjects has been investigated to great extent by a number of different studies [5]. However initial computer vision methods into colour imaging and image processing tended to utilise the basic machine colour spaces which had been determined by the sensitivities of imaging devices. However over time colour space research has attempted to bridge human colour vision and computer vision space models. As such a number of colour spaces have been devised, which can be grouped into a much smaller number of space types, those of machine colour spaces, perceptually based spaces and

physiologically based colour spaces. Perceptually based colour spaces can then be further split into two types: those of perceptual uniformity and non-perceptual uniformity.

Perceptual colour spaces were derived from investigations into the cone sensitivities of the human retina. Colour spaces such as the 1931 CIE Yxy space were derived from the tristimulus XYZ observer colour matching functions [5];

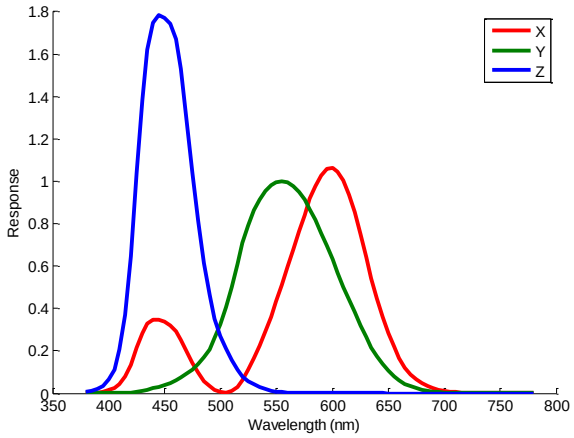
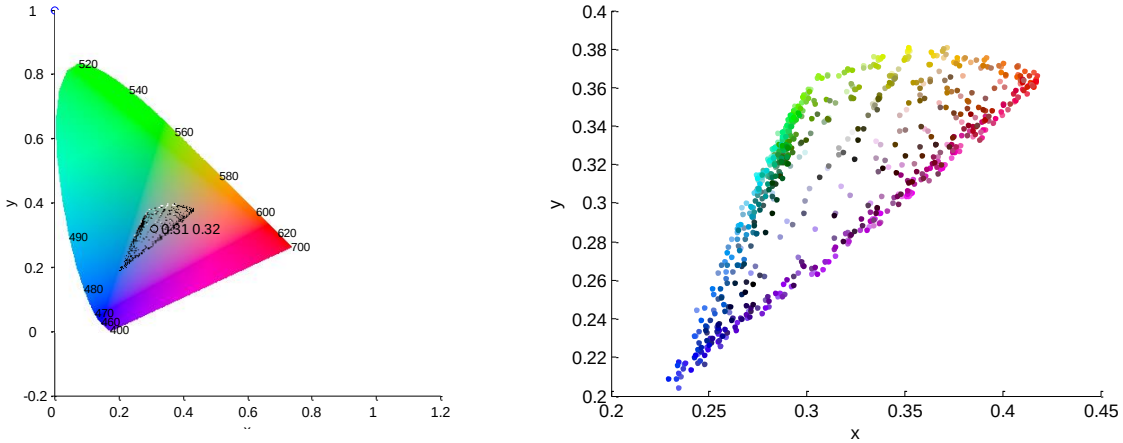


Figure 5 - XYZ observer colour matching functions

$$x = \frac{X}{X+Y+Z} \tag{4}$$

$$y = \frac{Y}{X+Y+Z} \tag{5}$$

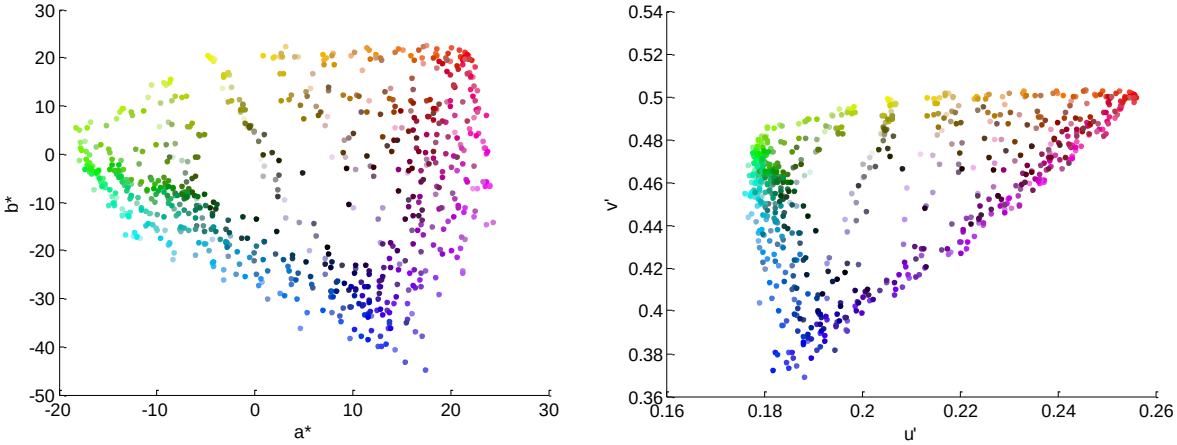
$$Y = Y \tag{6}$$



a) b)

Figure 6 - 1931 CIE xy space showing a) illumination D65 and the display gamut, of the Mirasol reflective display under D67, and b) a larger version of the gamut (collapsed across the luminance plane).

Whilst the 1931 CIE xy diagram illustrates the gamut of colours that are discriminable by a human observer, the distances between unique chromaticities in the diagram do not relate to the perceivable colour change between two chromatic points. In order to account for the non-perceptual nature of xy space, in terms of perceptual uniformity across distances, further spaces were developed; La*b* and Lu'v'.



a) b)

Figure 7 – The same Mirasol display gamut in a) CIELAB, and b) CIELU'V' spaces, both collapsed across the luminance plane.

Within La*b* space, Euclidean distances directly relate to perceptual differences between chromatic points. The space also utilises a von Kries transform, with a Euclidean distance, denoted as ΔE, of 1 equating to a JND of 1. However this perceptual uniformity is based upon the assumption of a D65 whitepoint. As such under other illuminations this perceptual uniformity does not exist and so care needs to be taken when utilising this colour space. The luminance value is a percentage representation that compares the luminance of the colour to that of the chosen white

point. L values are within the 0-100 value range apart from for specular highlights, which exhibit an L value greater than 100. Luv space has two different varieties; that of Lu*v*, like La*b* space, can be considered to be perceptually uniform with the assumption of a single illumination. It utilises a Judd type colour transform.

In the human visual system, the second stage of colour encoding involves colour-opponency, originally formulated by Hering [5]. This encoding creates a different representation for colours using the LMS cone responses to create the axes of variation in luminance, red-green and blue-yellow. Colour opponent space has also been described in detail within [40]. Brainard has explained the use of two versions of the colour opponent space. With [40] defining the one we have utilised. As we are using the colour opponent space the choice of adaptation point strictly controls the chromaticity of the origin.

From the dawn of analogue colour imaging the workflow of using three primaries to form a colour image was inspired by knowledge regarding the three sensitivities of the human eye and the additive colour model. As such imaging sensors have been designed to follow this workflow, and has resulted in basic machine colour spaces also having 3 primaries. The group of machine colour spaces includes that of RGB for images and monitor displays, along with its derivatives, and RGB or CMYK for printing devices, among others. These derivatives include rg chromaticity space as described below;

$$\boldsymbol{r} = \frac{\boldsymbol{R}}{(\boldsymbol{R}+\boldsymbol{G}+\boldsymbol{B})} \tag{7}$$

$$\boldsymbol{g} = \frac{\boldsymbol{G}}{(\boldsymbol{R}+\boldsymbol{G}+\boldsymbol{B})} \tag{8}$$

Along with reduced dimension spaces such as RB [41, 42] and colour ratio spaces such as R/G [43]. Other colour spaces have been devised from RGB sources, for use within TV transmission for example, such as the YCbCr, YIQ, YUV and YES spaces. These latter spaces split the luma (RGB derived luminance) and chroma information, with the chroma information split into opponents. However all of these machine colour spaces have no relation to how humans perceive colour, both in terms of colour gamut and perceptual uniformity across changes in colour values. This means any transform in these spaces does not have a clear and easily understandable translation to the human perception of the same change. They exist due to the chosen dimensions connected to the imaging and or printing device sensitivities. This is noted by Kakumanu et al. [43] its separate definitions of different colour spaces. Such colour spaces can be either

device dependent or device independent. Device dependency relates to whether a colour triplet, if we assume RGB, recorded or shown by a colour device for a chosen stimulus is the same across different devices. If the display or capture primaries for devices are different then the same RGB triplets will have different perceived chromaticities. The gamut of colours, either to be displayed or captured, within these devices will differ. To counter this sRGB space was developed with specified primary values and whitepoint [44].

Whilst the overall gamut of colours is set by the choice of primaries and backlighting the number of unique colours that can be captured or displayed within it is determined by the choice of colour depth. Currently the most common colour depth is 8 bits per channel, with a total of 16.78 million colours. At present however displays and graphics devices with colour depths of 10 bits per channel and greater are beginning to come to the consumer market, enabling a much higher colour resolution within the colour gamut. Whilst the sRGB standard can work with these enhanced colour depths, this has led to the development of newer colour spaces such as scRGB [45] and xvYCC [46].

2.5.2 *Colour image formation on a 3 channel sensor*

Previous sections have discussed sensor and processing colour space, yet the capturing itself of colour is also of importance. The recorded intensity for a given pixel is from the combination of the incident illumination (E_λ), object surface reflectance (S_λ) and the spectral response of the imaging sensor (Q_λ);

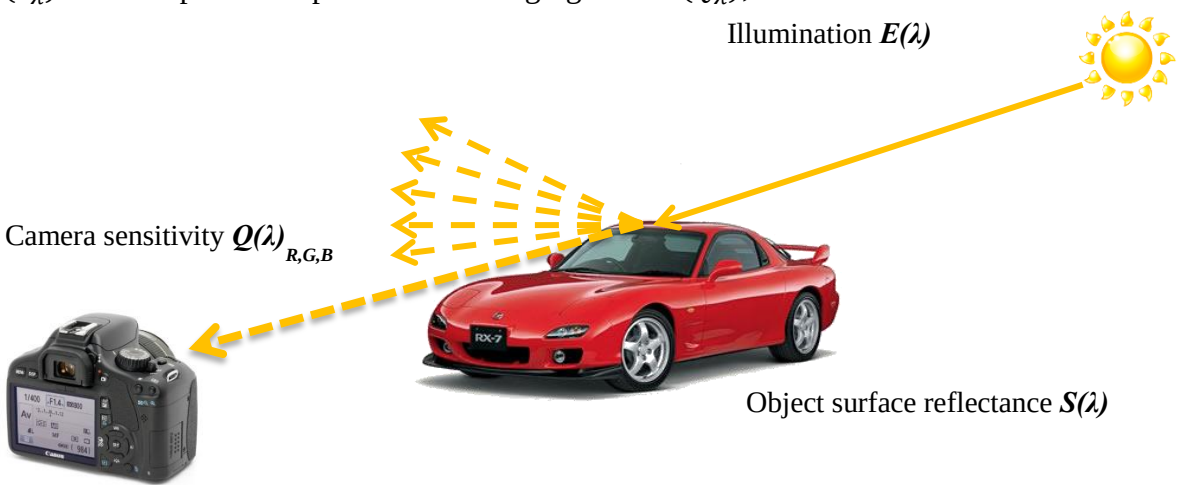


Figure 8 - Colour image formation

$$\boldsymbol{I_{R,G,B}} = \int_{\omega} \boldsymbol{E_\lambda S_\lambda Q_\lambda} \tag{9}$$

Equation 9 is a simplification of the dichromatic reflection model proposed by Schafer [47]. Where the interface (S_i) and body reflections (S_b) of the surface are combined to form $S(\lambda)$. Each of these components is affected by a number of factors, using the symbols introduced by Shafer;

$$\boldsymbol{S}(\boldsymbol{\lambda}) = \boldsymbol{S_i}(\boldsymbol{\lambda}, \boldsymbol{i}, \boldsymbol{e}, \boldsymbol{g}) + \boldsymbol{S_b}(\boldsymbol{\lambda}, \boldsymbol{i}, \boldsymbol{e}, \boldsymbol{g}) \tag{10}$$

Where, λ is the wavelength of light, i is the angle of incidence between the surface and illumination, e is the angle of surface emittance, and g is the phase angle (angle between surface normal and viewing angle).

2.6 Colour constancy methods

The theory of chromatic adaptation [48-50] has long proved to be the basis for many colour constancy algorithms within the computer vision field, due to both the ease of translation of the theory to 3 channel imaging devices and also the algorithmic simplicity of a linear model. The diagonal model evolved as a method to implement von Kries' equation in image processing in order to explain chromatic adaptation as a linear process. However as stated in the previous section chromatic adaptation on its own does not fully account for the performance of colour constancy as it neglects other perceptual factors at play within the phenomenon.

A number of algorithms utilise the von Kries transform after estimating the illumination in some manner. The most basic of these estimate the illumination by using certain specific scene statistics as cues to the illumination. The Retinex algorithm [3, 21, 51, 52] uses local area contrast, for example across two edges, to estimate illumination and remove its effect. This group of methods then utilise diagonal transform matrix to remap pixel values on a channel independent basis:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{11}$$

where, for a single pixel, R, G and B are the original intensities, and d_1, d_2, and d_3 are the independent channel rescale values needed to calculate the rescaled values R', G' and B' respectively. One subset of methods (typified by Buchsbaum [20]) relies on the grey world assumption , which assumes that for all images, under a canonical known illumination, the mean surface reflectance, and thus mean pixel value, is equal to grey. Buchsbaum [20] then adjusts the average of each channel independently to this grey value, performing a linear von Kries transformation. Another initial method for

illumination estimation is that of the white patch, or max RGB algorithm, which is referred to, with further advancements, in [53, 54]. Here the illumination estimate is based upon the assumption that a white patch exists in every scene. Thus the highest intensity pixels in each channel are found and translated to their maximum value. All remaining pixels are then translated by their corresponding channel shift. These two scene assumptions, that the average scene reflectance is grey, and that a white patch always exists in the scene, can be noted as the weak points of the algorithms. If a picture is taken of a seascape on a calm and sunny day we would not expect the average scene reflectance to be grey. However whilst we may easily state that these assumptions are unrealistic, we must accept why they have been made. Scenes may be made up from a near infinite set of surface reflectances, and a much smaller set of illuminations [48, 55]. This near infinite level of variability in scenes has meant that due to processing constraints, and for the need for the system to produce an estimate some scene assumptions must be made.. Scene assumptions also exist within these algorithms, they also both assume that within the scene only a single and uniform illumination exists. This then treats the complex 3D scene represented within the 2D image as a 2D scene. This area will be tackled later.

2.6.1 *Single pixel methods*

If we consider the two aforementioned algorithms (grey world and max RGB), they can be thought of as two implementations of the Retinex algorithm, except over a global spatial region. This difference between the use of the global and local regions in an image to estimate the illumination is a key difference across different colour constancy algorithms, and also across different image processing algorithms. How this global and local pixel information is used again separates different algorithms. That these algorithms use generally a single, or a very small number of individually monochromatic surfaces on which to base their illumination estimate is also another weakness. Brainard [56] explains how the use of a single surface, or simple sets of individually monochromatic surfaces (colour regions separated with no stepped edges) on which to base illumination estimation is not only incorrect, but also overly simplistic when compared to natural scenes. The grey world assumption does not take into account any of these surfaces at a separate level, and as such ignores the actual individual surfaces in the scene. The max RGB modus operandi of using spatially separate maximum channel intensities may indeed result in using more than a single surface for the estimate. However, this then means that each surface is only represented

by a single colour descriptor; a further abstraction from human perception and simplification of the scene. This surely complicates the task of separating surface reflectance and incident illumination. Even if the algorithm was adapted to use a single pixel, and thus surface, as the illumination estimate the same problem would exist.

With this knowledge over time alterations to these two algorithms have been made, resulting in a number of adapted versions. The grey-edge hypothesis [57], proposes a slight variation of the grey world hypothesis, in that the average differences between surface reflectances is achromatic. This produces a better performance than the original grey world and max RGB algorithms, and with a low computational complexity cost across a large dataset [58]. Funt et al. [53, 54] both investigate the max RGB algorithm further. Funt et al. [53] investigated whether exposure, or image dynamic range, has an effect upon the performance of the max RGB algorithm. This was due to the low dynamic range of images captured at a single exposure level in 8 bit space. It was found that image dynamic range, in conjunction with gamma ramping, does indeed have an effect upon the performance of the algorithm. This also makes complete logical sense as processes such as gamma ramping and clipping destroy important scene information.

2.6.2 *Histogram based methods*

The methods mentioned above either utilise single pixel information for illumination estimation, or the average value of channels. This can be considered to be a simplified representation of the colour information in the image. Single pixel methods ignore the overall distribution, and cumulative distribution of pixel values across the image. The previous single pixel methods followed the theory of channel independent adaptation as laid out within the von Kries transform. The use of a single pixel does however leave any illumination assumptions open to influence by both extreme scene characteristics and extreme sensor values. Extreme scene characteristics refers to contents such as specular highlights, whilst extreme sensor values can be connected to sensor malfunctions such as hot pixels or scene characteristics exceeding the performance capabilities of the sensor or incorrect capture settings, resulting in overexposure, or under exposure and noise. In un-calibrated imaging systems, chromatic aberrations due to lens characteristics can also cause colour casts within an image. However another level of image representation, in terms of pixel values, does exist: the histogram. The histogram allows the cumulative distribution of pixel intensities in the image to be visualised, yet it inherently ignores the spatial distribution of these intensities across the image. This means that histogram methods, much like the

pixel based estimation algorithms, ignore the specific object contents of the image. Histogram equalisation methods have generally been utilised to improve the contrast of an image due to an inappropriate exposure level across regions of the image. However strongly coloured illuminations also result in a colour cast upon a recorded image, and as such have a direct effect upon the histograms of the image. As such methods to remove illumination colour and intensity effects can be considered to be designed in the same vein as colour constancy algorithms. Previous histogram equalisation (HE) methods aim to maximise contrast across the entire image, whilst assuming certain scene characteristics, with no targeting of specific image contents.

As a breakdown of previous research, HE algorithms can be placed into two main groups: global and local approaches. Global HE approaches equalise the image using knowledge of the entire image's histogram. This is much like the global method of processing that is exhibited by methods which assume a single illumination and treat all pixels in the same manner, such as the Max RGB and grey world algorithms. On the other hand, with local approaches, the image is equalised using only local region histogram information. What this means is that these two methods assume different illumination characteristics across a given scene. The global method assumes that illumination is completely uniform across an image, while the local method works on the basis that the illumination will vary across, regions (in most cases rectangular) of the image. Even the local algorithm assumption can still be seen as an oversimplification of the scene contents captured in an image. These assumptions show the direct similarities between colour constancy and histogram equalisation methods. Furthermore these two approach methodologies can then be broken down into two algorithm types: automatic and manually controlled, with manual methods allowing the user to adjust certain processing parameters.

Most histogram equalisation methods follow the lead of the original HE (global) algorithm by Gonzalez [59]. Here the probability of each possible pixel intensity is first computed, and this is then used to calculate the cumulative distribution function (CDF) across the range of possible intensities. This CDF is then modified to equalise the distribution of probabilities. This is very simple approach that achieves a level of histogram equalisation, but the fact that the shape of the histogram is altered in accordance with a fixed statistical rule means that pixels which have the same intensity in the original image may have different values after processing. This means that this random pixel redistribution results in a great number of artefacts, with contrast

enhancement occurring sporadically across an image. This performance feature was noted by numerous researchers and led to global method research moving in the direction of altering a histogram related to features of the original histogram. An example of this is weighted threshold histogram equalisation by Wang et al. [60]. Examples such as [60] introduced the use of user defined thresholds into histogram equalisation techniques. Wang [60] uses the same basis of PDF and CDF calculation as does [59]. However the PDF is scaled between user-defined upper and lower thresholds. This results in equalisation between these limits, which if they are carefully chosen can yield much better results than [59], with fewer artefacts. However the existence of these defined thresholds results in the process needing to be carried out a number of times, which is not acceptable for an automatic solution. The values of the upper and lower thresholds which will result in the best possible equalisation result depend solely upon the image histogram and thus the contents of the scene, which have been ignored up to that point. This is a very common theme with equalisation methods and which has served as the main inspiration for the work presented here.

Further global methods have also been proposed [61, 62]. These are again variants of [59], but aim to counteract the unwanted brightness alterations caused by [59]. All of these methods work to try to create a linear cumulative density function with regards to pixel intensities. Kim [62] introduced a method called brightness preserving bi-histogram equalisation, which will be referred to as BBHE from this point. BBHE is different from normal HE in that it firstly splits an image's histogram into two histograms, using the mean value as the threshold point. This is an operation introduced to reduce the large brightness change caused by [59]. Histogram equalisation is then performed upon these two histograms, which are then placed back together to form the overall histogram. This is designed to preserve brightness by processing the brighter and darker sections of the histogram separately. The work by Chen [61] is a variant of this BBHE algorithm [62], aiming to minimise the mean brightness change introduced by [62], to prevent excess brightening with certain images. It proposes to do this by performing BBHE on the image after finding a threshold, which minimises the mean brightness change (error) between the original and processed images. This thus preserves the characteristic brightness of the original image. While these algorithms may not always produce the best results visually, their attempt to preserve certain image characteristics is important to note.

As was stated previously, local equalisation approaches have also been proposed. The

original paper by Hummel [63] performed histogram equalisation across an image using a neighbourhood of pixels approach. This does produce very different results in comparison to those produced by histogram equalisation, but can lead to different areas of the same object being edited differently, depending upon their surroundings. This effect can be mitigated through variation in pixel neighbourhood sizes. Stark [64] introduced a more complex form of local processing, using the frequency domain. However all of these previous methods involve severe modifications to an image's histogram, which then means that the contents are also severely altered. Caselles et al. [65] introduced the thought that the preservation of histogram shape in local areas should be of utmost importance, due to the importance of the connected nature of the scene in local areas. Caselles at el. [65] proposed what is most likely the most complicated histogram equalisation methods, which also works in an iterative fashion, but which produces good results. However it was only proposed for grayscale images, like all of the previous work.

Originally histogram equalisation methods were proposed solely for grayscale images, however as time has gone on and with the abundance of colour images the need for colour image equalisation methods has become apparent. Due to the more complex nature of the relationship between the three channels in an RGB colour image the application of the previously described algorithms to each channel independently results in the introduction of a number of artefacts. With a grayscale image algorithms such as [59-65] may introduce simple salt and pepper noise artefacts, however with RGB images these artefacts can be more apparent due to the chances of them exhibiting abnormal colours. The separate processing of different channels can also lead to large changes in colour ratios for objects, resulting in very abnormal colour shifts and therefore in object colours in the processed image being very different from those in the original scene, thus destroying scene information. Due to these problems, colour image equalisation techniques have had to be developed separately resulting in a number of new approaches, most of which follow on in philosophical terms from those previously proposed. Work by Limare [66] is perhaps a perfect example; it is a very simple algorithm, but one which is quite powerful. The simplest colour balance algorithm, or SCB, sets the top and bottom 1% of pixels (in bin terms) to 255 or 0 respectively. The remaining pixel values are then redistributed using an affine transform, which utilises their value and the previously calculated upper and lower 1% threshold levels. This ensures a level of contrast enhancement by altering the histogram distribution. Morel et

al. [67] proposed two algorithms for contrast enhancement, limited slope histogram equalisation (LSHE) and local contrast enhancement (LCE). The importance of this paper is that it begins to discuss the importance of ensuring that images, after processing, still look natural. It gives a brief survey of artefacts introduced by the most well-known equalisation algorithms for both grayscale and colour images. It proposes LSHE and LCE as schemes which will not introduce the same features into images. However both of these algorithms still alter the characteristic shape of a histogram, again altering the information. Han et al. [68] proposed a 3D approach to histogram equalisation, by attempting to equalise the luminance component of an image.

With these similarities between the two types of algorithm, contrast enhancement and colour constancy invariance, some work has attempted to bridge the two [69]. Finlayson et al. [69] worked upon the basis that incident illumination will cause a colour cast, and thus shift, upon the histogram distribution in each channel. The method works, once again, upon the independence of each channel as in the von Kries transform, and the assumption that under a white illumination, assuming a good exposure level, each channel histogram will fill the maximum dynamic range at equidistant intensity levels. The existing intensities in the histogram are counted and then equally spread across the dynamic range. This results in the order of intensities, or monotonicity, to be preserved thus retaining the relationship between intensities, and thus contents, on a channel-by-channel basis. However the method does have some weaknesses. The first weakness is that of the method of normalising each channel. Here preserving and equally distributing the existing intensities, without regard to their cumulative total. This can lead to a decrease in performance if certain scene features exist. Scene features such as specular highlights are not exhibited in a histogram by a large impulse at a single pixel intensity. They are exhibited by a distribution around clipping, with the tail of this distribution sometimes only resulting in very small cumulative pixel totals for certain intensities. For this algorithm, which bases normalisation upon a small number of large cumulative total intensities, this can lead to a drastic reduction in normalisation scaling.

2.6.3 *Gamut based methods*

A third way in which to analyse the colour data within an image also exists, that of utilising the relationship across the colour channels by analysing the overall colour distribution. Forsyth [70] proposed using the colour information at every single pixel within the image to help estimate the illumination, assuming a single uniform illuminant. He proposed that under a given illumination the total number of colours - or gamut -

that is able to be produced is finite. As such under a known illumination the maximum possible gamut can be set at a specific amount. Thus if we find the diagonal transform which maximises the fit of the image gamut within the canonical we find the most likely transform. The assumption within this method is not that a particular surface (e.g. a white surface) exists within the particular scene, but that a single convex hull can appropriately represent all possible surface reflectances under a known illumination. Whilst Forsyth [70, 71] described the approach in detail he did not show how to implement it. Therefore, a number of different methods implementing the gamut mapping method in different ways have been devised. Forsyth [70, 71] described the gamut to be transformed as a 3D one, if we assume a 3D colour space such as RGB. However, Finlayson [72] noted that working within a 3D space, and transforming a 3D gamut is computationally intensive, and had previously proposed the use of a 2D space, i.e. the chromaticity plane [72]. Finlayson [72] showed the use of this simple 2D space was the same as utilising the originally proposed space of higher dimension. However the assumptions, and weakness, of Forsyth's approach were laid out within [73]. Finlayson et al. [73] stated that, on top of the single uniform illumination assumption, the inclusion of any specular or shape (shading) colour information would also have an effect upon the composition of the image gamut under the unknown illumination, as these are not related to the true surface reflectances present within the scene. This then has the subsequent effect of incorporating this error in the resulting diagonal transform. Further to this the original algorithm allows for an extremely large number of illumination estimations to be made, with no regard to their likelihood. This likelihood of illumination is feasible if we consider images taken under natural capture conditions, and was acknowledged and utilised by [73] to simplify and increase the accuracy of the estimated transforms.

Finally the original algorithm also allows for the possibility of the ordinality of colours in a given gamut to be altered, which [72] shows not to be an accurate assumption. Moving on from this it also ignores the fact that certain illuminations may also alter the relative distribution of chromaticities within the gamut itself. If the method utilises a convex hull to create the representation of the image it signifies the acceptance that the most extreme chromaticities are the most informative, and that all points within the gamut exist within the image. This means that the existence and distribution of chromaticities within the hull are ignored, thus ignoring the contents of the image at a lower level.

Following these studies, a number of different gamut mapping variants were developed, with different numbers and types of axes used for gamut representation. The first of these combines gamut mapping with the white patch method as described, with additions, in [53, 54], tackling the weaknesses of the basic approach. With the methodology of the Max RGB function described previously it can be imagined that any extreme pixels will cause the resulting illumination estimate to be in error. Specular highlights will affect all three channels for a given pixel, whilst a hot pixel will affect one channel for a given pixel (assuming a single channel hot sub pixel). Joze et al. [74] attempt to counter any problems caused by specular highlights and over exposure for the illumination estimate by altering the choice of pixels used for estimation. Rather than using the highest intensity pixels in each channel, a method which is susceptible to error, Joze et al. choose to ignore pixels which exceed 90% of the dynamic range for each channel. The top 1, 2 or 5% (T%) of remaining pixels in each channel are then utilised as the qualifying white patch, and then perform a gamut mapping algorithm based on this white patch gamut. The use of this 'white patch' gamut, rather than the overall image gamut, is also of importance. As the white gamut is inherently smaller, the gamut better represents the distribution of colours within this 'white patch' region, as there will be fewer points within the gamut's convex hull which are not actually found within the image. This means that the white patch is more informative. The performance of this method was shown to be better than the standard Max RGB and gamut mapping methods. However it must also be noted that the illumination does not always exist within this white patch gamut region, and so the performance of this algorithm is again directly related to the contents of the scene, much like its predecessors, although this method is far more robust.

Tominaga et al. [41, 42] propose a variation on Forsyth's gamut mapping method. Here they do not attempt to utilise the method of estimating an appropriate diagonal matrix to transform the gamut of the image into the gamut of the canonical set. Rather than attempt to identify the illumination and its tri-stimulus definition they have developed a method to estimate the colour correlated temperature (CCT) of the illumination using the gamut. Similar to [72, 73], they utilise a two dimensional colour space for gamut representation. However they opt to utilise RB sensor space, a simple dimension reduction from the RGB space in [70, 71], after applying an intensity scaling based upon the brightest pixels. The use of the convex hull also signifies the acceptance that

the most extreme chromaticities are the most informative, and that all points within the gamut's convex hull exist within the image.

A reference dataset was developed of the canonical image gamuts under a range of different illumination CCTs, with illumination CCTs converted into the mired scale;

$$CCT_{mired} = \frac{1\times10^6}{CCT} \tag{12}$$

Each image gamut under an unknown illumination was then correlated to the reference dataset using the following calculation;

$$r_i = \frac{A_{Ii}}{\left(\sqrt{A_I A_i}\right)^2} \tag{13}$$

where A_I is the area of the image gamut, A_i is the area of the reference gamut and A_{Ii} is the size of the overlap between the two. This is a computationally simple approach, and does not make any attempts to assume how colour constancy works (such as by utilising the von Kries transform). However it does have problems.

Firstly if we only consider illuminations on the basis of CCT we do not account for spectral differences across CCT metamers. Some metamers can have drastically different spectra, as shown below using the system outlined in [29]:

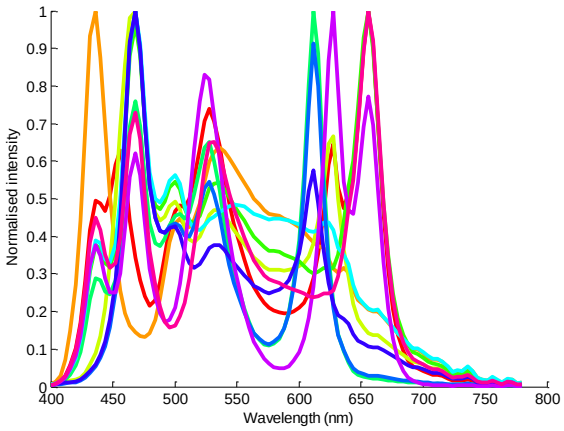


Figure 9 - A set of 10 D65 metamers, illustrating their varied spectra.

Object irradiance will similarly vary dramatically under different metamers, and thus recorded RGB colour may also vary. Thus if any colour correction is to be undertaken with the illumination estimate, chromatic errors in the resulting image may occur. However if the algorithm is designed to be used only with outdoor naturally illuminated images, the algorithm can sidestep this problem. An assumption of natural daylight illuminations means that no CCT metamers will be encountered due to the relatively

invariant, broad-band spectral nature of natural daylight illuminations. However under artificial illuminations, specifically indoors, these will be encountered, and as such this restricts the use of the algorithm.

Secondly whilst the mode of gamut correlation is robust, as it takes both overlapping and unique gamut areas into account, the use of solely the convex hull may have problems. As stated previously the use of the convex hull means that the distribution of points within the gamut is ignored, with only the extreme points being used. This works well when illumination changes result in large convex hull gamut changes, however if this is not the case illumination estimation accuracy may suffer.

The theory of relating the colours present in an image to those that can appear under an illumination is not only linked to gamut mapping methodologies. The idea of using the colours present in an image to estimate the illumination is also within the colour by correlation method [75]. This correlates the colours within an image to the colours that can exist under different illuminations, bypassing a problem of convex hull mapping methods in that only the colours existing within the image are used to estimate the illumination. This is in contrast to convex hull methods that implicitly assume all colours within the hull exist.

2.6.4 *Multiple image methods*

All of these previous methods are designed on the premise of using a single image with which to estimate illumination. This makes sense when designing from the standpoint of photography. The human visual system does not work on this single shot premise however. The visual system is constantly updated with imagery and the influence of this image stream on its achievement of colour constancy is still unknown. The utilisation of multiple images for processing for different purposes is the basis of the field of computational photography. These approaches offer a great deal of potential for future colour constancy methods.

The correlation between colours that exist under different illuminations, noted previously, has been taken further with a computational photography technique. The chromagenic camera method [76] utilises two or more images taken of the same scene with two known and characterised filters. The relationship between RGB values recorded under these two filters for an illumination is known:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M_{filter} \begin{bmatrix} R'' \\ G'' \\ B'' \end{bmatrix} \quad (14)$$

Here the unfiltered and filtered images of the same scene can be considered as the scene captured under two different illuminations, but linked by the same incident illumination. RGB triplets in both of these images are related by a diagonal model, determined by the incident illumination (not that introduced by the filter). The behaviour of M_{filter} for a given RGB triplet as such, is known, a priori, for a set of illuminations. Thus consequently illumination estimation across them can be undertaken by selecting the illumination with the minimum error. As such this method can be considered a sub set of a machine learning based solution to the colour constancy problem. Performance was seen to vary depending upon the choice of filters. Filter choice must be carefully selected to maximise the colour differences across different illuminations. This method can be seen as a different manner in which to utilise the theory behind gamut mapping and colour by correlation methods. Its only disadvantage relates to the need to capture more than one image of the scene with a change of filter. However with the capabilities of camera devices two images can be captured with very little delay between the two images, minimising the probability of changes between the two images, as in now commonplace with HDR systems. This approach can also be extended to different definitions of the utilised filters. These can be optical filters, or introduced illuminations, if you consider the combination of the unknown and known illuminations to behave as such a filter. The task of changing filters can also be performed relatively quickly with electronically controlled tunable filters, such as those in hyperspectral cameras. However the risk of incoherent scene registration between two images was noted in [77], which has introduced a method to allow chromagenic colour constancy to be performed on a single image. This involves the application of the second filter to only the spare G channel sub-pixel as is commonplace with Bayer sensor distributions. Application of a filtered R or G or B pixel to this spare sub-pixel is alternated. This does result in a spatial resolution decrease for the imaging sensor due to a larger basic mosaic size.

Fredembach et al. [78] also tackles some of the underperforming aspects of the original chromogenic algorithm, in that under some illuminations large estimation errors occur with the original model when estimation was based upon certain colours. [78] improves the method through investigation of the performance of the original method. It was found that the brightest, but not most saturated, RGB values resulted in the best performance of the algorithm. Therefore the bright chromogenic algorithm only utilises

pixel values which are within this region. The method also tackles the image registration problem of the original method by assuming that the brightest pixels across both images will be the same surfaces.

Finlayson et al., and Shrestha et al. respectively [76, 78] use multiple images taken with two different filters to estimate the illumination, but other methods exist which utilise a number of frames to estimate illumination. Wang et al. [79] tackled the challenge of illumination estimation using all of the frames from a video stream. As previously noted, colour constancy methods have overwhelmingly been designed for frame-by-frame processing of images. However if an algorithm is heavily dependent upon utilising certain specific image features and this method is applied to a series of frames, illumination estimates can vary from frame to frame. This results in inconsistent white balancing across a period of time. To counter this, Wang et al. [79] splits a video stream into different scenes, depending upon detected histogram and colour changes, and then averages the illumination estimate across these scenes before correction. This results in a consistent white balancing across scenes. However this method is entirely post video capture, and as such cannot be thought of as similar to the process within the human visual system.

2.7 Comparing human studies and computer vision models

All of these methods listed above have treated the imaging sensor as a direct and accurate replacement for the spectral sensitivities of the cone cells, while basic knowledge regarding the human visual system shows these not to be directly comparable. The methods also inherently assume that the human visual system utilises only sensory mechanisms based on the responses of the L, M and S cones.

All of the previous methods have utilised the RGB data from the imaging sensor and not attempted to pre-process this in any way. The theory of chromatic adaptation assumes that the sensitivities of the three sensitive cones are independent from each other [80], but if we consider the spectral sensitivity of the channels of a generic imaging sensor we can see that their sensitivity responses are not independent [81]. Some response overlap does indeed exist, and therefore, for example, if there is an overlap between the green and red channel responses then the resulting red and green intensities at a given pixel will be correlated. Finlayson and others showed that spectral sharpening does indeed improve the performance of diagonal transform based methods [35]. The performance of the diagonal model has also been investigated by Finlayson [82], who states that the

diagonal model performs well with Forsyth's gamut mapping method [70] unless surfaces within the image are not encapsulated within the canonical gamut. In cases like this Finlayson et al. state that the mapped gamut should be translated such that all points within the mapped image gamut will subsequently exist within the canonical. This results in an adapted form of the diagonal model, the diagonal-offset model:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix} \tag{15}$$

where, for a single pixel, R, G and B are the original intensities, d_1, d_2, and d_3 are the independent channel rescale values and o_1, o_2 and o_3 are the offset values needed to calculate the rescaled values R', G' and B' respectively.

We can consider that use of local and global pixel information, or solely looking at pixel values, does indeed use the contents of the image in an abstract way. However the manner in which it uses this information is also of key importance. Algorithms such as grey world, and max RGB, along with their derivatives can however be considered to ignore the real perceptual contents of an image. Under real conditions the human visual system, it can be reasonably assumed, achieves colour constancy not through the perception of individual pixels, or spatial regions but through the combination and resolution of these regions into different surfaces and objects.

This question of whether human colour constancy improves with certain specific objects, whether known or unknown, has long been put forward. Whilst certain studies have found certain conditions may indeed aid our colour constancy [26, 83], most test conditions do not always relate well to real world conditions:

Figure 10 - Comparing a grey, mondrian and real world scene.

The difference between these test and real world conditions is not only related to the scene composition on a global scale, but also connected to scene contents. On a day-to-day basis our visual system interacts with known objects, and the theory that colour constancy is affected by interaction with known objects is also an interesting one. As has been previously discussed, many histogram equalisation and colour constancy

methods have ignored the specific scene contents, ignoring what may be an important part of the mechanism in the HVS. They have however worked upon the basis that the higher the number of unique surfaces in the scene the more likely the illumination estimation is to be accurate. They also assume that all of the surfaces are matte, in order to assume uniformity of shading across each surface. Without this shading uniformity each surface could be split into different surfaces.

The smaller the total number of unique surfaces that exist in the room the higher the number of likely illuminations; as such this makes it easier to select a more probable illumination. Illumination estimation in scenes with one unique surface has long been considered a problem [84]; a white wall under a given coloured illumination looks the same as a wall of that same colour under a white light. However some image processing methods have attempted to tackle this challenge [84]. Finlayson et al. [84] attempt to see if a single surface can act as an accurate cue to illumination using the dichromatic model. Whilst it does indeed produce good results in comparison to the max RGB and grey world methods, the assumption of a single surface in a scene is highly different to what the human visual system experiences on a day to day basis. Whether the human visual system can actually achieve colour constancy under these scene conditions is another question entirely. However, if we consider the 'single surface' stimuli used within [84] we must question what the definition of single surface is. [84] defines single surface with its use of a green plant, as such any surface within the same colour naming group and object have been assumed to be within the remit of the single surface definition. However as [84] notes, the plant image contains other information such as specular highlights. This means that this single surface is not uniform in terms of shading or in terms of colour, based on visual checks of the images.

This assumption of using a single surface for illumination estimation, whilst not correct in this case, has been taken further by numerous studies across both psychophysics [15] and computer vision [16, 85]. Hurlbert et al. [15], as mentioned previously utilises a single object surface, but places importance upon investigating whether the polychromatic nature of real objects aids human colour constancy. This does indeed test and replicates stimuli encountered in day to day scenarios much better than previous assumptions of multiple matte surfaces. The usefulness of polychromatic surfaces in machine vision has also been proven and utilised for camera calibration [86]. Polychromatic surfaces are also key to and exist in abundance within computer vision

scenarios which can be aided by accurate colour correction and/ or constancy such as facial recognition [43].

2.8 Polychromatic contents based processing

2.8.1 *The use of skin*

Kakumanu et al. [43] outline the importance of colour correction of facial images for facial recognition methods. A lack of a real solution for colour constancy has led to the ignoring of colour information in most facial recognition algorithms. If we consider general imaging, most colour inaccuracies introduced during capture will be noticed for objects that people know, as with machine vision recognition algorithms viewing known objects [17]. On a day to day basis people generally take pictures of people, which means that many pictures can be assumed to contain images of human skin. Typical colour constancy methods have assumed the presence of other image contents or scene statistics and largely ignored specific contents of images, except for a few instances [87]. Although some methods have achieved a good level of performance, the question of whether specific image contents aid machine colour vision still exists.

The intertwined nature of skin and illumination colour in images has been investigated from a number of angles. These extend from human skin detection under unknown illuminations [85, 88-92] to illumination estimation using skin as a cue [93-96]. The underlying variation in skin chromaticities under different illuminations has been utilised to help design sensors which can better detect skin across different illuminations [97].

Störring et al. [89, 91, 94-96, 98] carried out a large number of investigations into the variations of skin chromaticities under different illuminations. Störring [96] goes into detail regarding illumination CCT estimation using human skin and its dichromatic reflection model. It works within the rg chromaticitity space, with co-ordinate values from an RGB camera calculated in the following manner;

$$r = \frac{R}{R+G+B} \tag{16}$$

$$g = \frac{G}{R+G+B} \tag{17}$$

$$b = \frac{B}{R+G+B} \tag{18}$$

The method works on the principle that, for a given melanin level, skin colour will vary within a dichromatic plane.

It noted how human skin is nearly achromatic, but also how it cannot be assumed to be a single surface. Across a given patch of skin the surface reflectance function will vary due to factors such as uneven melanin levels and differing blood flow levels. This means that human skin is, like natural fruits [15], a polychromatic object, and as such should contain more cues to help estimate illumination. Whilst skin may be inherently polychromatic, [96] also states that, as is expected, a very low level of specularity exists within images from its samples. In order to build this model the variation in chromaticities for skin of different melanin levels were modelled into iso-melanin-lines. As such when an image of skin under an unknown illumination is presented the melanin content of the skin is first estimated and the position of the skin on this estimated iso-melanin-line is used to then determine the illumination.

There are a number of points to be made regarding this approach, the first regarding the choice of colour space. The rg chromaticity space was chosen due to its intensity invariance; thus the chromaticities across a given patch of skin are constant in coordinate location for a given illumination even across different intensity levels. This space has an advantage in that it is intensity invariant but also device independent. However it has no connection to human colour perception, and as such cannot relate to mechanisms within the visual system.

Second is the error measure for estimated CCT. Störring et al. [96] compute the Euclidean (RMS) difference between the estimated and actual illumination CCT, whilst also showing the maximum error. However the usefulness of this metric needs to be considered, because whilst the Euclidean distance may be accurate, even with the shape of the Planckian illumination locus, perceptually it has no relevance. This is because the perceptual difference between two sets of equally spaced illuminations, in Kelvin (K), varies depending upon the location of the two illuminations along the Planckian locus.

Whilst Tominage et al. [41, 42] noted this, they chose to utilise the mired scale for comparing the difference between their actual and estimated illuminations. This is useful for showing the difference between two illuminations, and as such the value of a filter needed to equalise the two, but again does not relate to the perceptual difference between them.

Thirdly the philosophy of estimating the CCT of the incident illumination also needs to be discussed. As previously noted and stated different CCT illumination metamers can have very different spectral distributions, and thus different resulting irradiances from the same objects. However this method can work well if we assume all images are taken outdoors under natural daylight conditions, where the CCT generally specifies the spectrum. However under artificial (man-made) lighting, CCT metamers of daylights do exist, and can have differing spectral distributions.

Störring [95] extended the previous work to investigate the use of specular highlights on skin as a cue to estimate illumination, as have Kwon et al. [99]. Störring [95] tests the use of facial specular highlights across different lighting geometries. It utilised PCA analysis to separate body and surface reflectance components and non-specular and specular components respectively. In order to test their hypothesis an image database of faces captured with differing capture angles and incident illuminations was constructed. Different illumination CCTs were produced through the utilisation of different filters in front of a fixed illumination source. This light source was a 3200K light bulb, whose spectrum was not shown.

Algorithm performance was again compared using a non-perceptually based metric, that of the Euclidean distance between the estimated and actual illumination in RGB space and the angular error between the two. Kwon et al. [99] also utilise the rg chromaticity space to investigate the ability of an algorithm to estimate illumination using specularities on human skin.

In order to gauge performance, the metrics of absolute CCT difference (ΔCCT) and relative error (E_r) have been utilised:

$$\Delta CCT = CCT_{Actual} - CCT_{Estimate} \tag{19}$$

$$E_r = \frac{\Delta CCT}{CCT_{Actual}} \times 100 \tag{20}$$

Both of these error measures are again not perceptually relevant, and as such not useful in comparing performance relative to the experience of a human observer. This is also the same for the final performance metric D_{uv};

$$D_{uv} = \sqrt{(u_{Actual} - u_{Estimate})^2 - (v_{Actual} - v_{Estimate})^2} \tag{21}$$

Whilst the two methods above achieve some level of accuracy with regards to illumination estimation they cannot achieve colour constancy in the same manner as the human visual system. Furthermore we have seen how performance of the above algorithms is not considered with respect to the human perception of colour. This is in part connected to the intended end use of these algorithms being within a machine vision facial recognition algorithm.

The nature of skin chromaticities within LMS space has been investigated by Cotton et al. [100], who investigated this in order to detect other skin features such as tumours. Their investigations discovered that all skin colours lie on a 2D plane in LMS space, allowing for a dimension reduction in colour space and a reduction in computational complexity. This shows that within LMS space skin lies within a predictable region of the colour space, which should aid skin detection methods. Regions where skin colours exist have also been found in a number of other colour spaces. Soriano [90] shows the skin locus within rg chromaticity space, and [43] has covered the detection of skin across rg, HSV (and variants), orthogonal spaces, CIE spaces along with other colour spaces (colour space mixtures). It can be said that if colour space transforms are uniform then if a region of skin colour exists in one space it will also exist in another, as such this wide range of findings is largely expected. However it must also be added that even though these regions of skin colour exist, they are not regions of the colour space exclusive to skin colours, as such skin detection based solely on chromaticity will encounter some degree of error. This has been noted by researchers previously; Ruiz-del-Solar [101]. The fact that skin colours lie along melanin axes [85, 96] across illuminations means that skin may move in a predictable manner across this skin locus.

2.9 Performance analysis methods across the two fields

Across the previous section overviewing the literature of developed methods for achieving computational colour constancy a few words have been paid with regards to methods for gauging algorithm performance. These methods for gauging algorithm performance are based upon both the aim of the algorithm and generally the colour space of algorithm operation. Some methods attempt to recover the colour of the illumination, whether in RGB space or another, whilst others attempt to do so by estimating the CCT of the incident illumination. Barnard et al. [102] compare a number of different colour constancy algorithms and gauges performance based upon the rg estimates of the methods. To do this it utilises the metric of rg angular error and rg Euclidean distance between the actual illumination and illumination estimate;

$$\theta_{error} = \cos^{-1}(\widehat{\iota_{act}} \cdot \widehat{\iota_{est}}) \quad (22)$$

$$\Delta_{rg} = \sqrt[2]{((r_{act} - r_{est})^2) + ((g_{act} - g_{est})^2)} \quad (23)$$

Where $\widehat{\iota_{act}}$ and $\widehat{\iota_{est}}$ are the normalised rg or RGB vectors of the actual and estimated illumination, and r_{act}, g_{act} and r_{est}, g_{est} are the r and g values of the actual and estimated illumnations respectively. Other papers, which test developed methods in RGB space, utilise the same metric in RGB space. The RGB angular error and Euclidean distance measure are very commonly used to compare performance of different methods. Whilst they may quantitatively compare the illumination estimates in a method which directly relates to the method of adapting the raw image, these metrics do not relate to the perceptual difference between the actual and estimated illuminations. In fact, the same value of angular error and Euclidean distance in different regions of the RGB gamut can result in varied perceptual differences.

If we consider methods which attempt to estimate the CCT of the illumination [41, 42, 96], whether actual CCT in K or the mired value scale, differences in CCT scale values are utilised. This method however again has the disadvantage of not being related to the perceptual chromatic difference between the estimated and actual illumination.

Perceptual methods of gauging chromatic and colour differences do exist, and can be taken from psychophysical research in the colour constancy field [25]. The use of the $\Delta E_{a^*b^*}$ and $\Delta E_{u^*v^*}$ metrics are perceptually accurate as described previously. As mentioned earlier $\Delta E_{a^*b^*}$ is only accurate under the illumination of D65, however $\Delta E_{u^*v^*}$ is accurate across illuminations.

2.10 Existing image datasets for research

In order to standardise the performance evaluation of developed algorithms and also to aid the development of new techniques a number of different image databases have been created and used for study into colour constancy. These databases are designed for different purposes, and as such they contain varying levels and types of data. The use of a database may be to test object recognition algorithms, colour correction algorithms or to test other image processing abilities such as pattern recognition. As such the supplied data and images should give a particular advantage to an algorithm within its field. General colour constancy algorithm databases should provide a number of images of the same object(s) under varying degrees of illumination, but there are also a number of other pieces of information that a database should supply;

1. Objects
    a. Number and types – The database should consist of images of a wide variety of objects exhibiting as large a range of features and characteristics as possible. The objects used should take into account different reflectances, textures and likelihood to be present within photos.
2. Illumination and specific characteristics
    a. Colour – Details regarding the colour of the illumination used should also be supplied. This could range from a simple RGB of the illumination to a detailed spectrum description.
    b. Intensity – The strength of the illumination should also be supplied to the end user along with any variation across the scene.
    c. Number – A variety of lights and colour in one scene along with specific data on the variation of the spectrum across that scene could aid algorithms aimed to work with multiple illuminants.
    d. Geometry – The geometry of the light(s) in the scene should also be noted and supplied.
    e. Uniformity – The uniformity of illumination across the scene should also be measured and noted as to allow a complete description of the entire scene.
3. Scene characteristics
    a. An in depth description of the setup of the experimental scenes should be supplied.
    b. Scene type – Indoor/ Outdoor allowing more realistic testing of proposed algorithms.
    c. Canonical references within scenes – Every scene (as far as possible) should also have a canonical version to allow the results of any processing to be easily and quickly checked with performance metrics. This could also be carried out by using a colour checker in each scene for example or by having an image of the scene under the canonical illumination for the database.
4. Types of images supplied
    a. Tri band imagery – Images from normal RGB cameras
    b. Hyper spectral images
5. Characteristics of camera(s) used for image acquisition
    a. Spectral sensitivity(ies) of used camera(s), and sensor layout (Bayer/ Foveon etc...).
6. Image characteristics

a. Resolution of images – Varying degrees of resolution supplied to aid the speedy download and processing of images.
b. Quality – As any colour constancy algorithm used in a camera or system will use RAW image data this should also be supplied along with compressed versions.

7. Image capture camera settings
    a. Tri band camera - Gain, zoom, exposure, white balance and resolution settings.
    b. Hyperspectral – Spectral and spatial resolution.
8. Image data type
    a. If the database is solely for the purpose of colour constancy then only colour images should be supplied, however their varying levels of image bit resolution could also be made available.
    b. If other uses such as object recognition are to be used the binary and greyscale versions of images could also be supplied.

A table comparing current databases can be found within the Appendix. From the table it can be seen that previous image datasets [103-111], for technological reasons, have largely been of RGB images. However more recently hyperspectral imaging has become more commonplace. Captured scenes have ranged from a number of objects in different orientations [103], to natural scenes [104, 105, 107, 111] and objects under different illumination conditions [104]. However no dataset has, as yet, combined hyperspectral images of the same scenes under a large number of illuminations. This leaves an opportunity for potential work.

2.11 Bridging human colour perception and computer vision

From the prior review of both psychophysical and computer vision research we can see that the challenge of colour constancy is of utmost interest to both fields. We have also seen that some overlap between the two fields does exist and can be easily noted. Such examples of this are the direct relation between the diagonal model and von Kries chromatic adaptation theory, moving towards learning based methods which are based upon the importance of prior knowledge and perception. Research within the overall image processing field has been incorporating a greater number of results from the psychophysical field for a number of years, and this will continue to occur as the development of technologies aiming to automate currently human observed tasks increases.

Whilst current RGB based imaging sensors have been shown to provide a good level of performance the development and cheaper commercialisation of novel imaging technologies will enable an until now unparalleled series of development opportunities to not only match the capabilities of the human visual system, but also surpass it. This thesis aims to reinforce not only the applicability of trichromatic imaging sensors, but also the importance and the possibilities of using them to implement perceptually accurate and based algorithms. It also aims to show the possibilities of algorithm development based upon hyperspectral imaging and how this data can aid in situations where the human perception of colour and machine vision merge.

2.12 Summary

From this literature a number of areas of research interest have been uncovered:

1. Bridging psychophysical and computer vision research.
    a. The potential combination of psychophysical research findings into image processing algorithms, whether through designing algorithms to utilise specific features or aiming to replicate visual processes from the ground up.
    b. By carrying out psychophysical testing, can we uncover mechanisms in the human visual system which can aid the design and performance of computer vision approaches?
2. Replication of the human visual system.
    a. Colour constancy is a mechanism within the visual system, so can we achieve better results by replicating it further than has been achieved so far?
    b. By replicating the operational colour space of the human visual system, can we better match human performance?
3. Contents based image processing.
    a. The potential use of specific image contents, , and potential links to the human visual system, are of great interest. Specifically, can we design an algorithm that can utilise specific image contents to estimate illumination? Can we also design methods to preserve image contents whilst enhancing them?

Chapter 3. Luminance Skewness Based Monotonic Colour Normalisation

3.1 Introduction

The review has shown the state of the art approaches to colour constancy investigation and processing and also the link between contrast enhancement and colour constancy methods (both illumination estimations and colour normalisation). Inconsistencies between these fields, or the lack of integration between them has led to the development of different solutions. These are algorithms, that while do improve certain visual features of an image, do ignore the perception of the scene for a human viewer. Whilst RGB based sensors and camera systems cannot be expected to capture images to the same level as the human visual system we aim to show, in this chapter, how the captured information can still be processed to improve within-image contrast whilst utilising certain aspects of human visual perception to do so.

Within this chapter we show a method to contrast enhance images, both grayscale and colour, by utilising a statistical metric that has be showed to be linked with human gloss perception which also preserves scene contents. The work presented within this chapter has its foundations in previous work [69], which showed the performance of the monotonic transform when dealing with images under strong colour casting or over/under powering illuminations. The work here will tackle these images by introducing a new variant of the transform to deal with images exhibiting highly skewed luminance histograms, which have been linked to gloss perception in humans [112-114]. The variation of luma skewness and other luma histogram statistics with illumination colour and power will first been shown. We will show that the luma (RGB derived luminance) skewness monotonic transform produces contrast enhanced and roughly illumination invariant results that are also perceptually texture invariant.

3.1.1 *Algorithm Inspiration*

The connection between the tasks of contrast enhancement and colour constancy were noted previously within Chapter 2, with papers such as [69] beginning to bridge the gap between the two areas of image processing. As stated previously this connection between the two different challenges is due to the colour casts that some, generally chromatic, illuminations introduce into images [69]. Whilst colour constancy within image processing is an attempt to reproduce a mechanism within the human visual system, other image processing tasks are of a slightly different nature. The challenge of colour constancy is inherently related to sensor capabilities and a lack of comparable

processing to that within the human visual system. A number of other challenges also exist within the image processing field which, whilst are not connected to the replication of a process in the human visual system, are linked to different capabilities of the imaging system. In the human visual system this relates to the eye, while it relates to the entire camera system in the image processing field.

The challenge of contrast enhancement is directly linked to scene characteristics exceeding camera capabilities, whether due to incorrect capture settings or baseline camera capabilities. Once an image has been captured at an incorrect exposure level post processing is the only method through which contrast between surfaces and objects in the scene can be either increased or reduced. The manner and degree to which contrast is altered in these images will result in different perceptual results for a viewer. Links between general psychophysical research into visual perception and image processing, whether grayscale or colour, were initially thought of as important [3]. Most image enhancement algorithms after this however began to deviate away from giving consideration to human perception [60-65, 68, 115]. Recently the importance of perception has been taken on board and led to the development of many algorithms [66, 67], specifically with the rise of interest within colour image processing. When dealing with colour contrast enhancement techniques are at risk of introducing their own colour casts if colour channels are treated separately.

Before embarking upon any type of image processing we should understand the scene that is represented by the image. The main aspect of this is the acceptance that an image is a 2-D representation of a complex real world scene. This real world scene is complex in that it is 3-D, can contain any number of unknown surfaces in any number of orientations relative to the imaging device and be under an unknown number and geometry of illuminations. Figure 11 shows an example of a scene under the single uniform and non-uniform illumination assumptions currently used by histogram equalisation methods. This 3D scene is recorded as a 2D image by the imaging device. As is shown by Figure 11 while some of these objects may appear adjacent to each other in the final image they may be separated by some depth distance in the actual scene. As such they could be under very different illumination spectra and intensities. Both global and local histogram equalisation methods ignore this for the sake of simplicity, our novel method will also ignore this. It must also be stated that objects that are within the same depth plane from the camera can also still be separated by a large horizontal distance. This again means that they could be under differing conditions, a

fact which is utilised by local methods [63-65].

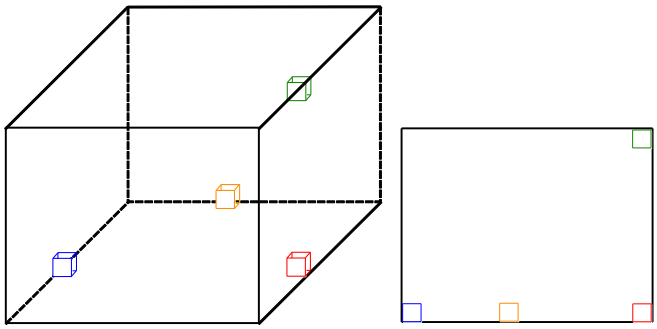


Figure 11 - The representation of a 3D scene in a 2D image.

The question of what a good exposure for an image is, should also be discussed. Historically a simple definition of a good exposure would be an image which contains pixel intensities across the entire dynamic range. However this ignores the contents and 3-D nature of the image scene. A variation in illumination and illumination geometry in conjunction with object geometries can mean that while an image may have a maximal dynamic range, objects in certain spatial locations may be over or underexposed. It must be said that an adequate exposure level further depends upon what the user wishes to visually achieve, it may be that they only aim for the main subject of the image to have adequate contrast.

Global equalisation methods, such as [59] do not try to tackle this greater problem, while local methods do accept this spatial illumination variation to a degree, and treat local neighbourhood regions of pixels independently. Although this can improve things further it is a simplification of the spatial relationship between objects falling on adjacent sensor pixels, with regards to differing depths, but one that is perfectly acceptable.

Nevertheless present research is aiming to work towards camera systems, which can record images as minimally different to what we perceive. HDR imaging algorithms are an example of this, with camera systems aiming to record an image with intensities over a dynamic range closer, for a given scene, to that recorded by the human visual system. Work on colour constancy, the ability of the human visual system to correct for illumination has also been underway for many years [3]. However, even with extended dynamic range capabilities, under/ over-exposure can still occur.

When devising a histogram equalisation method we must also have in mind the aims of histogram equalisation. The main aim is to, of course, equalise the distribution of pixel intensities across the entire available dynamic range, and thus increase the contrast between pixel intensities. Historically it has also gone hand in hand with the equalisation of pixel intensity probabilities. However the image and its contents should be preserved if the image is to have any relation to the original image, and thus the original scene. We have previously mentioned that an image is a 2-D visual representation of a real world scene. Moving on from this a histogram can be described as a further abstract description of this scene. It represents the distribution of pixel intensities across the entire image and as such represents the contents of the scene. A given histogram is thus characteristic of a given scene, and any editing of the distribution will alter the resulting image, increasing the differences between it and the original scene. Virtually all current algorithms alter the shape of the image histogram during equalisation and in doing so create numerous artefacts and inconsistencies in the final image.

When this is carried out with a grayscale image noise artefacts can be introduced to an image, however when altering the three channels of an RGB image separately the colour ratio of the three channels for a given pixel can be changed, altering the actual colour of the object. This in turn will alter the luminance histogram of the image and each object further changing the apparent texture and the perception of the scene for an observer [112-114]. If we wish to have, after equalisation, a maximally accurate representation of the original scene an algorithm should attempt to minimise the occurrences of these. However we cannot state that solely preserving the shape of a histogram will preserve the contents of an image. The rank order of pixel intensities, in a given channel, is also characteristic of the contents. Thus in order to completely preserve the contents of an image both the histogram shape and rank order of intensities must be maintained, we will refer to this a monotonic transform.

Caselles et al. [65] produced an example of an algorithm that preserves the shape of the histogram across local areas, while altering the rank order of the pixel intensities, changing the image contents. When dealing with single channel images on top of our previous assumptions this works well. However, when introducing this method to multi-channel RGB images it can be imagined that independent channel histogram modification in this manner will result in slight colour alteration if differences between the channels exist. However in practice these alterations will be much smaller in size,

more natural looking and more acceptable than those introduced by other methods, minimising the key aim of content alteration during equalisation.

3.1.2 Pixel Intensity Monotonicity

The monotonicity of pixel intensities in an image was noted previously within Chapter 2 when discussing the theory of gamut based methods and different implementations. In the previous section the effect of exposure and illumination upon a histogram were noted. It can be recognised that if the effect of the illumination is large enough the order, and thus monotonicity of the pixel intensities, under one illumination will differ to those under another.

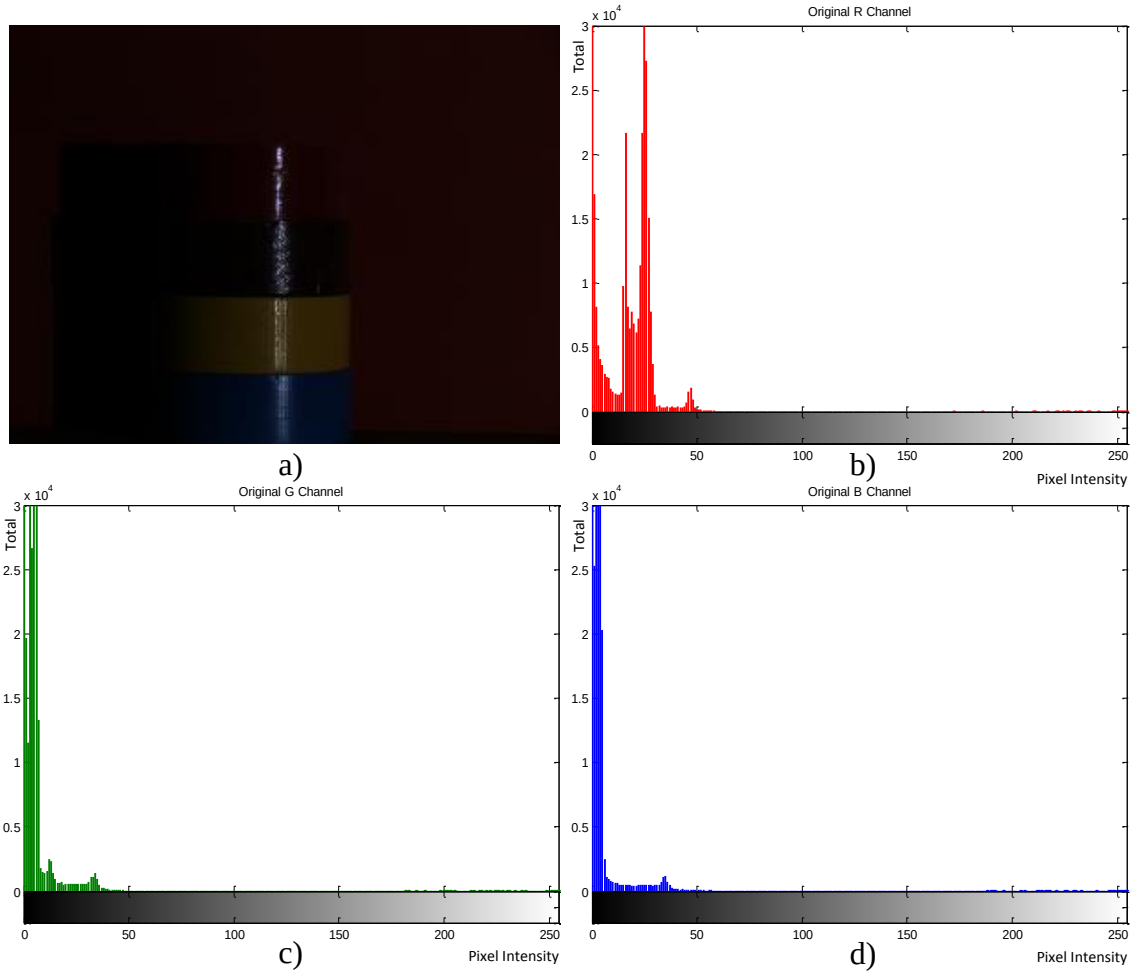


Figure 12 - An example of a) an underexposed image, b) the R channel, c) G channel, and d) B channel histograms.

This links histogram equalisation to the colour constancy problem in image capture. The field of colour constancy attempts to counter the effects of illumination on an image by attempting to either find an illumination invariant representation of a scene or estimate the illumination. Histogram equalisation can thus be considered a special case of the

colour constancy problem, in that the image was captured under a very low, or high, intensity and unknown illumination.

Previous sections have outlined the importance of histogram shape and rank intensity, another way to describe this would be that through contrast enhancement we wish to maximise the level of mutual information between the original and resulting image. This will ensure that the resulting image is a true representation of the original image.

Generally research introducing histogram equalisation methods the visual results of the novel algorithm are compared to those from existing methods. This gives a perceptual comparison of the results, however if we are to use the definition that the preservation of intensity rank order is of the utmost importance a metric which compares the histogram of the original and that of the resulting image needs to be introduced. Visual comparisons have been used as most image quality metrics, such as PSNR and SSIM [116], have been designed to compare a resulting image to a known ground truth version. In the enhancement and equalisation area we have no ground truth known data, as an image to be processed is treated as unusable in this way. However we have previously stated the case for using the histogram of an image as a representation of the scene, and the importance of histogram shape preservation.

3.2 Original Monotonic Transform

The previous section has outlined the basis for the monotonic transform [69], however it should be described in greater detail regarding its derivation and the scene assumptions it is designed to operate under. This transform assumes that there is a single uniform illumination in the scene and that all scene contents are on the same 2-D plane in space. It has also been designed to work with both single (grayscale) and multi-channel (RGB) images as the proposed transform is designed to be applied independently to each channel in an image.

The implementation of the monotonic transform is a simple re-sampling of an image by normalising its ranking orders. The process first calculates the intensity order for all pixels in an image in each of the different channels. The new pixel values are re-scaled, according to the intensity order, between 0-255 (assuming an 8 bit image). This results in normalised images that are invariant to lighting and dependent only upon the image pixel intensity order. A description of this, enabling a speedy and efficient implementation, of this is provided below;

1. Store, in ascending (with rank/ position $n_{R,G,B}$) order, and count the number of unique intensities ($u_{R,G,B}$) present in an RGB image ($I_{R,G,B}$).
2. Calculate the step position increment ($g_{R,G,B}$) required to equally distribute the existing intensities across the entire dynamic range.
3. Transform a given pixel's intensity in the original image ($I_{R,G,B}(x,y)$) by multiplying it's rank order position by the calculated step position increment $g_{R,G,B}$.

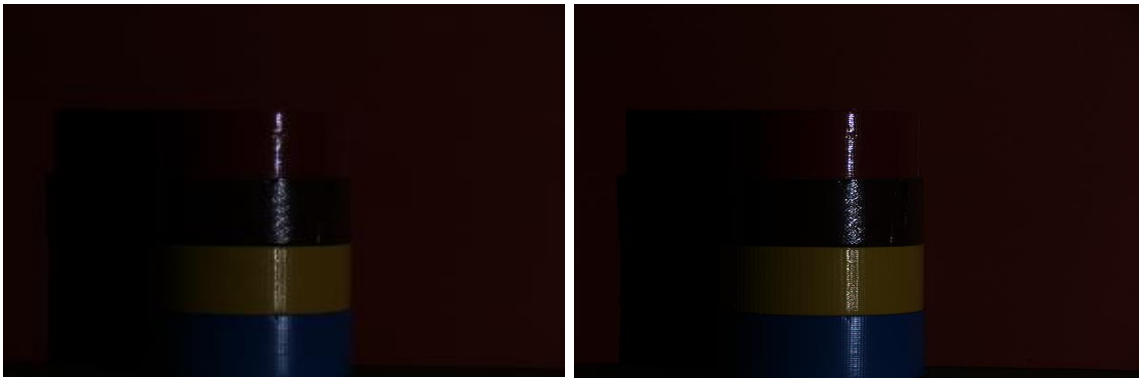
Mathematically, this can be described as (again assuming an 8 bit image, with a maximum pixel intensity of 255);

$$g_{R,G,B} = \frac{255}{(u_{R,G,B}-1)} \tag{24}$$

Leading to the translation of pixel $I(x,y)$, with an intensity value having a rank order position n;

$$I(x,y)'_{R,G,B} = n(x,y)_{R,G,B} \times g_{R,G,B} \tag{25}$$

Where $I'_{R,G,B}(x,y)$ is the monotonically transformed pixel intensity of $I_{R,G,B}(x,y)$. Examples of the monotonic transform on an image and its component channels are illustrated in Figure 13 from [104].



a) b)

Figure 13 - The a) originally under-exposed image, and b) the result of the monotonic transform.

However the main weakness of the monotonic transform can be deduced from the described method. As previously described, the monotonic transform will equally distribute the intensity components of a channel across the entire dynamic range. As this is the only operation being undertaken it does mean that if an image already has a number of unique intensities approaching the maximum possible number its effects will

be minimal. With the transform doing nothing, when the image has the maximum number of possible intensities, 256 for an 8 bit image. However this weakness also results in one of the strengths of the algorithm. This simple transformation of pixel intensities with a preservation of rank order results in a more natural looking image with no large introduced components. Further to this the performance of the monotonic transform should thus outperform other methods when dealing with under and over exposed images.

3.3 Luma based Monotonic transform

The previous section mentioned briefly the main weakness of the monotonic transform when dealing with a 'well' exposed image. Whilst an image containing a large number of components at every single possible intensity may provide a good image, this performance problem will exist even if only 1 pixel in an image contains a specific value. Thus an image which is noisy or has a heavily skewed histogram will lead to degradation in transform performance relative to existing algorithms.

Heavily skewed histograms can be caused by a combination of either under or over exposure in specific scenes. Material textures and lighting geometries thus also contribute to the list of causes in conjunction with an incorrect exposure level. Figure 14 from [104] shows an example of the skewed luminance histogram caused by a combination of underexposure and scene properties.

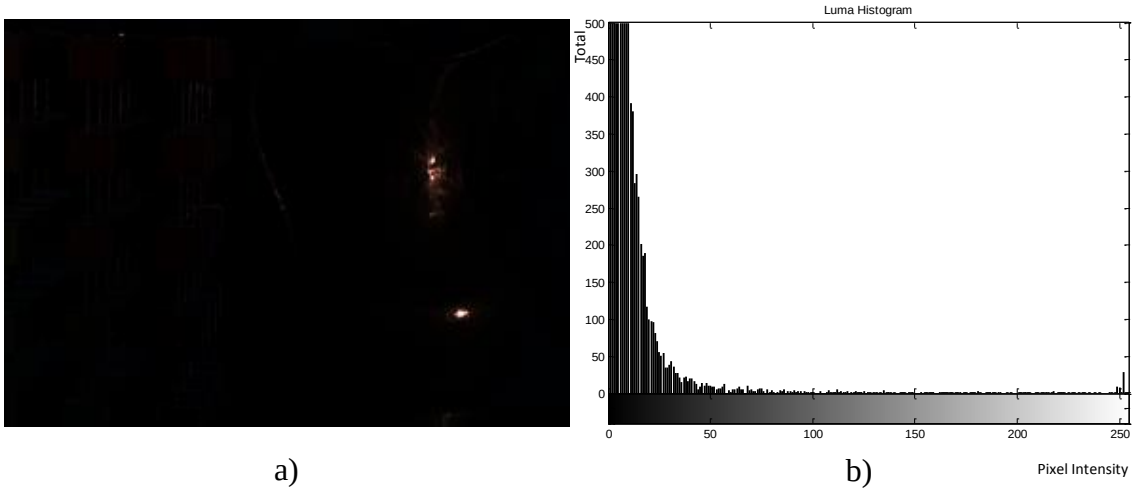


Figure 14 - An example of an a) highly luma skewed image, and b) the luma channel of the image.

As Figure14b shows the R channel is heavily skewed towards the lower intensities. However it can also be seen that while the vast majority of pixel intensities are in the lower region the histogram, and thus image, exhibits a small tail of pixels with

intensities up to a value of 255. These originate from the specular regions on the image, with the range of upper values being directly caused by the shape feature of the tape in conjunction with the lighting geometry and multiple surface reflectance functions. The result of the standard monotonic transform on

Figure 14a, comparing the resulting histogram of the red channels is shown in Figure 15 below.

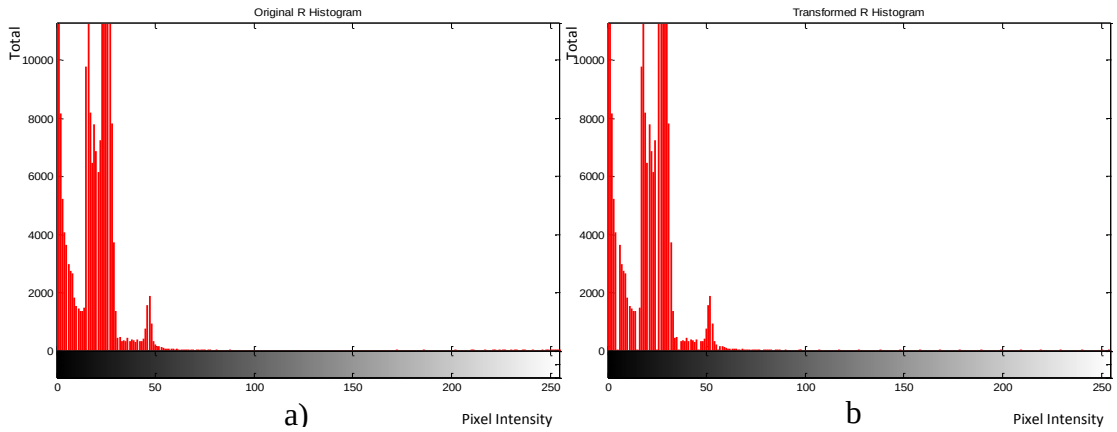


Figure. 15 - Comparing the red channel histograms for a) the original image in Figure 13a, and b) the monotonically transformed image in Figure 13b.

As can be seen the tail of intensities causes a decrease in performance with the transform, whereas other methods would not be affected. The lack of image modification is apparent when viewing Figure 16 showing the difference between the original and monotonically transformed images, with the most visible difference occurring with the specular highlights in the indicated region. A slight expansion in the main histogram components can be noted however, but the visual effect of this is minimal.

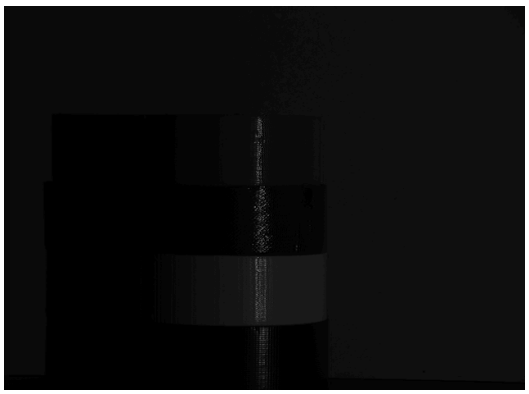


Figure 16 - Pixel differences introduced into Figure 13a by the monotonic transform for the Red channel boosted by a factor of 5.

The challenge posed by such image statistics to the performance of the monotonic transform can also be seen in a number of other images from [104];

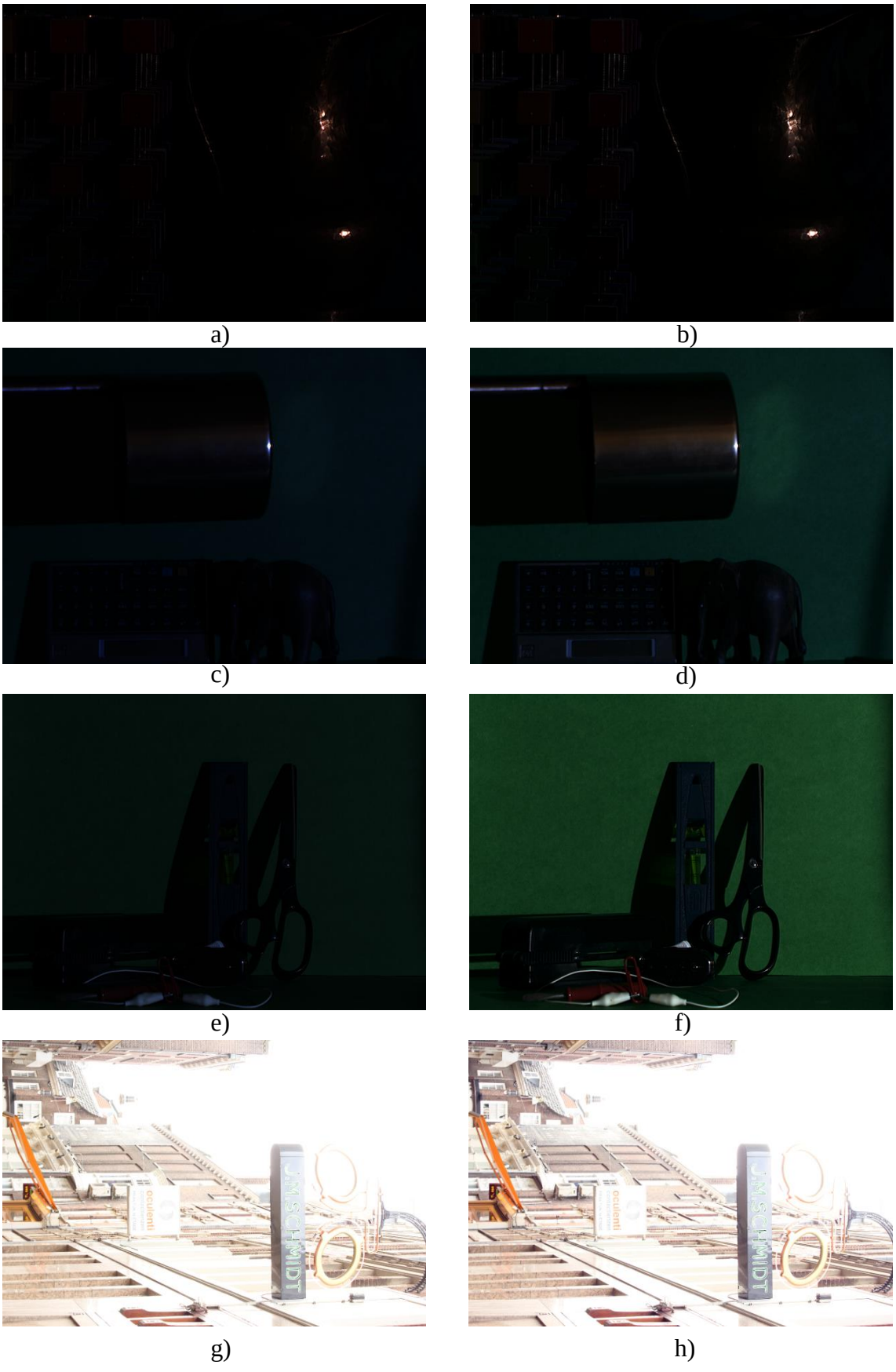


Figure 17 - Examples of original images (a, c, e and g) the respective results of the MT algorithm (b, d, f and h).

3.3.1 *Human perception of gloss*

The perception of gloss by the human visual and visual perception systems has long been an area of interest for researchers. Recently a great deal of work has been carried out to investigate if certain specific luminance statistics play a role in the perception of gloss. Sharan and Motoyoshi both [113, 114] reported that luminance skew in images was an indicator of gloss perceptions when tested with human subjects. They utilised a rendered version of an image in which the mean luminance vale of the image was kept constant, but the skewness of the image was varied through the use of linear and non-linear reflectance values. They also utilised a database of images taken of materials with different surface properties. However both Anderson and Kim [112, 117] put forward the idea that the statistic of skewness cannot be the cue that the human visual system uses to assess surface glossiness. They investigated the same theory by utilising the same set of basic images, however in order to see if histogram skewness alone was a cue they altered the locations of the specular highlights. To achieve this they rotated the locations of the highlights by set degree value offsets and then presented them to observers. This operation was carried out without any consideration towards the resulting highlights existing in natural locations.

If the results from both Sharan and Motoyoshi [113, 114] were correct, in that luminance skew aids human gloss perception then it would be expected that gloss perception would be preserved. However it was found that gloss perception was not preserved. Anderson [112] puts forward the idea that skewness on its own cannot represent glossiness and moves to put forward evidence that local surface information in conjunction with highlights aids surface gloss perception [117]. Anderson [112] states that an image histogram cannot represent such a higher order feature, but does represent the geometric distribution of luminances across the surface. The relationship between this and the skewness is what is of importance to human gloss perception.

As previously have put forward, and what was stated in [69] and backed up within [118], the distribution of intensities is a characteristic of the scene or surface in the image even across changes in illumination. These agree with Anderson [112], however they also agree with the findings of [113]. Whilst at first glance the two studies appear to disagree their findings can lead to another explanation. Anderson [112] places importance upon a joint consistent relationship between local spatial luminances and the overall luminance skew, in comparison to [113]. In the images tested within [113] these were consistent, and thus also natural. However the images in [112] lacked this

consistency, and this was of course on purpose. However the question to be asked is whether surfaces such as the ones tested within [112] existed in the natural world. As such skewness in conjunction with local luminance fluctuations around specular highlights always occur together, the validity of separating the two then comes into question. If the two are never separated under real conditions then luminance skewness, whilst not an absolute definitive quantitative cue, can be harnessed to roughly estimate the glossiness of a surface, or surfaces.

The calculation of glossiness in [113, 114] as a global image descriptor matches how images are effected by over-/ under-exposure. As such images which exhibit rapid changes in pixel intensity, such as with specular highlights, have a characteristically larger skewness value. Whereas images containing gradual increase in luminance value, such as images of shaded object, do not.

3.3.2 *Incorporating perceptual findings into an algorithm*

Second to this the degree of skewness can also illustrate the degree to which an image has been under/ over-exposed. We previously noted that images exhibiting slightly increased, absolute, luma skewness values caused problems for the monotonic transform. Analysis of an images' luminance or luma counterpart also has another advantage. Previously it was mentioned that a challenge of colour image equalisation/ enhancement was to minimise the introduction of colour casts/ abnormal colours during processing. Methods which treat each channel separately for contrast enhancement disregard the relationship between the channels during under and over exposure. Colour normalisation and colour constancy methods can treat each channel separately on the basis on von Kries type chromatic adaptation assuming a good level of exposure. However a challenge is presented if we are trying both contrast enhancement and colour normalise an image at the same time. The luminance/ luma channel is related to all 3 channels in an RGB image, and as such if we adjust enhancement solely with respect to this then we will be adjusting each of the three channels by the same degree. This may mean that the colour cast in a highly casted image may not be totally removed, but contrast will be improved. On from this utilisation of the monotonic transform will preserve the kurtosis of both the luminance and individual colour channels.

Within [113] skewness was calculated using the following formulae;

$$meanL = \left(\frac{\left(\sum_{m}^{x=1}\sum_{n}^{y=1} I_L(x,y)\right)}{N}\right) \quad (26)$$

$$\sigma = \sqrt{\frac{\sum_{m}^{x=1}\sum_{n}^{y=1}(I_L(x,y)-meanL)^2}{N}} \quad (27)$$

$$Sk = \left(\frac{\left(\sum_{m}^{x=1}\sum_{n}^{y=1}(I_L(x,y)-\sigma)^3\right)}{(\sigma N)^3}\right) \quad (28)$$

Where *meanL* is the mean luma value of the image, σ is the standard deviation of the luma value and *Sk* is the skewness value.

However investigations into the variation of luminance skewness and exposure level were first required. In order to investigate this images across [58] were analysed for their luminance statistics. These statistics included mean luminance, skewness, standard deviation and the pixel count at most populated luminance intensity.

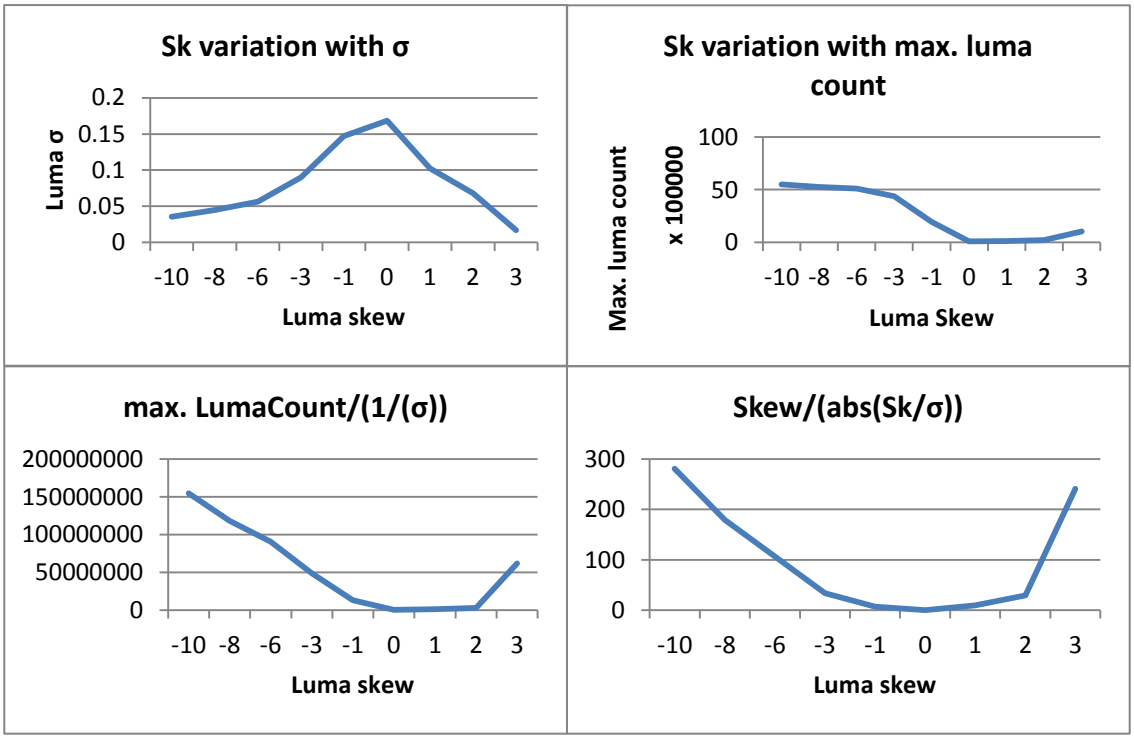


Figure 18 - An overview of the statistical relationships between different luma statistical measures

From the graphs above a method to counteract this was devised using the luminance statistics (skew, skewness, standard deviation and mean) of the image to gauge the skewness and skewness characteristics. If this is over a set threshold an edited version of the monotonic transform will be used.

There are two possible ways an image's histogram may be skewed, towards either the maximum or minimum intensity values. This depends upon over/ under exposure and

the colour of any illumination in the scene. Any algorithm should be able to deal with both of these possibilities and adjust the modification appropriately. As stated previously the level of luminance skewness will be compared to a set threshold to decide whether to utilise this skewness monotonic variant. The differing luminance and thus skewness characteristics of under and over-exposed images means that two different thresholds are required. The threshold for an image exhibiting negative skewness (over exposed) will be much smaller (in absolute terms) than the threshold for an under exposed image. With our algorithm here using values of -1 (*negSkT*) and 5 (*posSkT*), these have been chosen due to experimentation with a number of images, but can be altered. This much smaller threshold for negatively skewed images is related to the non-linear nature of sensor sensitivity as values approach saturation. As an imaging sensor approaches saturation pixels which would exhibit slightly different pixel values at a lower exposure level are grouped together into the same pixel intensity. This leads to a very lob sided luminance histogram, leading to a smaller standard deviation and thus skewness value as the difference between the mean luminance and each pixel values decrease rapidly.

3.3.3 *Algorithm operation*

The skewness detection monotonic transform works in much the same way as the monotonic transform. However if a high absolute level of luminance skewness is calculated then all pixel intensities with a total count, in the histogram, less than a set threshold (*Threshold*), dependent upon the luminance skewness (*sK*). This is due to the relationship between skewness value and total number of existing intensities within the image. As skewness increases in magnitude we note, from Figure 18, an increase in the total number of small cumulative total intensities that exist within the tail of the histogram distribution. As such it is the intensities within this tail that cause the original monotonic transform to fail. An important feature of this algorithm is that by using the count of unique intensities below the threshold value we are scaling the enhancement by the skewness of the histogram.

Having counted the number of unique pixel intensities below the calculated threshold (*Threshold*), depending upon the direction of skewness, our procedure again differs depending upon the direction of the skewness. We now use this number to preserve the most important components of the histogram, ignoring if the pixel total for a given intensity is above the previous calculated channel threshold. Put simply, if a positive

skewness is exhibited (under exposure) the most common pixel intensities (those required to preserve the shape of the histogram) will be closer to 0. As such if an intensity count (*u*) of 74 resulted from the threshold operation we give ranks to the first (starting at intensity 0 with a rank of 1) 74 non-zero count intensity values (intensities at which at least 1 pixel exists) in the image. With all other existing non-ranked intensities being set to a value of 255, preserving the rank order of the main components.

For images exhibiting a negative skewness (over exposure) this operation is reversed in its direction, due to the fact that the most important intensities are located closer to 255. Unlike with the original transform and when dealing with a positively skewed image the value of *u* here for all channels is determined by using the maximum calculated *u* from the R, G and B channels ($u_{R,G,B}$). The same value of *u* is used here for all channels due to the non-linearity of sensor channel sensitivities as they approach saturation. In practice this means that if we used an individual value of *u* for each channel the resulting image would have a strong colour cast determined by the sensor channel with the most linear response close to saturation. As such using the maximal value of $u_{R,G,B}$ results in an image with a minimal colour cast. We then proceed to give ranks to non-zero count intensities from 255 downwards (with the first intensity having a rank *n* equal to *u*, proceeding to count down to 1). This then means that any other existing non-ranked intensities will be set to a value of 0. This can be seen to be opposite in method to the procedure for positive skewness values. The transform is then carried out using these rank orders.

The algorithm can be described as;

1. The image (I), is converted into a luminance version of itself (I_L), if it is a colour image, using the Rec. 709 standard values.
2. The luminance skewness (*Sk*) of this image is calculated.
3. Compare *Sk* to *negSkT* and *posSkT*, setting *skewnessFlag* to 1 if *Sk* <= *negSkT* or *lumSkewness* >= *posSkT*.
4. Calculate the histogram threshold for each channel (*Threshold*), which is equal to the largest pixel count (*max LumaCount*) for a single bin in the histogram divided by the appropriate Threshold value.
5. If…
    a. *Sk > 0*;
        i. Count the number of qualifying unique intensities (*u*) in each

channel (those that are smaller than Thresh), giving unique channel positions/ ranks *n* to all non-zero count intensities starting from an intensity of 0.

ii. Set all non-ranked, non-zero intensities to the maximum rank value (*u*), resulting in a value of 255 after transformation.

b. *Sk* < 0;

i. Count the number of qualifying unique intensities ($u_{R,G,B}$) in each channel (those that are smaller than *Thresh*). Set *u* to *max*($u_{R,G,B}$) giving position/ rank *n* to all non-zero count intensities starting from an intensity of 255 with a maximal rank value, working downwards.

ii. Set all non-ranked, non-zero intensities to the minimum rank of 1, resulting in a value of 0 after transformation.

6. Using these new rank values continue from Step 2 of the monotonic transform.

The following equations show steps 1-6 in detail;

Calculating the luminance image, if an image with more than one channel is being used, where *x* and *y* are pixel locations;

$$I_L = (0.2126 \times I_R) + (0.7152 \times I_G) + (0.0722 \times I_B) \quad (29)$$

Using this value the algorithm decides whether to use the normal or skewness variants of the monotonic transform.

The threshold is then calculated for each channel depending upon the direction of *Sk*;

For positive *Sk*;

$$Threshold = \frac{maxLumaCount}{\left(\frac{|Sk|}{\sigma}\right)} \quad (30)$$

For a negative *Sk*;

$$Threshold = \frac{\left(\frac{maxLumaCount}{100}\right)}{\left(\frac{|Sk|}{\sigma}\right)} \quad (31)$$

If Sk==0 and no more than 244 unique intensities ($u_{R,G,B}$) exist in each channel:

$$Threshold = \frac{maxLumaCount}{\left(\frac{1}{\sigma}\right)} = maxLumaCount \times \sigma \quad (32)$$

If *lumSkewness* is negative (*lumSkewness* < *0*);

$$\boldsymbol{u_{R,G,B}} = \left((\boldsymbol{u_R} \vee \boldsymbol{u_G}) \vee \boldsymbol{u_B}\right) \quad (33)$$

We then proceed to calculate $g_{R,G,B}$ in the same manner as Eq. (1), moving on to rank and transform the intensities as laid out in Eq. (2). Except only qualifying intensities are normalised, and non-qualifying intensities are left at their original value.

The result produced by this modified algorithm for Figure 13 is shown in Figure 19. The performance difference when compared to the standard monotonic transform (Figure 16) can immediately be noted.

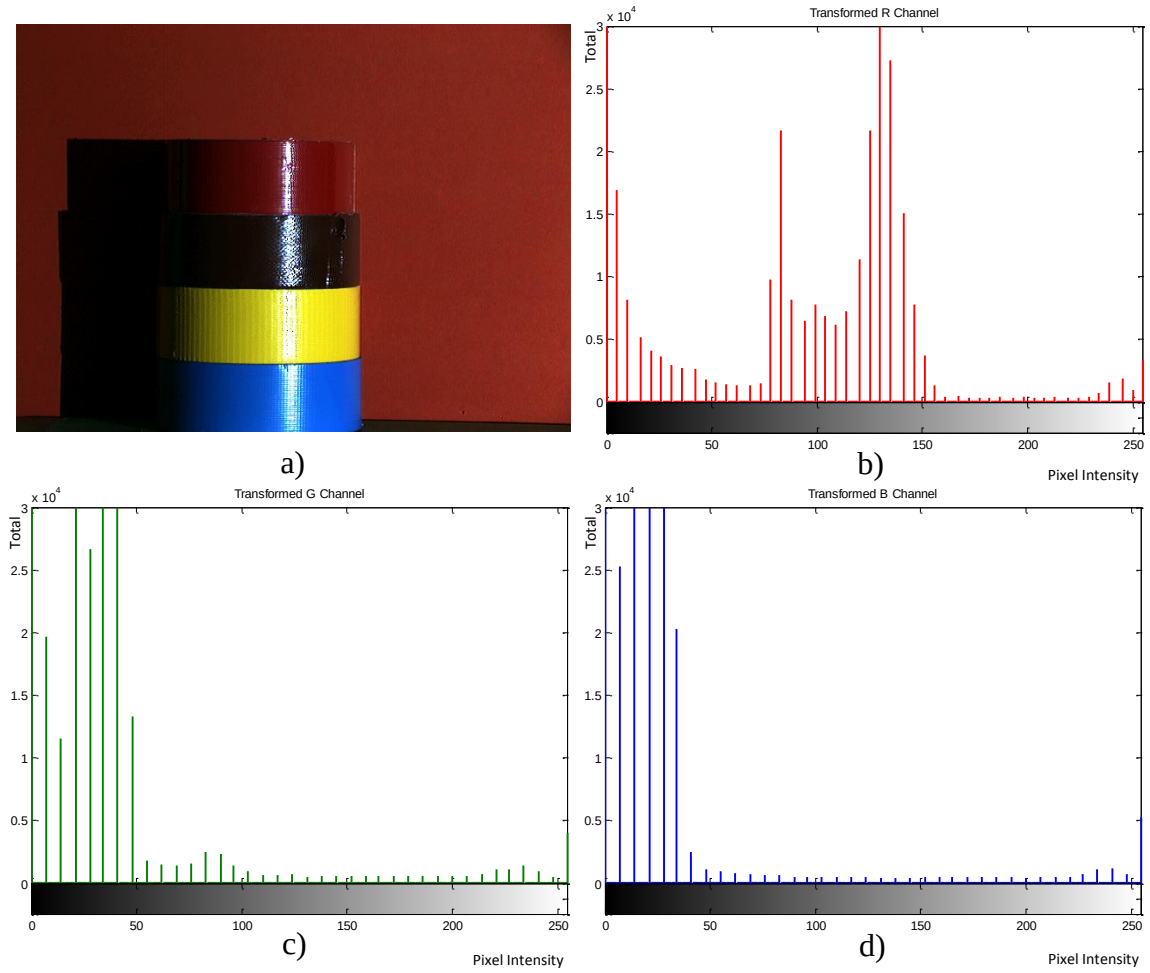
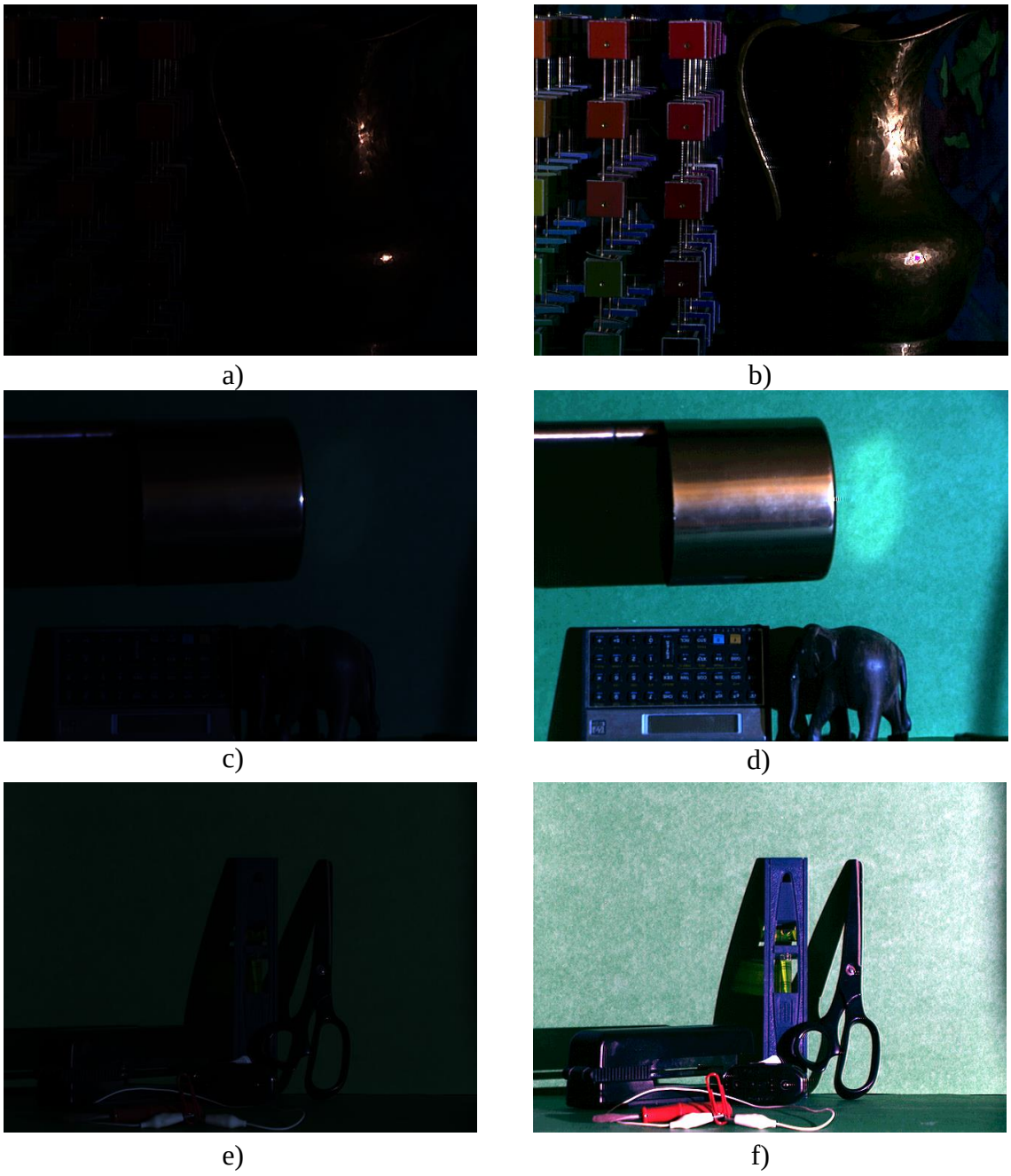


Figure 19 - The results of the luma skewness based monotonic transform a) resulting image, and b) resulting R channel, c) G channel, and d) B channel histograms.

3.4 Visual result comparison

When comparing results only a visual comparison will be given of the best results from the previously mentioned algorithms. However a review of the algorithmic results for each image will be given. A number of these images have been taken from [104], whilst others are images taken by the Author.



a) b) c) d) e) f)

Figure 20 - Examples of original images (a, c and e) the respective results of the luma skewness MT algorithm (b, d, f).

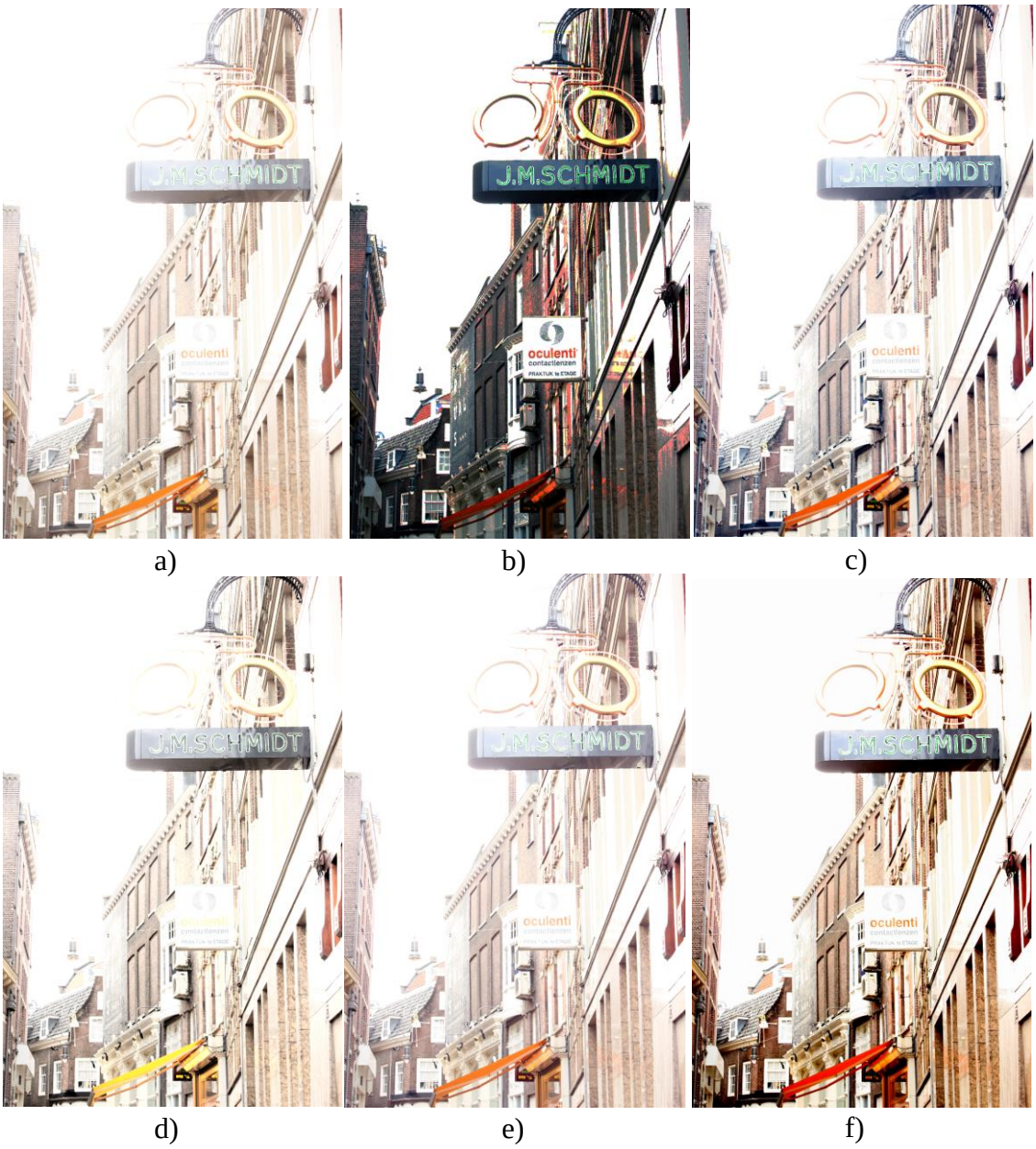


Figure 21 - An example of processing a) an over-exposed image, and the results of b) histogram equalisation, c) SCB, d) intensity preserving SCB, e) the MT and f) the luma skewness based MT.

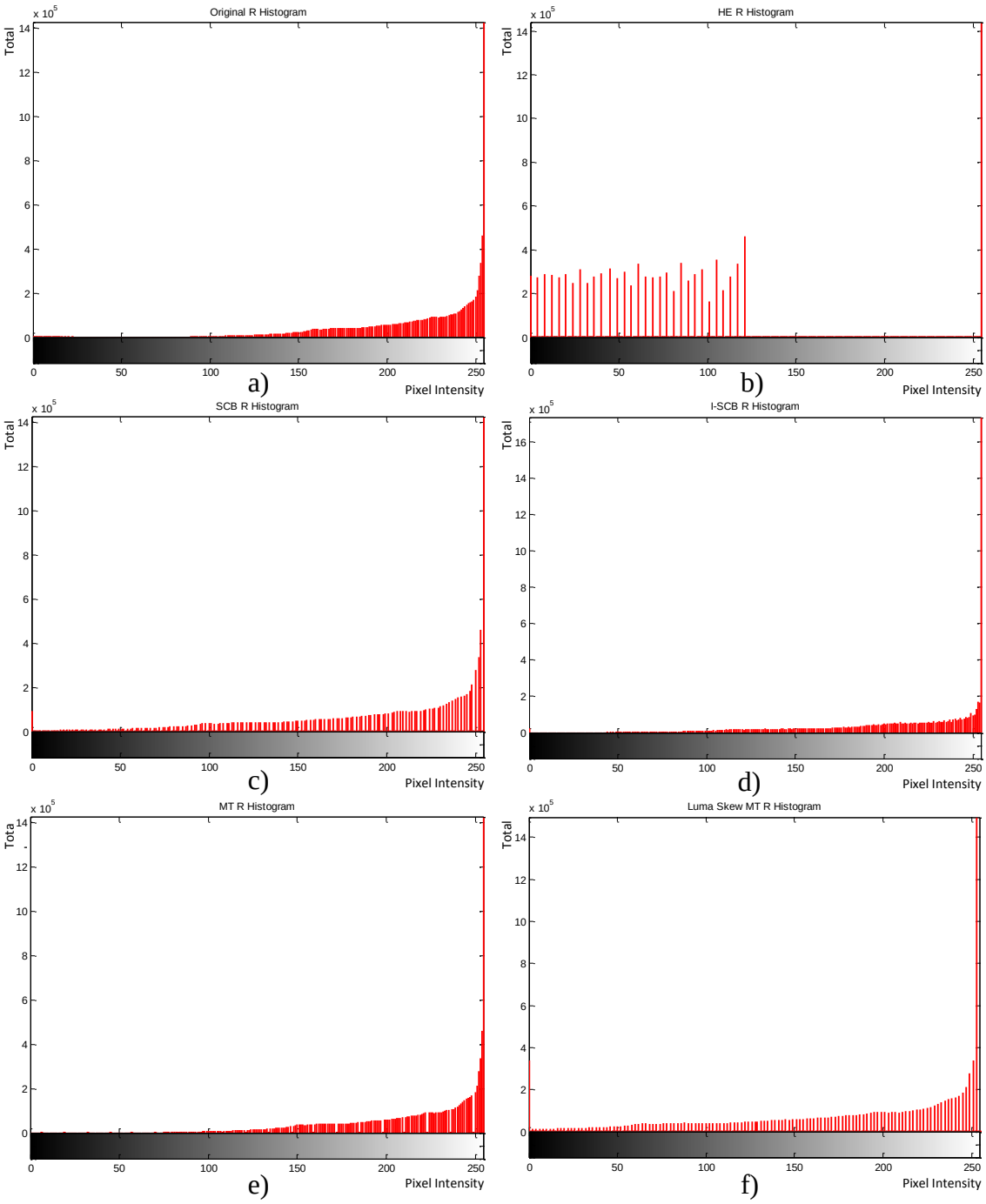


Figure 22 - Comparing the R channels of a) the original, b) the histogram equalised, c) SCB, d) intensity preserving SCB, e) MT and f) luma skewness MT images.

Figure 21 shows a comparison between a number of different algorithms. It can immediately be seen that the results from the SCB algorithm (Figure 21c) and the MT (Figure 21e) are comparable, with the SCB algorithm producing an image with marginally better contrast, due to the greater contrast stretching above and below the set minimum and maximum thresholds. Both of these algorithms perform far better than the other versions. Figure 22 illustrates the effects of these algorithms on the red channel of the image. Figure 22c and Figure 22e show the effects of SCB and MT on the channel,

and these are the only algorithms to preserve the shape of the original histogram. This in turn shows why they are the best contrast enhanced representations of the original images. SCB here provides the slightly better results, however in a severely underexposed scene this algorithm will be outperformed by the MT. As some of these algorithms, such as histogram equalisation were not designed for use with colour images it is more appropriate to compare the performance of our method on these types of image.

Again this is when dealing with under exposed images, however Figure 21 shows the results of these algorithms on an over-exposed image. When attempting to normalise over-exposed images we do not expect to get as good results compared to underexposed images. This is because of the reduction in recorded data due to sensor saturation and sensor non-linearity close to saturation. However we do still wish to achieve a reasonable increase in contrast, whilst again preserving the information that has been recorded in the scene; the shape of the histogram. Any other actions to edit the shape of the histogram would be tantamount to adding data to the scene resulting in a reduction in the mutual information between the original and edited images. It would then be incorrect to say that the edited image is a perfect representation of the original image, with the mutual information between the two images decreasing with every histogram shape changing process. Here it can be seen that while the MT does perform well, a few other algorithms also do reasonably. However, the fact that the MT does not introduce artefacts here is a major plus point.

However the skew monotonic transform does not always outperform other methods, as mentioned previously the method is designed to work on images where the luminance skew causes problems for the original monotonic transform. What happens if we attempt to utilise our novel method on images where the monotonic method already performs well?

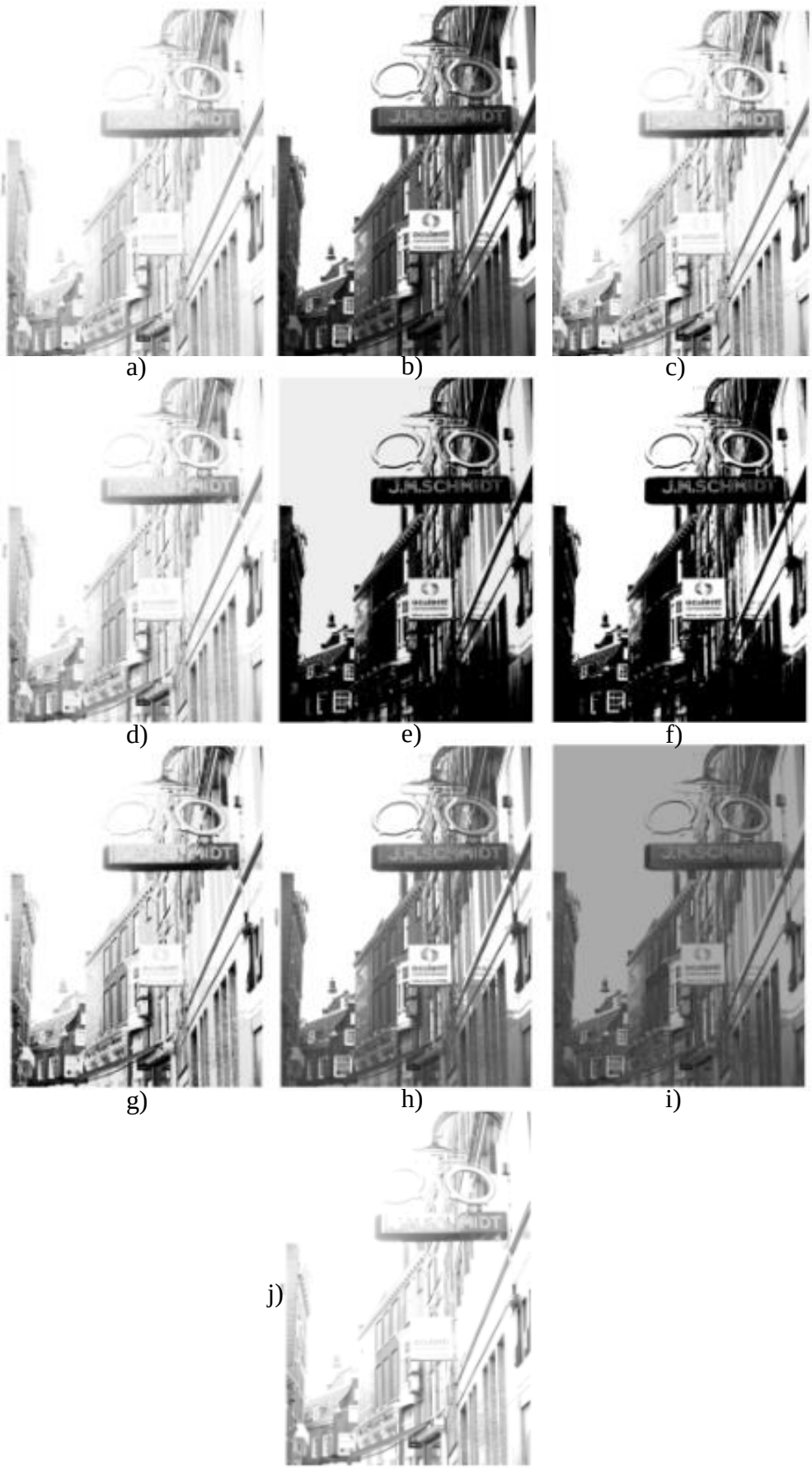


Figure. 23 - Comparing the results of different contrast enhancement techniques on original grayscale image a), using b) histogram equalisation, c) SCB, d) the MT, e) luma skewness MT, f) weighted threshold histogram equalisation, g) BBHE, h)

MMBEBHE, i) limited slope histogram equalisation and j) local colour histogram equalisation.

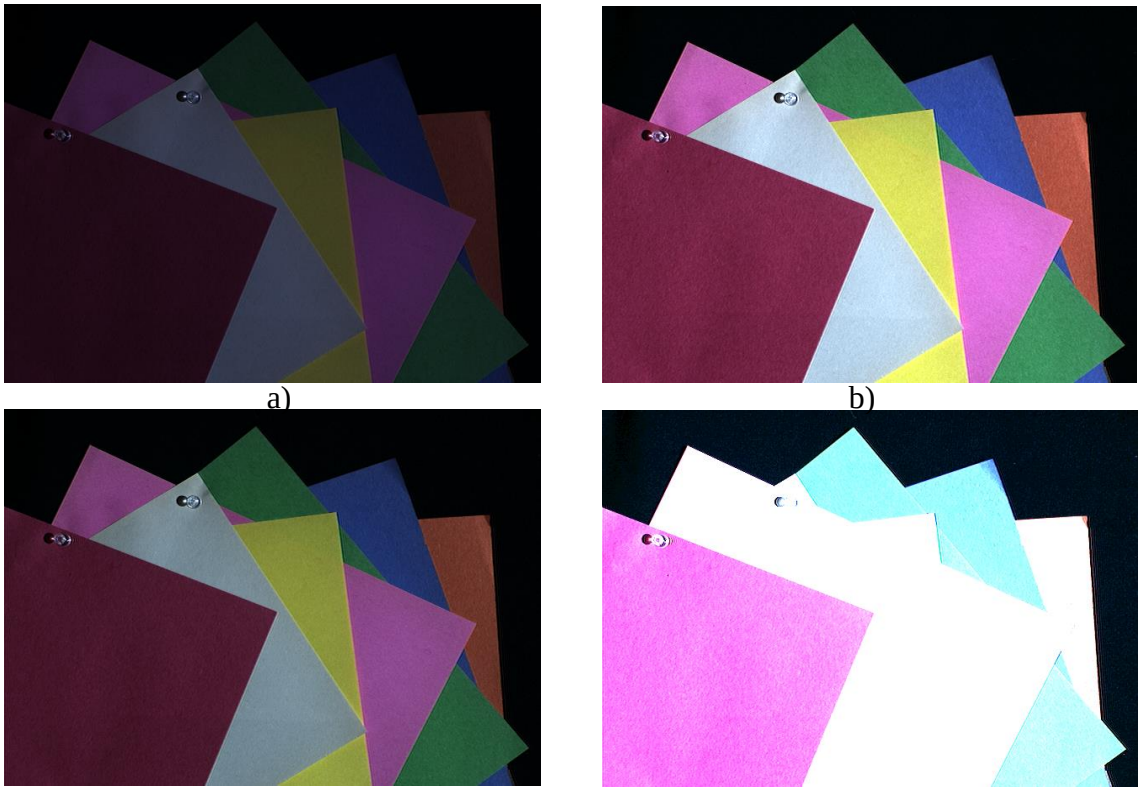
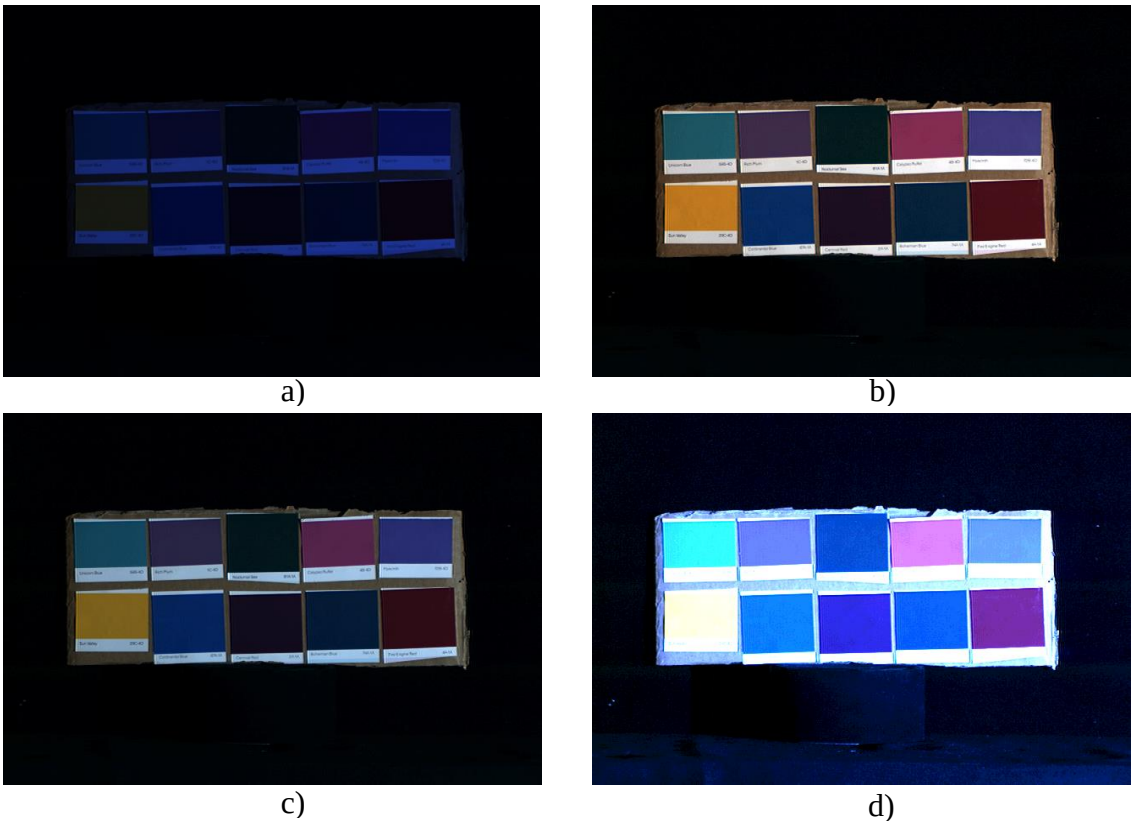


Figure 24 - Comparing the results of processing a) an image, using b) SCB, c) MT, and d) luma skew MT

Figure 24 and Figure 25 both show instances where the performance of the novel luma skewness MT algorithm is outperformed by both the original MT and the SCB method. Both of the original images here exhibited small luma skewness values of 1 and 3 respectively, and similar relative performance differences have also been noted with images of similarly small luma skewness values. In these instances care must be taken when utilising the luma skewness method as it can result in the over boosting of the colour channels due to the automatic hard coded threshold nature of the algorithm modelled from analysing image datasets.



a) b) c) d)

Figure 25 - Comparing the results of processing a) an image, using b) SCB, c) MT, and d) luma skew MT

3.5 Summary

The work presented within this chapter has illustrated how a general RGB space based algorithm can be improved through the utilisation of findings from psychophysical research into the perception of images. Whilst it cannot be said that the algorithm replicates the human visual system it does incorporate specific psychophysical aspects of it. It also utilises the luminance contents of the image to guide the normalisation process. In this it manages to achieve the three main aims and objectives of this thesis laid out within Chapter 2.

The developed method also manages to match the performance of state of the art normalisation algorithms Furthermore it achieves the specific chapter aim of preserving scene contents whilst doing so. With the basis of this chapter being the use of non-perceptually or physiologically correct image colour space.

Chapter 4. Illumination estimation and normalisation through skin chromaticity gamuts

4.1 Introduction

Previously it was noted that the vast majority of algorithms have attempted to replicate colour constancy whilst ignoring the specific contents of images. Previous methods have been firmly rooted within the computer vision area, by utilising machine colour spaces such as RGB with no direct link to human colour perception. Others have attempted to tackle this problem by using colour spaces with a degree of perceptual relevance, such as La*b*. However the perceptual uniformity of this space is preserved only under a specific illumination, CIE D65. The difference in chromatic representation and resolution of machine colour spaces in comparison to those of the human visual system are a major hurdle in the effort to replicate colour constancy. These computer vision colour spaces have been determined by the design of imaging sensors. However hyperspectral image acquisition allows us to side step this problem by capturing the ground truth spectral irradiances within a scene. This then allows precise cone responses to be calculated, and as such design algorithms which can utilise visual data at the same level as the H.V.S.

Colour constancy is a process within the human visual system, and as such, to replicate it, we should try to match the human visual system in terms of chromatic resolution at the input level. Furthermore the exact mechanisms the human visual system utilises to perform colour constancy are unknown. It is likely, though, that the objects and contents of scenes are generally of some aid. As stated previously, the use of specific perceived contents has been largely ignored historically by the majority of algorithms. As such we investigated the development of a method that could utilise common scene contents to perform a reliable colour balancing on images. Research following this same line of thought has been carried out previously [84, 119]. Bianco et al. [119] utilises human skin in sRGB space as the basis for illumination. It first attempts to detect skin pixels within an image and then compares the average of these skin patches to a reference skin chromaticity value, whilst also comparing the gamut of these skin pixels to a predefined canonical skin gamut. Using this it then estimates the sRGB colour of the illumination. This method also allows spatial variations in illumination to be estimated and corrected.

Whilst utilising skin in the same manner as our method, this work is heavily rooted in non-perceptual and non-physiological colour space. It does produce good results, however it cannot be considered to investigate the use of skin for illumination estimation as the work within this chapter has done. The motivation and approach of the work within this chapter has solely approached from the angle of investigating and replicating signals within the human visual system.

It seems evident that a large proportion of photographs contain people. Apart from the common nature of skin in images, skin can also be considered a known surface and object to the photographer. The research interest of the effect of known object colours to colour constancy in human subjects has also been previously noted [17, 18], and even with skin tones [120]. With these motivations in mind we set out to see if, using a physiologically correct colour space, human skin can be used as the basis for illumination estimation.

4.2 Foundations of the investigation

The following section will detail the design and collection of the image and illumination dataset used as the basis of the investigation into the variation of skin chromaticity gamuts across different illuminations.

4.2.1 *Data collection*

In order to test whether human skin could be used as a cue to the illumination a dataset was created. This covered a wide range of commonly and un-commonly encountered illuminations and a wide range of skin types and is novel in itself. Ground truth data capture was provided through the use of a spectro-radiometrically calibrated Specim V10E hyperspectral camera. This dataset, consisting of hyperspectral images of 19 different individuals' hands, was captured along with the spectral details of 39 different illuminations. These illuminations were provided by a tuneable LED illuminator and integrating sphere, as described in [29].

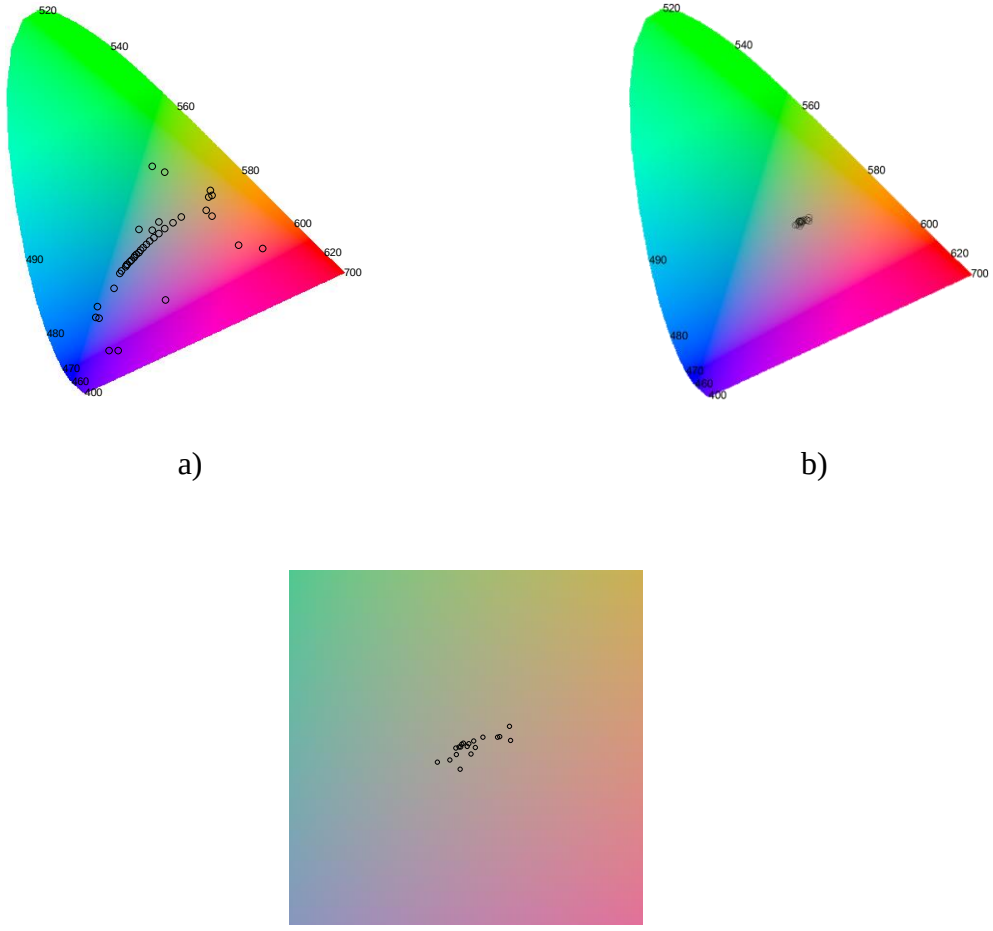
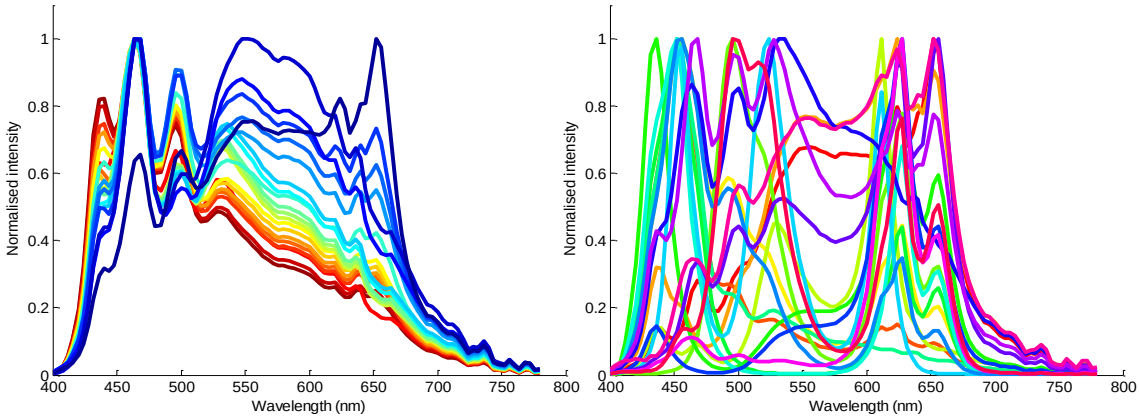


Figure 26 –The CIE xy co-ordinates of a) the database illuminations, b) the 19 individual average hand chromaticities, under D65, collapsed across the luminance plane, and c) a closer look at the chromaticity variation between these subjects.

In this set of illuminations 20 were metamers of the CIE daylight standard illuminations, varying from D40 – D250. The metamers chosen for use in the dataset for each of these daylight illuminations were the smoothest spectra produced by the LED system [29]. An overview of the metameric illumination matches for the daylight loci, across the aforementioned range, are shown below. The rise in prominence of components in the upper wavelength region (500nm onwards) showcases the changes as correlated colour temperature (CCT) increases from 4000K to 25000K.



a) b)

Figure 27 - Spectra of a) CIE standard D type illuminations, and b) non-daylight illuminations, used to construct database of hand chromaticity gamuts.

In addition to this range of daylight illuminations, 19 non-daylight illuminations were added to the database, these being either created *de novo* or as metamers of existing commercial illumination sources. This included a metamer of CIE Illuminant A and CIE fluorescent standard F12. Other illuminations were far more narrowband and were intended to provide extreme illuminations that would not be normally encountered in real world scenarios.

The 19 subjects were of varied ethnic origins from worldwide locales, and their average skin chromaticities varied correspondingly as shown in Figure 26b and Table 1;

	Mean Yxy under D65		
Subject #	Y	x	y
1	50.97	0.3582	0.3701
2	27.55	0.3827	0.3804
3	34.12	0.3729	0.3789
4	43.23	0.3677	0.3764
5	35.63	0.3720	0.3735
6	46.46	0.3670	0.3765
7	14.16	0.3675	0.3674
8	28.49	0.3882	0.3791
9	45.59	0.3658	0.3759
10	36.78	0.3737	0.3761
11	42.90	0.3682	0.3775
12	33.27	0.3835	0.3807
13	30.93	0.3633	0.3710
14	37.37	0.3704	0.3767
15	32.82	0.3877	0.3847
16	39.61	0.3768	0.3804
17	45.54	0.3710	0.3778
18	31.49	0.3659	0.3732
19	44.59	0.3687	0.3779

Table 1 - The average Yxy values of the 19 subjects under a set D65.

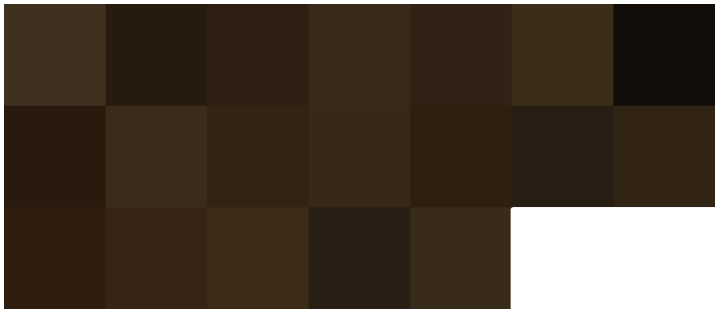
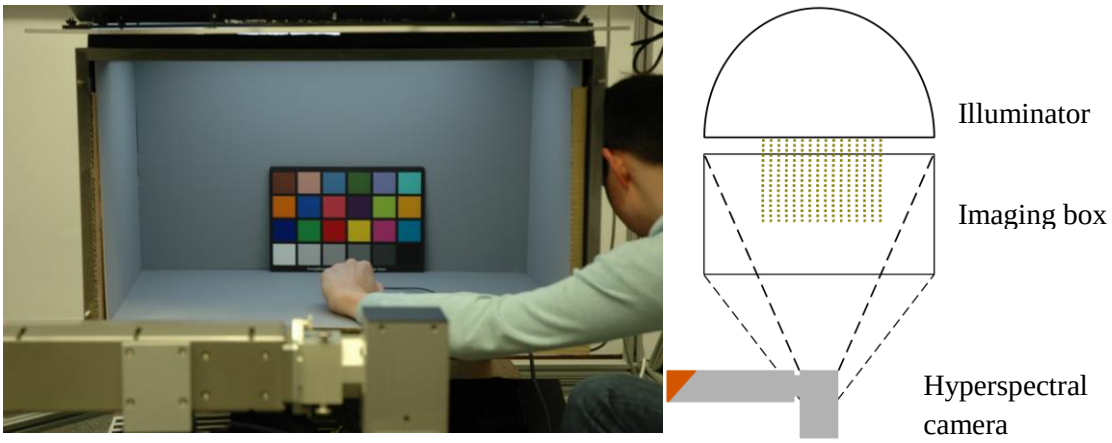


Figure 28 - RGB representation of the mean subject hand chromaticities

Each hand was captured using a Specim V10E hyperspectral camera which had been spectro-radiometrically calibrated to provide colourimetric outputs equivalent to a Spectrascan PR-650 as noted in [121]. Data was captured at a spectral resolution of 4nm,

from 400-780nm. An image of the hand scene captured under one illumination (D65); the spectrum of this was then factored out to result in the scene pixel-by-pixel reflectance being stored. This operation was undertaken by removing an illumination spectrum which was the average reflectance of a white patch under the given illumination. Thus this method will result in an accurate spectral reflectance shape being recovered, inaccuracies may result from any effects of spatial intensity variations of the original illumination within the box. However as our method does not utilise luminance gamuts this has no effect upon our data. Each of the 39 recorded database illumination spectra were then applied to this reflectance image to result in their related irradiance images. Gamut features, under each of these illuminations were then retrieved for a manually annotated hand area, which incorporated hand areas with variance in both illumination and capture angle. This was done in order to replicate real world imaging conditions as much as is possible within a laboratory environment. The use of the same single reflectance scene, on a subject-by-subject basis, removed any variances which would be introduced by separately imaging each subject under each illumination. An example scene can be seen below;



a) b)

Figure 29 - Database image capture in a) real terms, and b) in basic layout

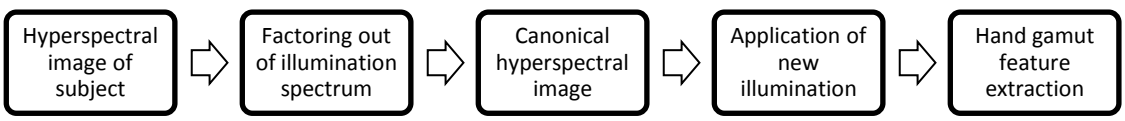


Figure 30 - Illustration of hand gamut feature extraction for a subject

Figure 31 - Comparing the 8 skin types for initial investigations under D40.

This data capture enabled the storage of the ground truth pixel-by-pixel irradiance values for each hand across the range of dataset illuminations. As such this data can be subsequently transformed into other colour spaces.

4.2.2 *Colour Space*

The choice of colour space is also inherently linked to the aim of human visual system replication. Previously it has been mentioned that machine colour spaces should be avoided due to their non-perceptually correct nature. As such we have looked toward the more perceptually and physiologically correct colour space models. What can be argued is that when dealing in these perceptual colour spaces we are aiming to replicate processing within the higher levels of the visual system, starting with the signal that is delivered to them after pre-processing in the retina and the subsequent early stages of the visual pathway. In computer vision terms this would be comparable, to some degree, to performing colour analysis work upon images that have already had their colour information down sampled, whether for the sake of compression or for other purposes. It is known that in the second stage of colour encoding in the retina, the cone signals are

pitted against each other in an opponent process, and these opponent signals are normalised by their mean activation level (averaged across both space and time) [40]. Here we use the cone-opponent contrast space defined in [122]. Because the choice of adaptation point strictly controls the chromaticity of the origin in this space [40], the choice of whitepoint (adaptation point) is of key importance to this work.

The importance of adaptation point is its use in the calculation of cone-opponent contrast value. The cone-opponent contrast values for each pixel in the image are calculated relative to the same adaptation point, and as such conversion is carried out to the same 'global' value. If we define our adaptation point as having cone excitations of L_a, M_a and S_a, we can calculate the cone-opponent contrast for a given pixel by;

$$\boldsymbol{L_C = \Delta L - L_a/L_a} \tag{34}$$

$$\boldsymbol{M_C = \Delta M - M_a/M_a} \tag{35}$$

$$\boldsymbol{S_C = \Delta S - S_a/S_a} \tag{36}$$

Where ΔL, ΔM and ΔS are the pixel LMS cone excitations, and L_C, M_C and S_C are the calculated cone opponent contrast values for the pixel, with regards to the adaptation point. The Luminance, RG and BY axis values are then calculated using the Eskew transform from [122];

$$\begin{bmatrix} \boldsymbol{Luminance} \\ \boldsymbol{RG} \\ \boldsymbol{BY} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0.78} & \boldsymbol{0.37} & \boldsymbol{0} \\ \boldsymbol{0.70} & \boldsymbol{-0.72} & \boldsymbol{0.02} \\ \boldsymbol{-0.55} & \boldsymbol{-0.25} & \boldsymbol{0.80} \end{bmatrix} \begin{bmatrix} \boldsymbol{L_C} \\ \boldsymbol{M_C} \\ \boldsymbol{S_C} \end{bmatrix} \tag{37}$$

4.2.3 *Whitepoint Selection*

The choice of whitepoint is not an arbitrary choice; it is a key indicator of the philosophical approach of this work. Whitepoints are generally taken to be the chromaticity of a defined white within the scene, but as colour constancy algorithms such as max-RGB [54] have shown, the danger of assuming that a specific pixel is the 'white' in the scene, and even the basic assumption that a white pixel exists in the scene is sometimes incorrect. We decided to investigate three different whitepoints, two of which would be applicable to real world imaging scenarios. The first of these whitepoints was chosen to be the illumination chromaticity; this setting is suited only to laboratory testing conditions in which the illumination chromaticity is known. This would show us the true variation of different gamut features with respect to changes in the illumination spectrum. The second choice was that of using the average hand Yxy as the whitepoint, to attempt to link changes in gamut directly to changes in hand

chromaticity under each illumination. This would make the method completely scene independent, focusing solely on the contents of skin. Thirdly we followed the grey world approach, using the scene average as the whitepoint.

Hand chromaticity gamut data for each subject was gathered using all three of these whitepoint methods under all illuminations. On top of this we decided to only utilise the variation of the skin gamut along the two chromatic axes, RG and BY, in order to investigate only the effect of changing illumination chromaticity, and thus ignore the effect of changing illumination intensity.

4.2.4 *Focusing upon the daylight locus*

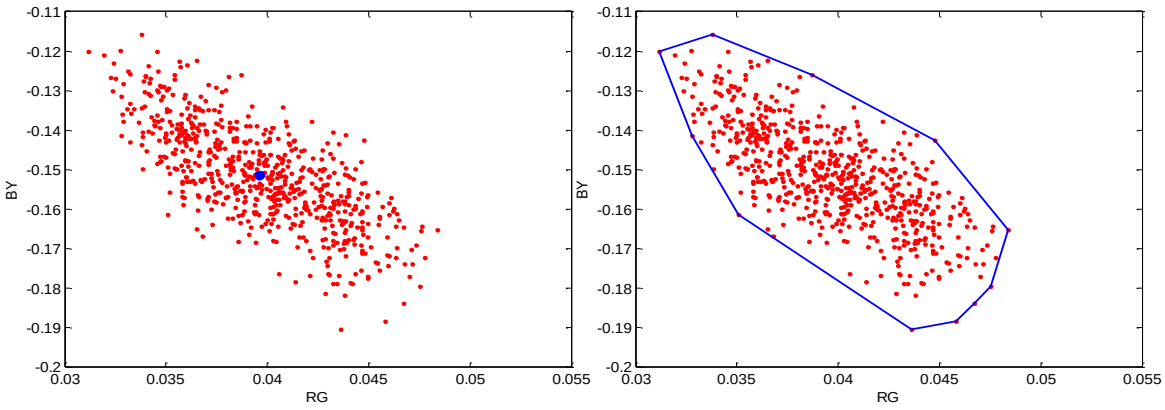
Whilst the illumination dataset contains a total of 39 illuminations it was decided that this first stage of work would concentrate on trying to see if human skin could be used as a cue to estimate illumination under only the daylight series of illuminations. This was because of the common nature of these illuminations in images. Thus the following work, and algorithms, has been developed under this assumption.

4.3 Visualisation and variation of skin gamuts

The next section will firstly demonstrate the typical distribution of skin gamuts within cone contrast space, alongside definitions of the different gamut features that will be used for characterisation. It will then move on to illustrate, across our database of subjects, the variation in each of these features across the daylight illuminations.

4.3.1 *Gamut data*

In order to track changes in skin gamut across illuminations, the gamut itself was required to be characterised using its features. Traditional gamut mapping uses a linear transform to translate the gamut of an image under an unknown illumination to an assumed canonical gamut. However from initial investigations it was noted that the skin gamut for subjects did not vary in a linear manner; therefore we investigated a group of specific gamut characteristics. These characteristics related not only to the shape of the convex hull gamut, but also to the distribution of chromatic points within the gamut. This was for a number of reasons, the first being that the convex hull is related only to the most extreme points of the gamut distribution, and although these may indeed give important information regarding major gamut distribution changes, the overall distribution may also give some cues. From considering a two dimensional gamut, a list of features was assembled for data collection and these are listed and illustrated below;



a) b)

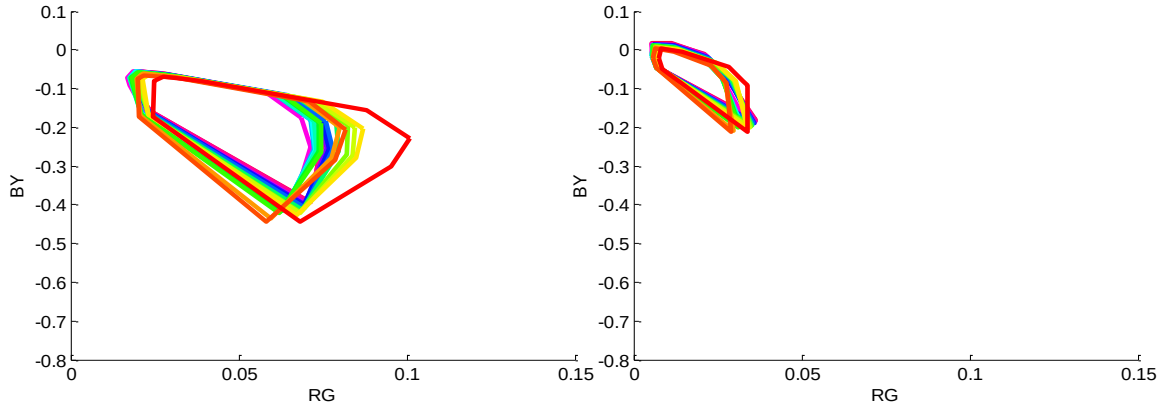
Figure 32 - Viewing a) the chromaticity points on a hand, and b) the convex hull of these points.

Gamut feature	**Explanation**
Area	Area of the hue gamut
Centre of mass	Distribution centre of mass
Extreme RG and BY points	The two points on the convex hull with the greatest distance between them.
Gamut spread distance	Distance between the furthest two points on the convex hull.
Gamut spread angle	Angle between this spreading line and the RG axis (x axis)
Extreme point angles	Angle between the origin and the two extreme RG and BY points.
Convex hull points	Location of each point on the convex hull of the gamut distribution.

Table 2 - An explanation of the main features gathered from each hue gamut

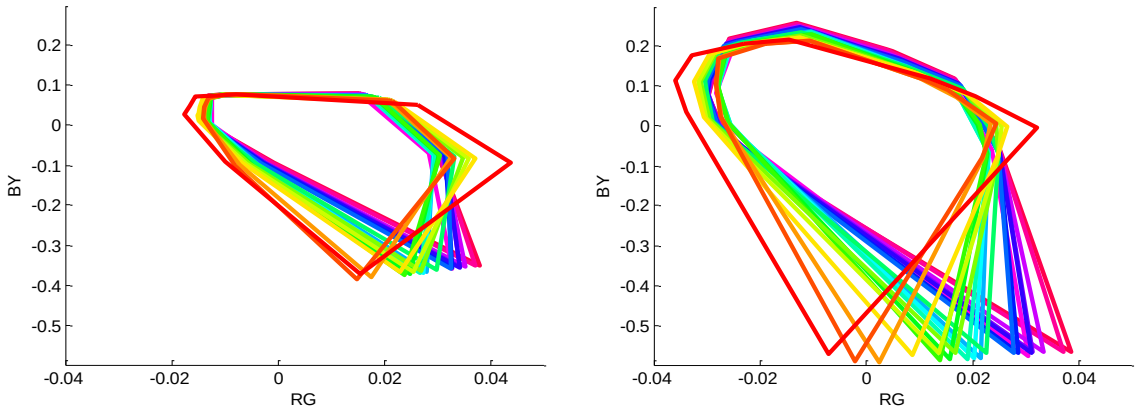
4.3.2 *Variation of skin gamut across the daylight spectrum*

Firstly we will illustrate how some of the collected features vary with respect to illumination and whitepoint, mostly with respect to the gradual change across the CIE daylight loci.



a) b)

Figure 33 - The changing skin gamut hull across the daylight locus for a) Subject #1, and b) Subject #7 using the illumination chromaticity as the whitepoint.



a) b)

Figure 34 - The changing skin gamut hull across the daylight locus for a) Subject #1, and b) Subject #7 using the hand average as the whitepoint.

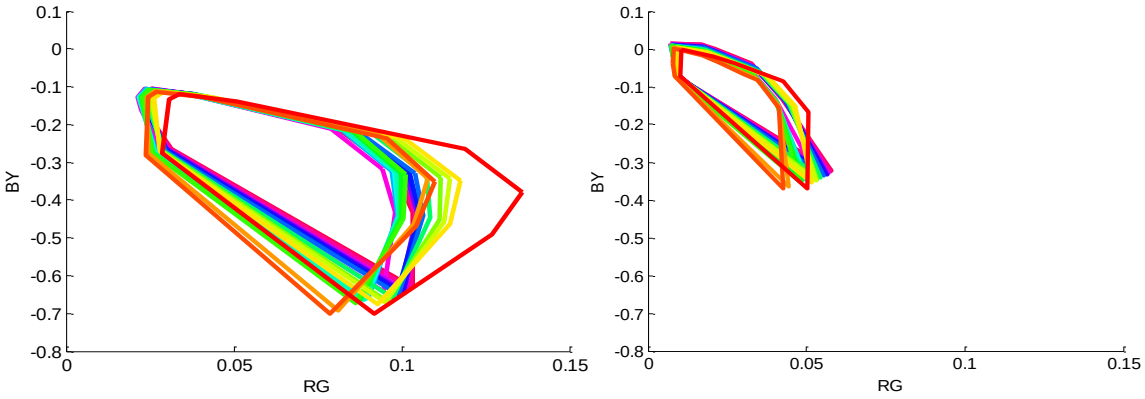


Figure 35 - The changing skin gamut hull across the daylight locus for a) Subject #1, and b) Subject #7 using the scene average chromaticity as the whitepoint.

As can be seen in Figure 34 the skin gamut convex hull for the subject above does vary with respect to illumination, even across the small changes between neighbouring D type illuminations. These changes are also exhibited across the three possible whitepoint choices. What can also be seen is the similarities in gamut shape between the outputs of two different whitepoint choices, those of the illumination and the scene average. It can be considered however that one of the whitepoint choices, that of the illumination Yxy, is a reference for other whitepoint based gamuts to be compared to.

The use of these differing whitepoints is also an exercise to support findings from previous papers that have discussed cone-opponent space [12, 27, 123] and the variation of polychromatic object gamuts [15]. Under a correct whitepoint assumption, such as the scene average (when compared to the skin average), certain features of the polychromatic gamut of a natural surface should stay constant across daylight illuminations. The constancy of these features is then directly related to the invariance of cone ratios within a scene. There when we view the changing skin gamuts under the daylight illuminations when using the skin average whitepoint we find that no features exhibit invariance. However the scene average whitepoint gamuts for both subjects in

Figure 35 show a greater degree of basic invariance in gamut shape and location of hull points. A greater discussion of the differing levels of invariance of the skin gamuts across different features will be had later.

As such the remainder of this chapter will only exhibit the results of using either the hand average or scene average as the white point for colour space transforms.

4.3.3 *Skin gamut variation for 8 subjects*

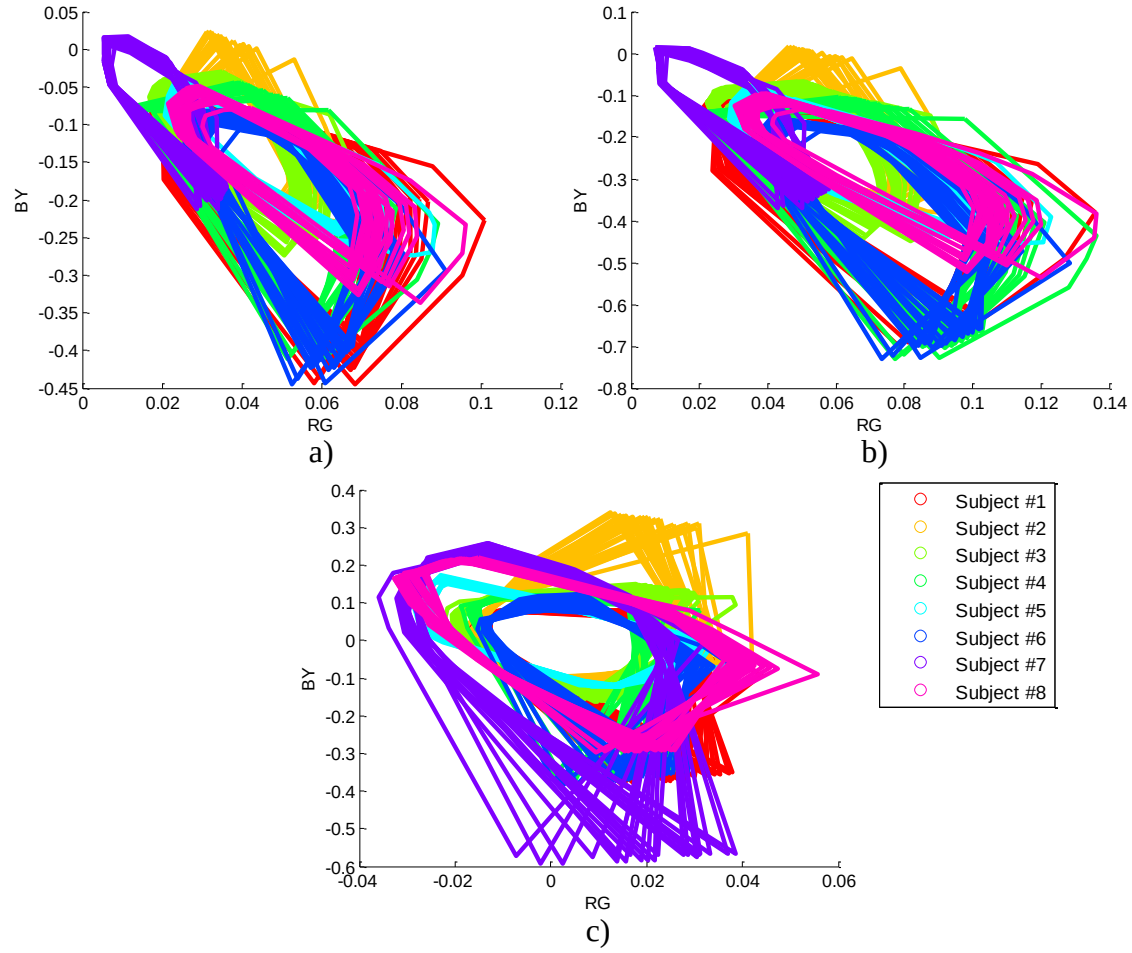


Figure 36 - Comparing the variation in hand gamut across subjects and the daylight locus for a) the illumination chromaticity, b) scene average chromaticity, and c) hand average chromaticity as the whitepoint.

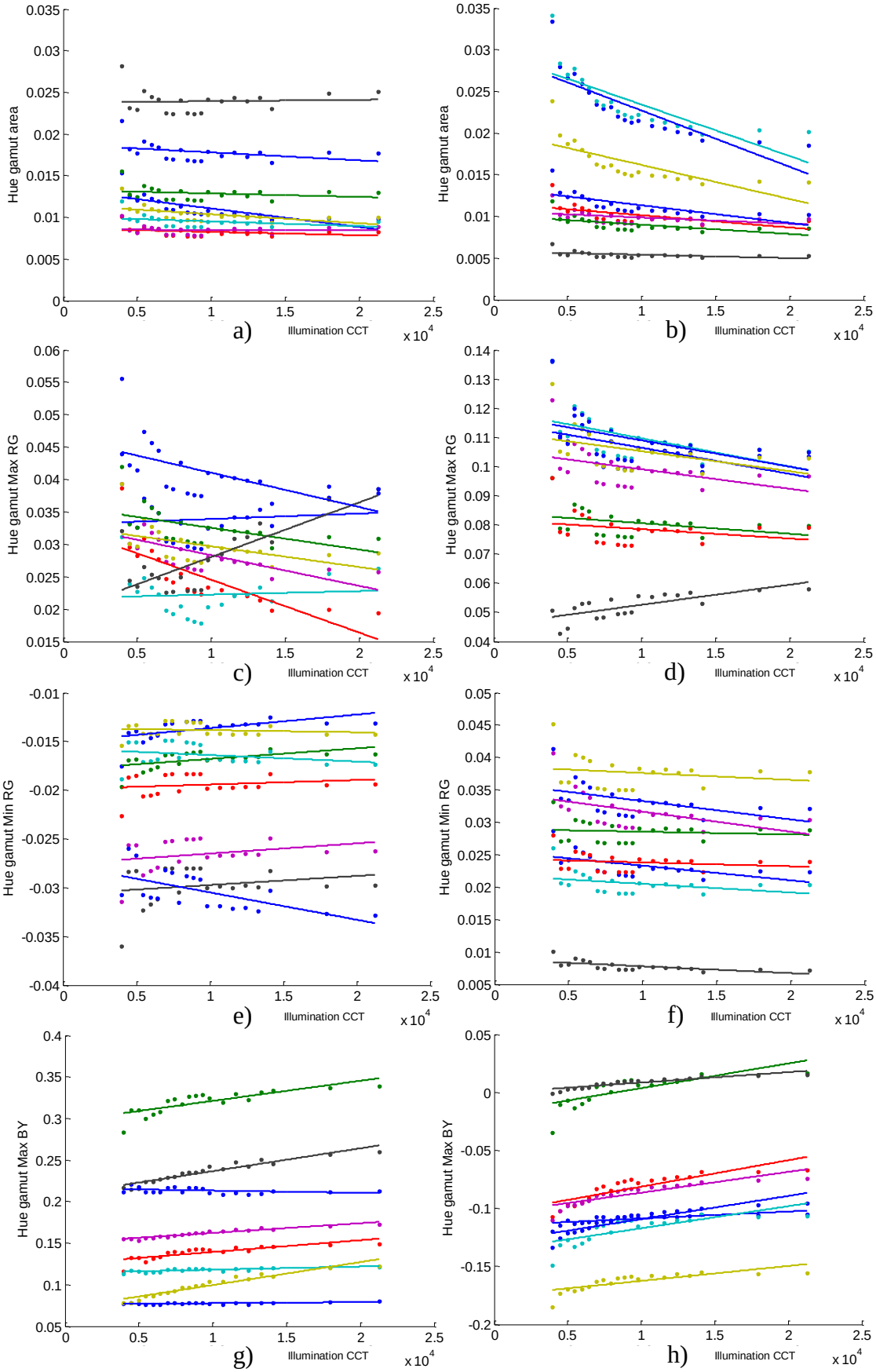
Figure 36 shows that across the daylight locus the gamuts of the initial 8 subjects (giving a wide spread of skin chromaticities) are located within a small area of the colour space. This matches previous research findings placing skin within a small region of other colours spaces [85, 100, 124]. Furthermore it can also be seen that the general shape of the gamut, on a subject-by-subject basis, does not vary to a large degree across the illumination spectrum. On from this it can be noted that the majority of changes in the convex hull are restricted to those caused by movements in some of the most extreme points in the gamut. Across all subjects it can be seen that the majority of movement takes place in the form of a shift along the RG axis in the green direction. This is interesting when considered alongside the somewhat stable nature of the left extreme of each subjects' gamut.

The similarities between the convex hull variations between the use of the illumination and scene average as the white point can also be seen. For this reason, and the fact that

prior knowledge of the illumination chromaticity would render the proposed method defunct we will from this point only compare the use of the hand average and scene average chromaticities as the whitepoint for colour transformation.

Further to this, in order to compare the remaining gamut features the plots will contrast gamut feature to illumination CCT.

4.3.4 *Gamut feature variation*



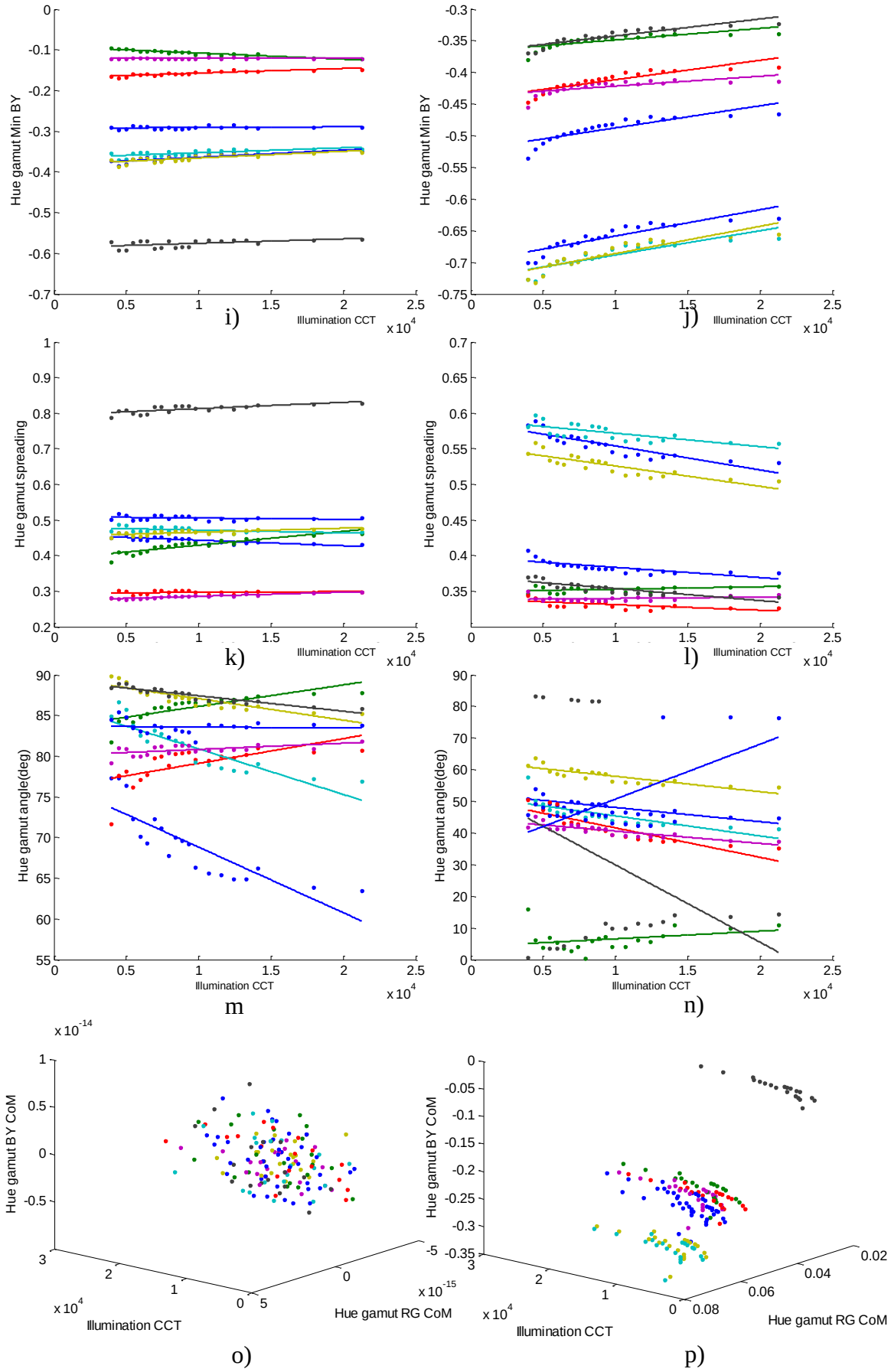


Figure 37 – Variation in hue gamut area across daylight illumination CCT for a) hand average, and b) scene average chromaticity as the white point. Variation of maximum

RG hull point across daylight illumination CCT for c) hand average, and d) scene average chromaticity as the white point. Variation of minimum RG hull point across daylight illumination CCT for e) hand average, and f) scene average chromaticity as the white point. Variation of maximum BY hull point across daylight illumination CCT for g) hand average, and h) scene average chromaticity as the white point. Variation of minimum BY hull point across daylight illumination CCT for i) hand average, and j) scene average chromaticity as the white point. Variation of hue spreading across daylight illumination CCT for k) hand average, and l) scene average chromaticity as the white point. Variation of hue angle across daylight illumination CCT for m) hand average, and n) scene average chromaticity as the white point. Variation in hue centre of mass across daylight illumination CCT for o) hand average, and p) scene average chromaticity as the white point.

As was noted above the general shape of the gamut on a subject-by-subject basis did not vary to a large degree across the illumination spread. However we then compared the variation in gamut area for these 8 subjects.

4.3.5 *Hue angle*

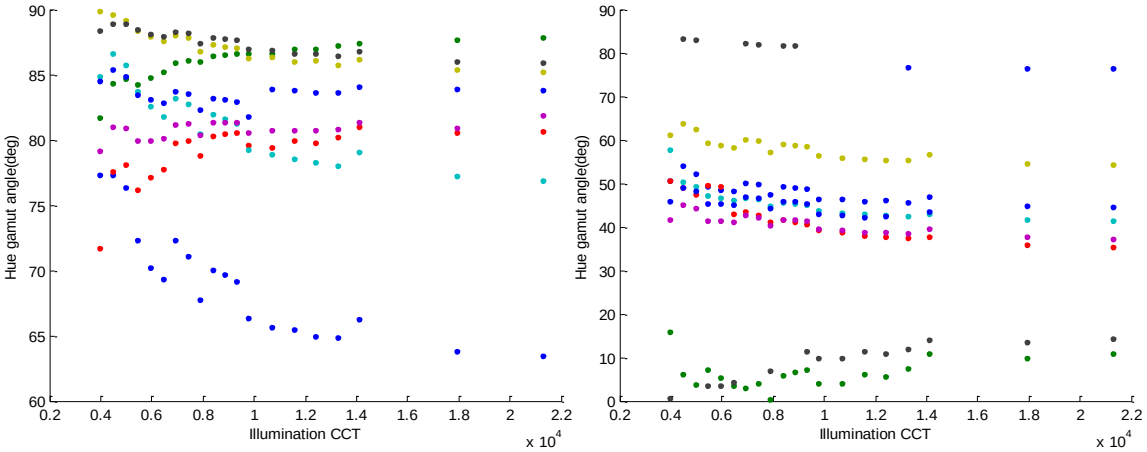


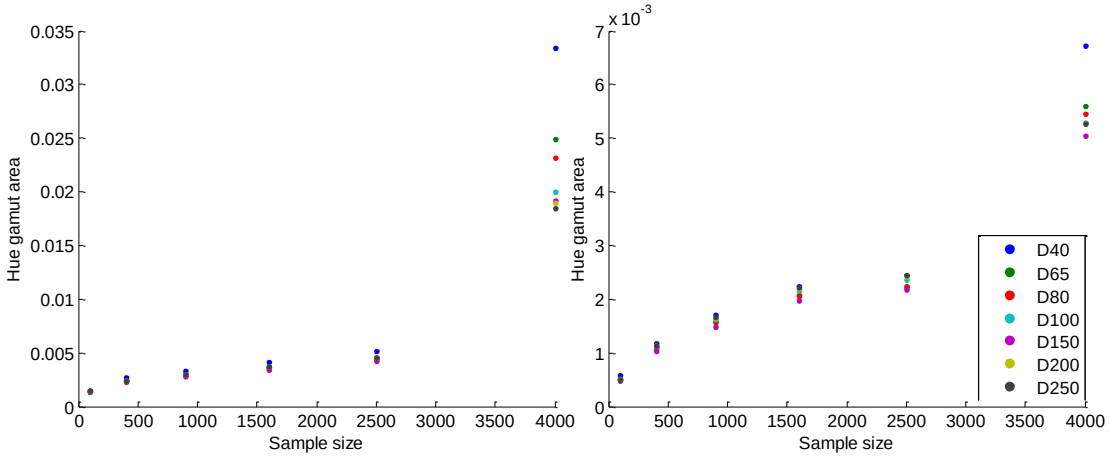
Figure 38 - Variation in hue angle across daylight illumination CCT for a) hand average, and b) scene average chromaticity as the white point.

4.3.6 *Patch size invariance*

As these features are extracted from colour data on a pixel-by-pixel basis the size of the sampled patch on skin will obviously have an effect upon the retrieved gamut. If we alter the sampled patch size and retrieve different features this will pose a major challenge to our hypothesis and resulting method. As such investigation into the effect of sampled size upon gamut features was undertaken. This involved the comparison of different patch sizes across our initial 8 subjects; 10*10 (100 pixels), 20*20 (400 pixels), 30*30 (900 pixels), 40*40 (1600 pixels), 50*50 (2500 pixels) and Entire hand (4000+).

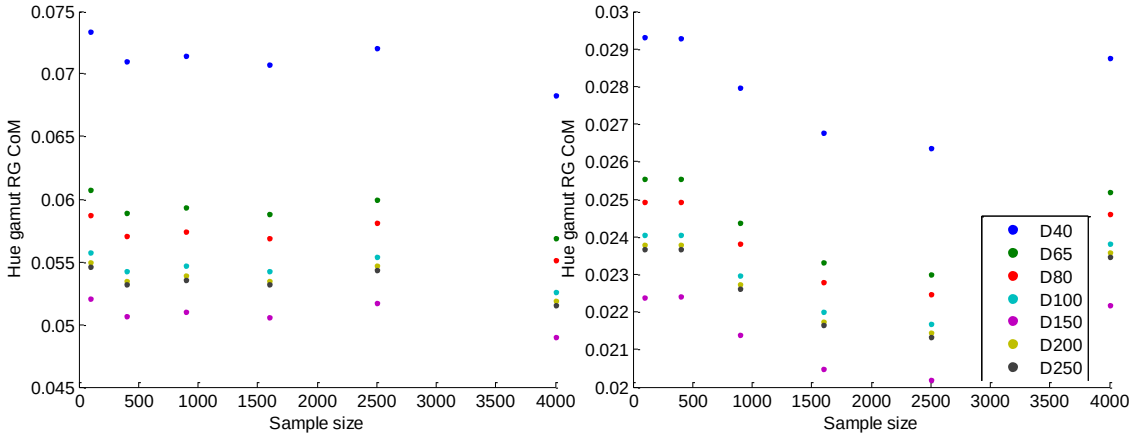
The variations in retrieved features for a number of subjects across these patch sizes are shown below.

Figure 39 - Illustration of the different sample patch sizes



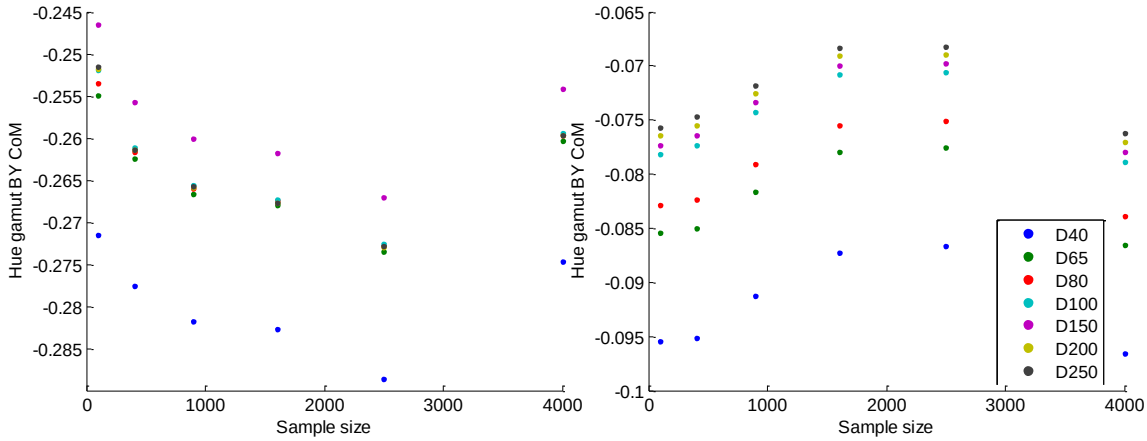
a) b)

Figure 40 - Comparing the variation in hue gamut area across sample sizes for a) Subject #1, and b) Subject #7 using the hand average chromaticity as the whitepoint.



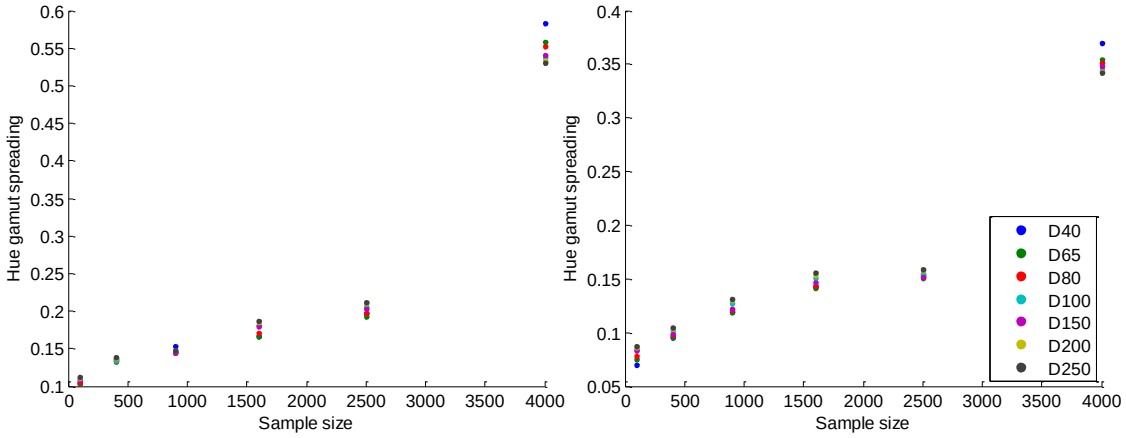
a) b)

Figure 41 – Comparing the variation in hue gamut RG CoM across sample sizes for a) Subject #1, and b) Subject #7 using the scene average chromaticity as the whitepoint.



a) b)

Figure 42 - Comparing the variation in hue gamut BY CoM across sample sizes for a) Subject #1, and b) Subject #7 using the scene average chromaticity as the whitepoint.



a) b)

Figure 43 - Comparing the variation in hue spreading across sample sizes for a) Subject #1, and b) Subject #7 using the hand average chromaticity as the whitepoint.

As was expected sample size can be seen to have an effect upon retrieved features for all subjects. The change in features across sample sizes for both subjects can be seen to be very similar. However the largest sample size of 4000 can be seen to show a different set of features, with this exhibited across the range of daylight illuminations. Figure 40 and Figure 43 show the direct effect of sample size upon the gamut area and spreading values respectively. Whilst Figure 41 and Figure 42 show the stability in the maximum RG and BY CoM gamut features across sample sizes. As has been seen any sample size above 4000 pixels roughly returns a consistent set of skin gamut feature characteristics. This total, when we consider the high pixel count of modern imaging devices, is an adequate one. However we must ensure that we utilise this knowledge to enforce a minimum sample size criterion.

4.3.7 *BY variation with luminance*

We previously stated that we chose to utilise only the hue axes of colour contrast space, RG and BY, whilst discarding the luminance component. The luminance component was discarded in order to ignore any effects caused by differences in illumination irradiance, so that all gamut features are directly related to illumination chromaticity. However we have discovered that the luminance and BY axes are linked due to the optical nature of human skin. Through testing with different luminance levels we have noted that as luminance increases the value of the corresponding BY co-ordinate also increases. This means that in uncontrolled scenarios the BY co-ordinates in cone contrast space should not be utilised to determine skin type or illumination estimates.

4.4 *Variations in features for different skin groups*

The previous figures have shown that on an individual basis skin gamut features do indeed vary with illumination. Across our small sub sample of human skin colours it should therefore be possible, at least in theory, to extract an estimate of the illumination from the skin gamut features. But the fact that skin gamut features vary with skin type, as made evident in previous sections (compare subjects 1 and 7, for example, in Figure 37 above), makes it necessary to determine skin type before carrying out illumination estimation. However if we are to produce a method which can operate with any skin colour we will need to model the viewed illumination dependent features for a more discrete number of skin colours, rather than the continuous spectrum of colours which exist. This will simplify the task of illumination estimation by reducing the dimensions of the problem.

In order to reduce our initial 8 subjects into a smaller number of skin type groups we investigated similarities in the way different subjects' gamut features varied across the set of illuminations in conjunction with known skin group. The first breakthrough came when analysing differences between subjects, where three main groupings of subjects can clearly be seen. This grouping was also reinforced when analysing the centre of mass variations using k-means clustering;

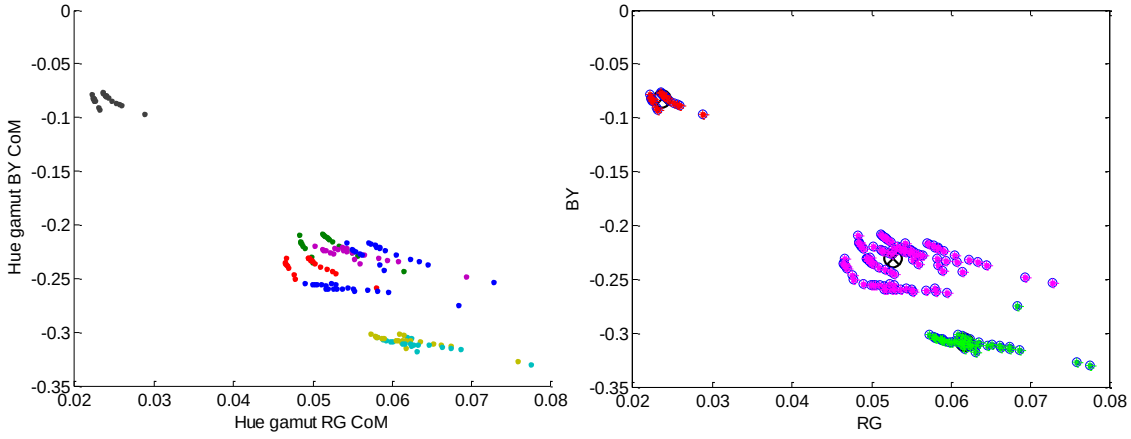
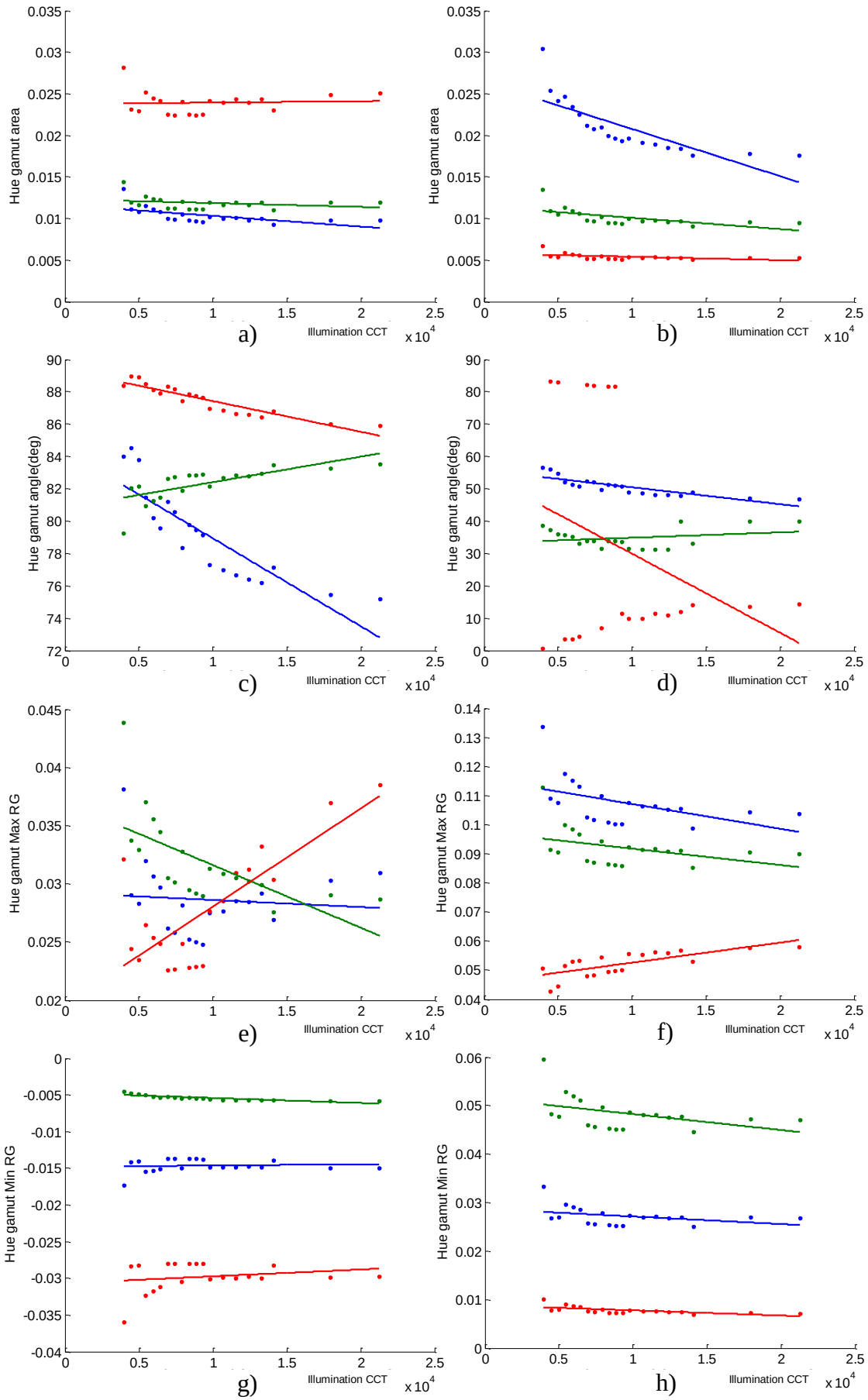


Figure 44 - Grouping our 8 subjects into skin type groups based upon a) centre of mass variation across the DL illuminations, using b) k-means clustering

These results show that we can indeed characterise skin type, to a certain degree, before performing any sort of processing. The k-means clustering here separated the data into 3 groups based upon the BY component of the centre of mass values across the daylight illuminations. However these same three groups can also be approximated using the RG components.

If we then group these subjects into the resulting three skin types and average their feature variations across the illuminations we return the following feature set values that will be used to estimate the illumination after skin type estimation:


Hue gamut area
Illumination CCT
x 10^4
a)
b)
Hue gamut angle(deg)
c)
d)
Hue gamut Max RG
e)
f)
Hue gamut Min RG
g)
h)

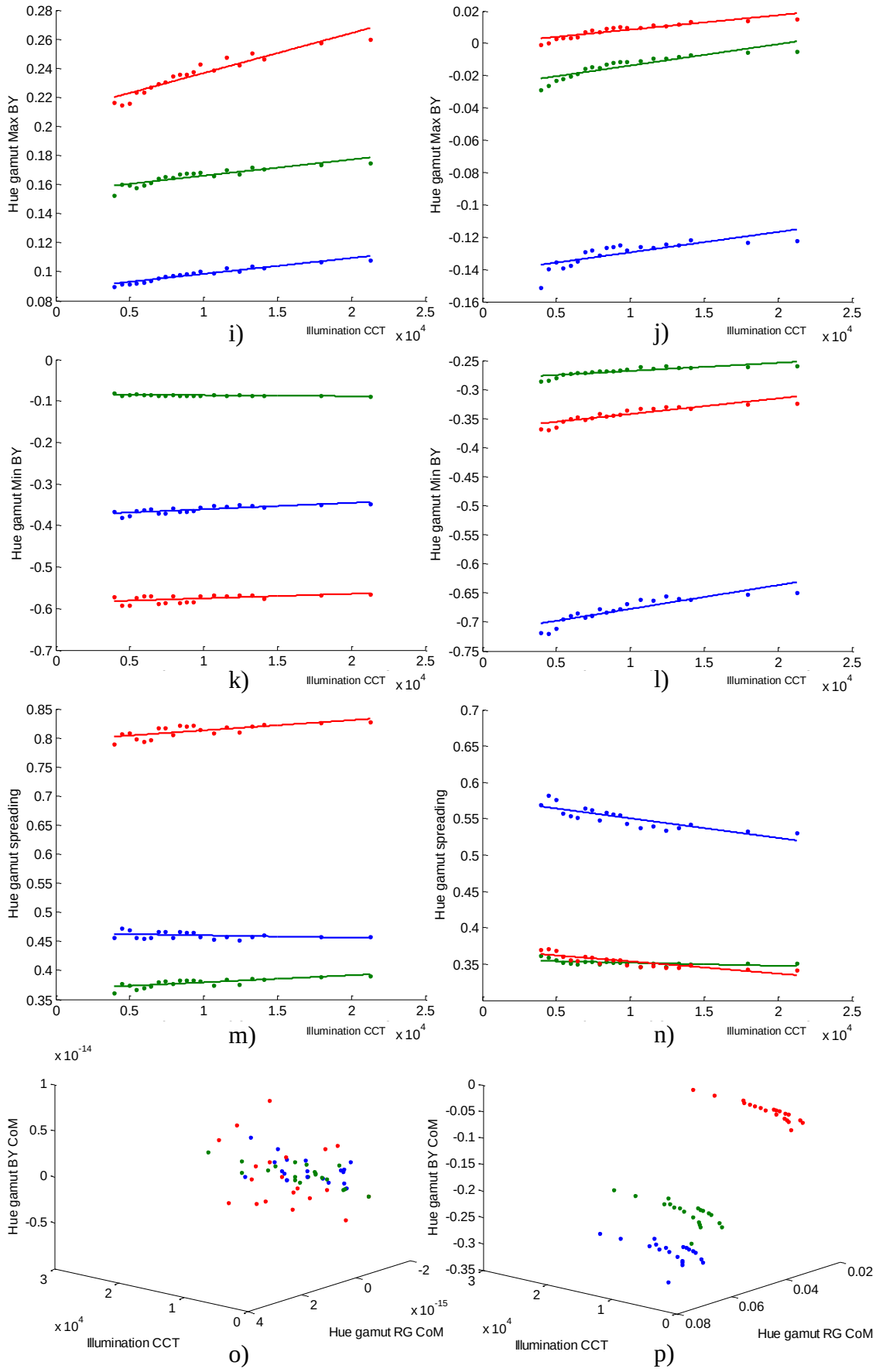


Figure 45 - Variation in hue gamut area for skin groups across DL illuminations using a) hand average, and b) scene average chromaticity as white point. Variation in hue gamut

angle for skin groups across DL illuminations using c) hand average, and d) scene average chromaticity as white point. Variation in max. RG hull point for skin groups across DL illuminations using e) hand average, and f) scene average chromaticity as white point. Variation in min. RG hull point for skin groups across DL illuminations using g) hand average, and h) scene average chromaticity as white point. Variation in max. BY hull point for skin groups across DL illuminations using i) hand average, and j) scene average chromaticity as white point. Variation in min. BY hull point for skin groups across DL illuminations using k) hand average, and l) scene average chromaticity as white point. Variation in hue gamut spreading across DL illuminations using m) hand average, and n) scene average chromaticity as white point. Variation in hue gamut centre of mass for skin groups across DL illuminations using o) hand average, and p) scene average chromaticity as white point.

We have seen in Figure 44 that the rough skin type grouping of a subject can be estimated prior to any image pre-processing though the use of the minimum RG hull point when using the average hand chromaticity as the whitepoint, and the RG CoM point when using the scene average chromaticity as the whitepoint. This means that using the colour contrast space we can classify human skin colour before any colour constancy/ white balance processing [43].

In terms of illumination estimation performance we need to firstly uncover which of our two remaining whitepoint methods result in a superior level is illumination discrimination across our illuminations. Although the ideal whitepoint method for use would be the skin average chromaticity of the patch itself we have seen from the previous figures that use of the scene average whitepoint results in greater monotonicity for changes in features across the range of illuminations.

Statistical analysis (using multivariate ANOVA) reveals that for certain gamut features there is a significant main effect of illumination, and an interaction effect between illumination and skin type, when the scene average chromaticity is used as the whitepoint (Table 4), but not when the hand skin average chromaticity is used (Table 3).

Therefore, from this point forward we consider only the scene average chromaticity whitepoint method.

We have also noted from Figure 45e, f, g and h that the maximum and minimum RG hull points also seem to vary in a predictable manner with respect to illumination. Other features, such as gamut area, and maximum/ minimum BY along with spreading value also appear to give a good cue to the illumination. Although across all of these features, as we increase CCT value we can see that neighbouring illuminations exhibit similar skin gamut feature values. This observation suggests that the success in estimating the

illumination using these features may be better for lower CCT illuminations. These statistical findings can be found within Table 4.

Even though it was previously stated that this work will solely utilise the scene average chromaticity as the whitepoint, statistical analysis of the hand gamut white point with regards to the aforementioned skin groups was also undertaken. This has shown, in Table 3 that, using the skin type groups proposed by the previous k–means clustering, the produced skin types exhibit a significant relationship with regards to gamut features. On top of this Table 3 also proves statistically that illumination cannot be related to subject or skin type when utilising the hand average chromaticity as the whitepoint.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Skin Type	Spreading	2.925	2	1.462	189.133	.000
	Hue Angle	872.122	2	436.061	9.837	.000
	Gamut Area	.003	2	.001	108.420	.000
	Max. RG	.000	2	.000	3.085	.050
	Min. RG	.002	2	.001	19.483	.000
	Max. BY	.357	2	.179	31.500	.000
	Min. BY	2.637	2	1.319	146.534	.000
Illumination	Spreading	.007	19	.000	.046	1.000
	Hue Angle	39.188	19	2.063	.047	1.000
	Gamut Area	8.930E-005	19	4.700E-006	.358	.993
	Max. RG	.001	19	4.336E-005	.902	.582
	Min. RG	.000	19	9.258E-006	.187	1.000
	Max. BY	.008	19	.000	.076	1.000
	Min. BY	.002	19	.000	.014	1.000
*Skin Type * Illumination*	Spreading	.004	38	.000	.013	1.000
	Hue Angle	292.403	38	7.695	.174	1.000
	Gamut Area	1.295E-005	38	3.408E-007	.026	1.000
	Max. RG	.001	38	1.478E-005	.307	1.000
	Min. RG	2.416E-005	38	6.359E-007	.013	1.000
	Max. BY	.003	38	6.662E-005	.012	1.000
	Min. BY	.002	38	4.150E-005	.005	1.000

Table 3 - The results of multivariate analysis comparing gamut features across illuminations and assigned skin types using the hand average chromaticity as the whitepoint.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Skin Type	Spreading	.632	2	.316	34.539	.000
	Hue Angle	18852.101	2	9426.050	49.912	.000
	Gamut Area	.003	2	.002	44.604	.000
	Max. RG	.058	2	.029	1834.217	.000
	Min. RG	.008	2	.004	89.149	.000
	Max. BY	.359	2	.180	132.586	.000
	Min. BY	1.856	2	.928	71.378	.000
Illumination	Spreading	.005	19	.000	.031	1.000
	Hue Angle	14323.882	19	753.889	3.992	.000
	Gamut Area	.000	19	7.697E-006	.226	1.000
	Max. RG	.002	19	.000	7.175	.000
	Min. RG	.000	19	1.130E-005	.261	.999
	Max. BY	.007	19	.000	.270	.999
	Min. BY	.022	19	.001	.091	1.000
*Skin Type * Illumination*	Spreading	.002	38	4.802E-005	.005	1.000
	Hue Angle	19522.032	38	513.738	2.720	.000
	Gamut Area	.000	38	2.648E-006	.078	1.000
	Max. RG	.001	38	3.545E-005	2.251	.001
	Min. RG	5.324E-005	38	1.401E-006	.032	1.000
	Max. BY	.001	38	2.254E-005	.017	1.000
	Min. BY	.001	38	2.635E-005	.002	1.000

Table 4 - The results of multivariate analysis comparing gamut features across illuminations and assigned skin types using the scene average chromaticity as the whitepoint.

Table 4 above illustrates the relationship between daylight illumination and gamut features across all skin types. If we ignore skin type, and consider all subjects on the basis of a common grouping we see significance across a number of features with the greatest significance being with the maximum RG hull point; $F(19,100)=7.175$, $p=0.000$. When considering skin type only we can see that skin type is inherently linked to all of the tested features, as is expected.

The interaction between skin type and illumination is the most important relationship for our aim of estimating illumination CCT from the values of certain features, or sets of features, the analysis reveals a significant skin type-illumination interaction for hue angle, $F(38,100)=2.720$, $p=0.000$, and for the maximum RG hull point $F(38,100)=2.251$, $p=0.001$.

4.5 Implementation of findings into a real algorithm

From these findings it can be noted that the most significant features for illumination estimation using our calculated skin type groups are the hue angle and the maximum RG hull point. However we must remember that these statistical analyses were carried out upon the average feature values across the subjects in each group, so we must consider their variance. If we look back to Figure 37m and Figure 37n we can see that within each of the three groups there is a great deal of inter-subject angle variation. This means that use of the hue angle on its own cannot be utilised to accurately define skin type. Figure 37c however shows a much smaller level of variation in maximum RG hull location across subjects. We will thus only investigate the efficacy of utilising the maximum RG hull location as a cue for illumination estimation.

4.5.1 *Skin gamut variation across all database illuminations*

The previous figures have shown that gamut features do vary across the daylight locus of illuminations. However previous work has shown that the chromaticity of skin is located on a locus, related to melanin levels [85], across different colour spaces [43]. As such we also investigated the location of a skin gamut under D65 in comparison to their locations under other daylight and non-daylight illuminations. We investigated this by looking into the variation in skin type centre of mass across our daylight and non-daylight illuminations.

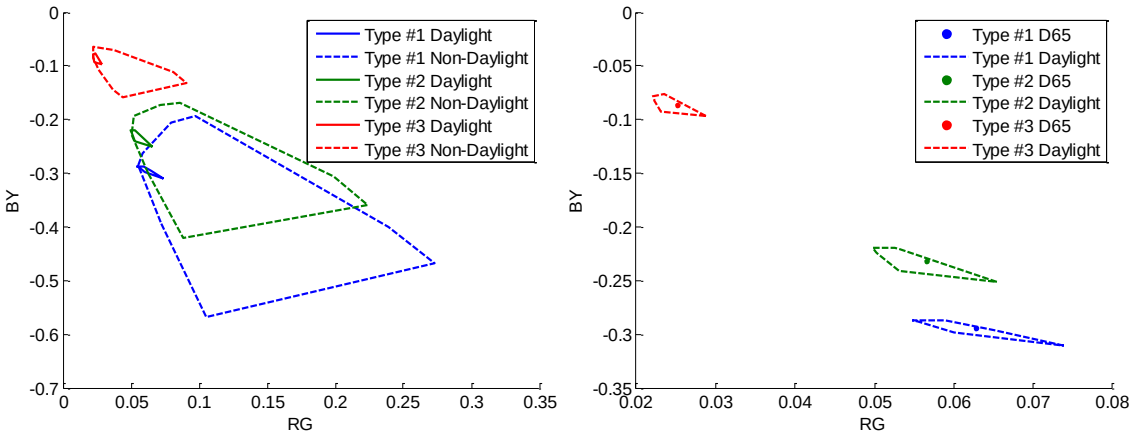
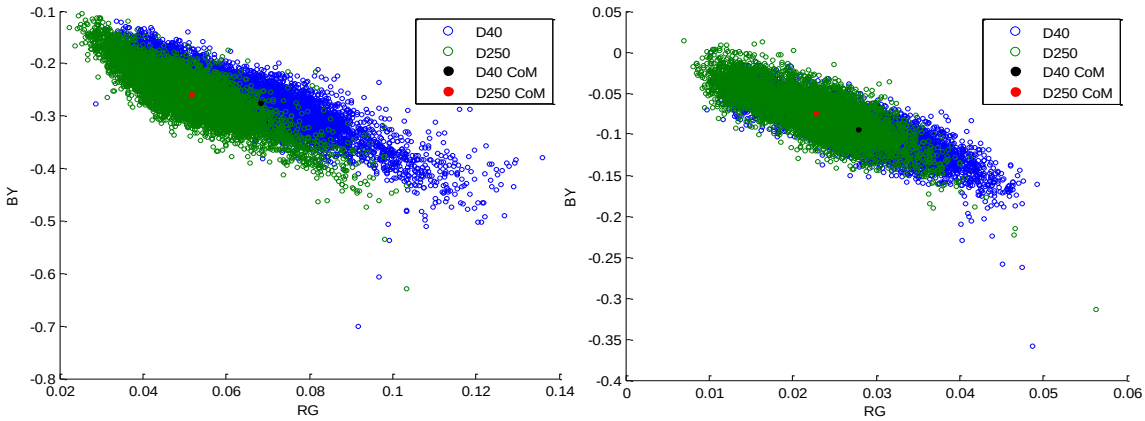


Figure 46 - Comparing the convex hull of skin type average centre of masses between a) daylight and non-daylight illuminations, and b) D65 and remaining daylight illuminations.

Figure 46a above shows the small region in which skin gamut centre of masses exist within each skin group in comparison to their location under non-daylight illuminations. Figure 46b illustrates the small region within the daylight CoM hull at which the CoM under D65 exists for each skin type. For all skin types it can be seen that their CoM is consistently present close to the centre of this overall daylight hull. The variations in CoM can also be seen to be affected to a greater degree within the RG component. This means that under daylight illuminations a correction could easily be applied to shift the centre of mass of a gamut, and the remaining chromaticity points towards this area. Furthermore across the three skin types it can be noted that the pattern of variation in centre of mass with illumination (whether daylight or non-daylight) is similar (i.e. the shapes of the convex hulls are similar), which most likely reflects the similarity in structural components of skin, despite small variations in their relative concentrations.

This opens up the possibility that a form of colour normalisation could be applied to images via shifting the characterised skin gamut to its appropriate D65 location, and applying an equivalent shift to the corresponding image gamut.

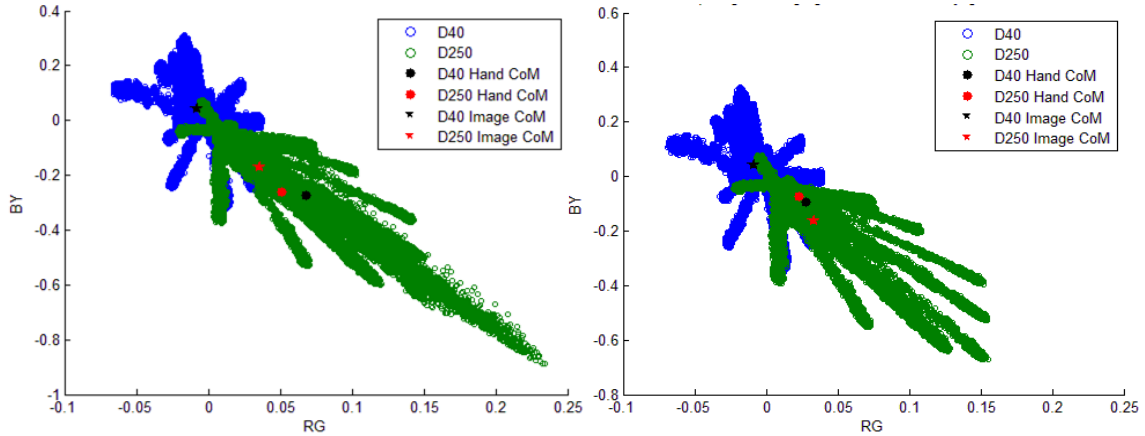
However before assuming we can shift the skin gamut and corresponding image gamut we must investigate how these vary across illuminations, in terms of chromaticity distributions, for different subjects. In order to uncover any large changes we have compared the skin and image gamut for subjects under both D40 and D250;



a) b)

Figure 47 - Comparing the skin chromaticity gamut across D40 and D250 for a) Subject #1, and b) Subject #7.

However if we plan on altering the overall image gamut using the same method we also need to view the changes in image gamut across these two illuminations;



a) b)

Figure 48 - Comparing the changes in image gamut under D40 and D250 for a) Subject #1, and b) Subject #7.

Figure 48 above shows that changes in skin gamut location are not equivalent to those for the image gamut. The distribution of chromaticities in each image can also be seen to be quite different. Distance between mean image and hand chromaticity also varies between the two illuminations across the two subjects. Thus if we shift the entire image gamut by the D65 centre of mass offset of the skin patch we may not fully colour correct the image. However we must note that these differences in gamut are related to

the scene average whitepoint utilised for conversion of the image into cone-contrast space.

As such if we shift the skin hand gamut into its position under D65, along with a subsequent shift of the entre image gamut, and then transform this back into RGB space using our newly assumed D65 whitepoint we will may gain a degree of colour correction.

4.5.2 *Illumination estimation using skin chromaticity gamuts*

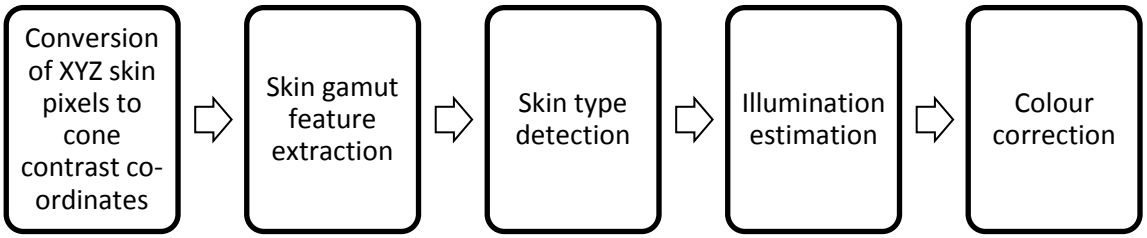


Figure 49 - Explaining the skin gamut based illumination estimation method

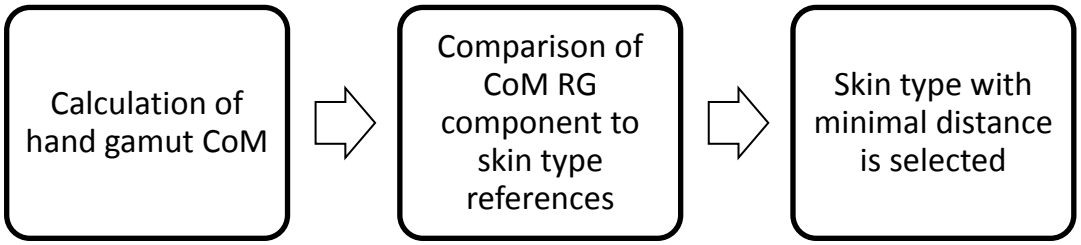


Figure 50 - Explanation of skin type estimation

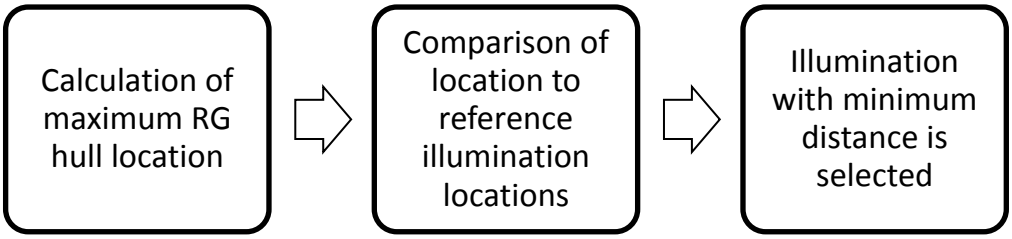


Figure 51 - Explanation of illumination estimation

4.5.3 *Colour normalisation using skin chromaticity gamuts*

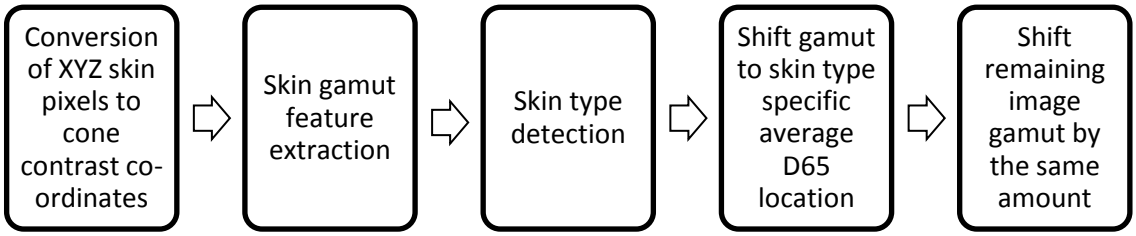


Figure 52 - Explaining the skin gamut based normalisation method

As previously shown in Figure 46, and discussed, we have discovered that within cone-contrast space human skin exists within a very small region, when assuming daylight illumination. This reinforces previous findings [43, 85]. The general lack of change in the shape of the chromaticity gamut distribution across the daylight illuminations shown in both Figure 34b and Figure 35 also means that a general offset shift in gamut centre of mass could be used to colour correct an image to a certain degree.

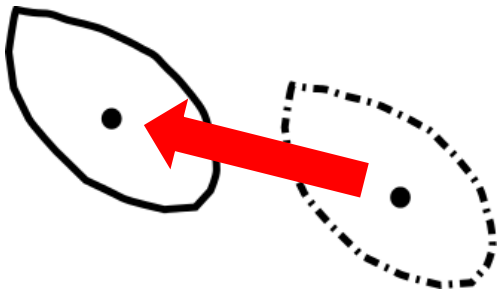


Figure 53 - Illustration of the proposed normalisation method

4.6 Illumination estimation algorithm performance

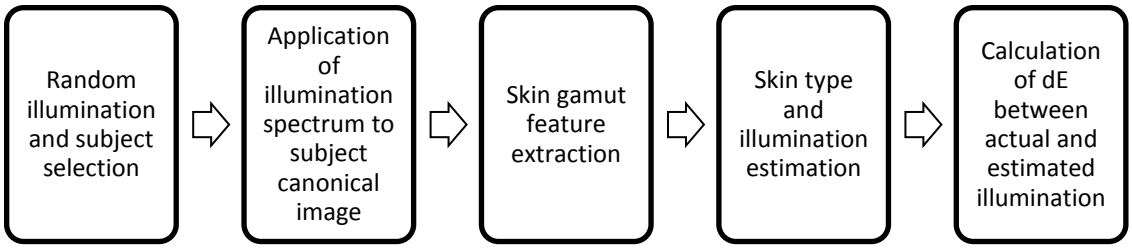


Figure 54 - Outline of testing methodology for the illumination estimation method

In order to test the efficacy of both the skin type estimation and illumination estimation methods, along with their combined performance the three methods were required to be tested separately.

4.6.1 *Testing methodology*

- The range of subjects for testing has been increased from the 8 utilised for the creation of the model to 19, meaning 11 subjects under the 20 illuminations can be treated as unknowns.
- The testing method involved the use of 10000 iterations where both subject and illumination are chosen at random.

4.6.2 *Extended subject dataset for testing*

Figure 55 illustrates the hand gamut centre of mass changes across the daylight illuminations for the extended subject list of 19 people as described in Table 1. As can be seen the vast majority of our subjects fit within our current skin type group clusters.

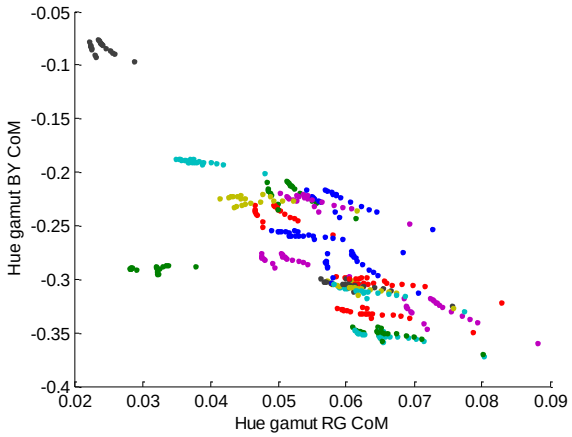


Figure 55 - Illustrating the gamut centre of mass changes for the extended range of subjects across the daylight locus, where different subjects are denoted by different colours

4.6.3 *Skin type estimation performance*

Skin type estimation performance is shown below in Figure 56 with skin types of the new unknown subjects set using the same k means clustering method as previously used.

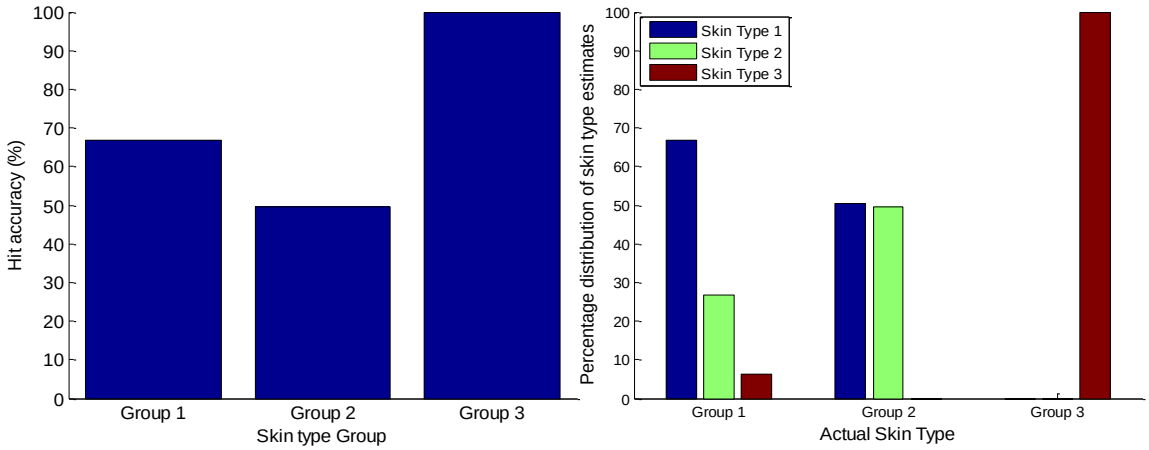
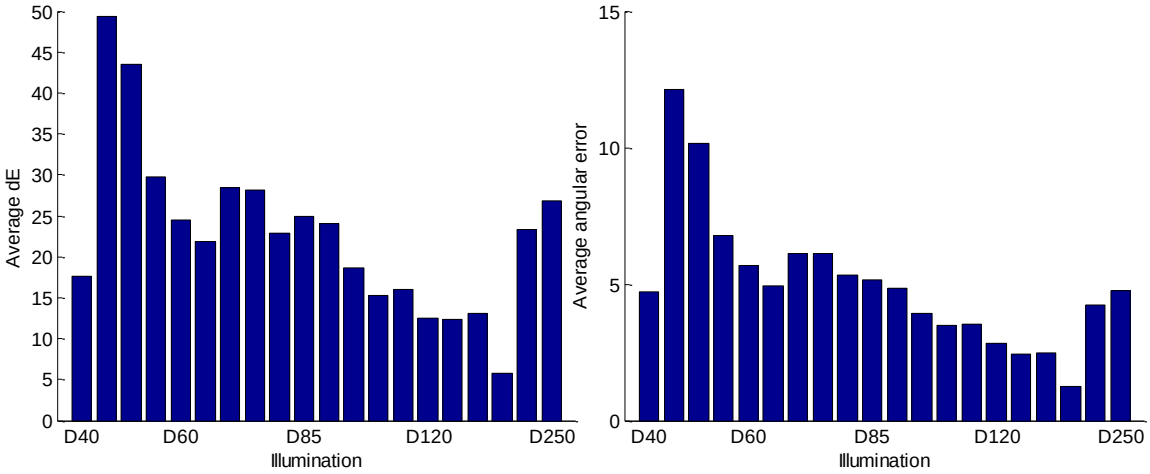


Figure 56 - Estimating subject skin type using a) hit rate, and b) distribution of skin type estimates for each skin type group.

If we first consider the skin type estimation accuracy of the proposed method, as shown in Figure 56 we see a good level of performance for skin type groups 1 and 3, whilst a lower level for group 2. The high level of successful hit rate performance for group 3 is explained by the presence of only 1 test subject within this group. However the performances of group 2 are of greater interest. We can note that subjects in group 2 are just as likely to be placed within skin type group 1. This performance can be explained by the centre of mass locations of the added subjects as previously seen in Figure 55. A number of these lie across the boundary between the two skin type groups. The effect of this performance will be discussed later on within the combined algorithm result section.

4.6.4 *Illumination estimation performance*

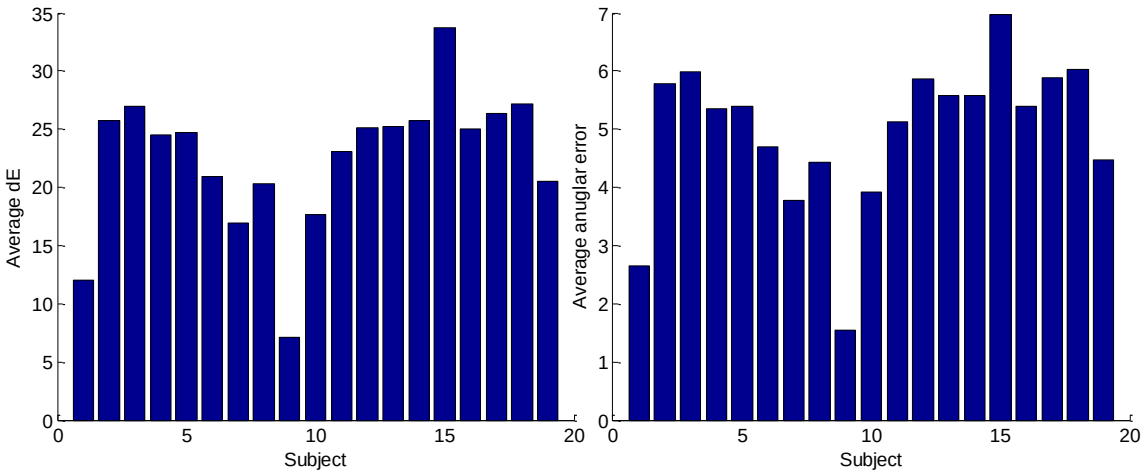
Illumination estimation is calculated both in terms of angular error and dE of the chosen illumination from the actual. RGB values to represent each illumination were measured from white patches in the sRGB version of the hyperspectral images. We consider performance per skin type, per subject, per illumination and across the known and unknown subjects to consider the overall performance of our proposed method. We also investigate where, for each illumination, the most common estimation choices lie along the daylight axis to gauge reasons for the performance of the algorithm. When estimating illumination estimation performance the algorithm was fed with the skin gamut of a random subject (from the full set of 19 subjects) under a random daylight illumination (from the set of 20), this is supplied alongside the correct skin type group for the subject. This allows us to test the illumination estimation performance under ideal conditions.



a) b)

Figure 57 - The illumination estimation performance of the algorithm on an illumination-by-illumination basis in terms of a) dE, and b) angular error.

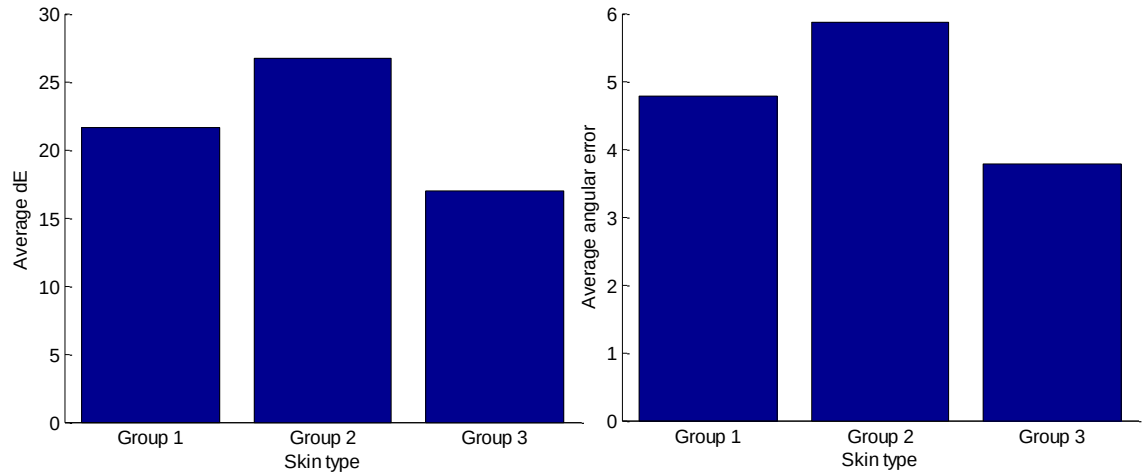
Figure 57 – Figure 60 illustrate the overall illumination estimation performance of our proposed method. Figure 57 shows that performance for illumination estimation improves as we go along the daylight axis from D40 to D250 (illuminations 1-20).



a) b)

Figure 58 - The illumination estimation performance of the algorithm on a subject-by-subject basis in terms of a) dE, and b) angular error.

Figure 58 illustrates that performance across subjects, and thus skin chromaticities, can be considered to be quite uniform. This in conjunction with the decent level of performance shows the validity of the method.



a) b)

Figure 59 - The illumination estimation performance of the algorithm on a skin type-by-skin type basis in terms of a) dE, and b) angular error.

Figure 59 confirms this skin type performance uniformity across all illuminations. As such it can be seen, in Table 5, that the illumination estimation method on its own results in a decent level of performance with a mean angular error of 5.05 and mean dE of 22.92.

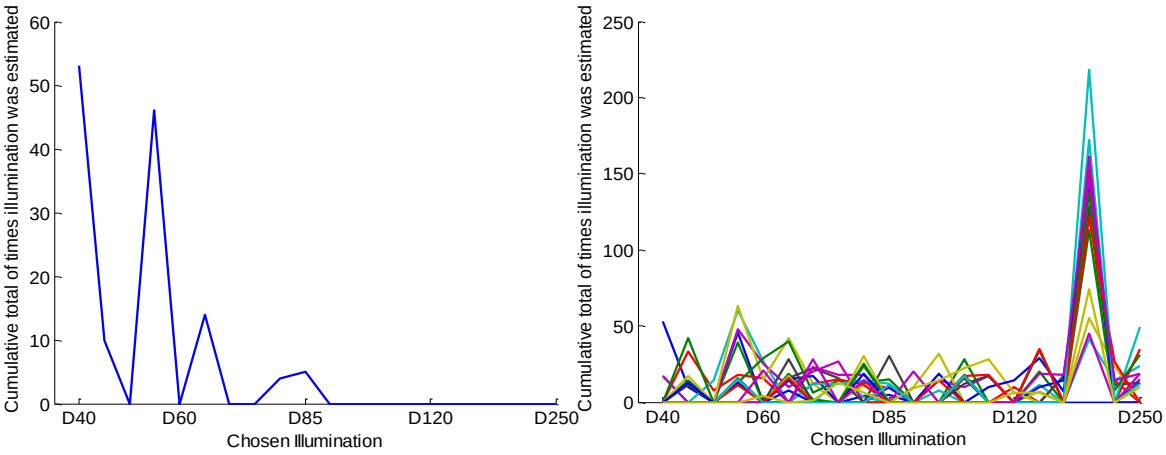
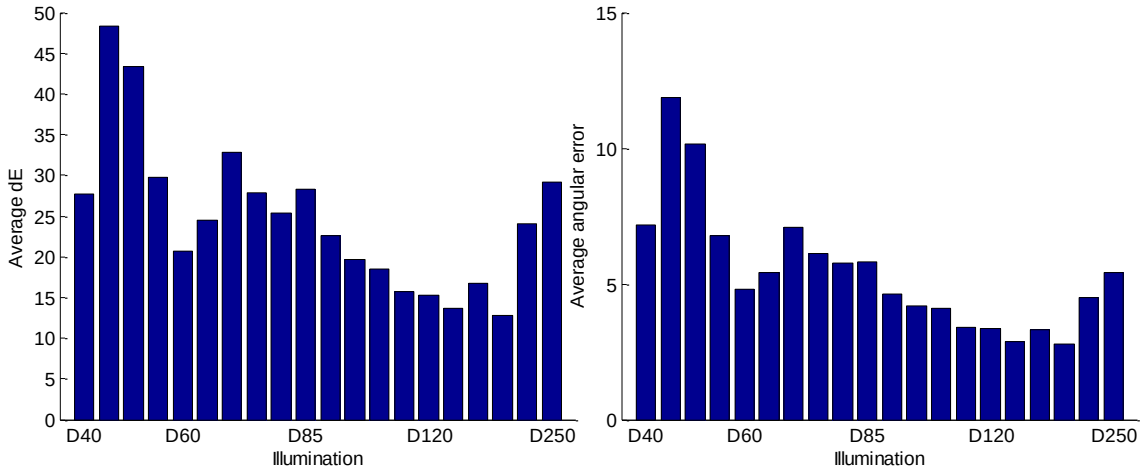


Figure 60 - Illustration of the most commonly estimated illumination for each actual illumination for a) D40 (Illumination #1), and b) across all illuminations.

These previous figures may illustrate the abstracted overall illumination estimation performance, however to further investigate performance we must consider where illumination estimation errors occur. This will aid us to further improve the algorithm in the longer term. Figure 60a shows that if we consider illumination 1 (D40) we can see the vast majority of illumination estimates are correct, however illuminations 4 and 6 can also be seen as quite likely to be chosen. Figure 60b shows that over the entire range of 20 illuminations illumination 18, D150, has a high likelihood of being estimated. This spike in illumination estimation choice will have a large effect upon overall performance. The reason for this performance must be considered. If we look back to Figure 45f we can see a stationary point in the maximum RG hull point across skin type groups 2 and 3.

4.6.5 *Overall algorithm performance*

Whilst the previous section proves that our proposed method exhibits a decent level of performance the overall algorithm, with skin type estimation working in tandem with illumination, must be tested. The figures below showcase the performance of the algorithm when the skin type estimation and illumination estimation methods are combined into one algorithm. This means that the algorithm is fed with the skin gamut of a random subject, again from our set of 19, taken under one of the 20 daylight illuminations. The algorithm then estimates the skin type, and then using this estimates the illumination using the reference features based upon the estimated skin type.



a) b)

Figure 61 - Comparing the results of the combined algorithm across illuminations in terms of average a) dE, and b) angular error.

Figure 61 shows that the combined method exhibits the same performance characteristics across the daylight locus of illuminations is the same as previously noted in Figure 58. The level of angular error performance when considered in its own right here is, whilst not as high as other methods have previously reported, a decent level. The kurtosis of the angular error distribution, as described by the statistics above, show that the majority of errors lie within the 0° - 12° region.

Algorithm	Error Measure	Min.	10th perc	Median	Mean	90th perc	Max.
Illumination estimation (known skin)	dE	0.00	0.00	18.95	22.92	54.77	75.10
Combined (skin and illumination estimation)	dE	0.00	2.96	21.50	24.69	56.91	87.65
Illumination estimation (known skin)	Angular Error	0.00°	8.54E-07°	3.57°	5.05°	12.60°	16.39°
Combined (skin and illumination estimation)	Angular Error	0.00°	0.49°	4.27°	5.44°	12.60°	19.63°

Table 5 - Comparison of algorithm performance for both the known skin and overall combined algorithm.

In order to adequately and accurately gauge the performance of colour constancy algorithms open access standardised image datasets are used. However due to the nature of our method we must utilise a dataset created by our imaging system.

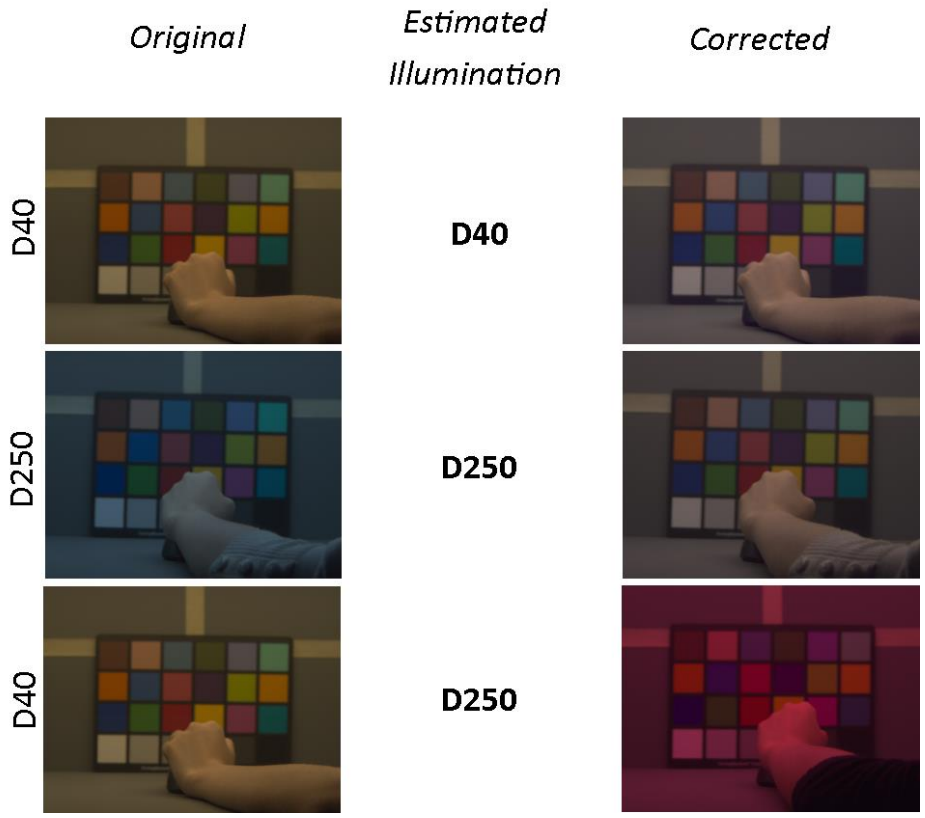


Figure 62 - Examples of the effect of correct and incorrect illumination estimation.

4.7 Colour normalisation algorithm performance

In order to investigate the performance of the skin gamut based normalisation method we will compare the processed average RGB values from a number of patches in the background of the scene and compare the angular error between these and their measured canonical values under D65. We utilise a series of 12 patches on the MacBeth Colour Checker Chart for comparison, and compare the average angular error across illuminations and subjects for our method.

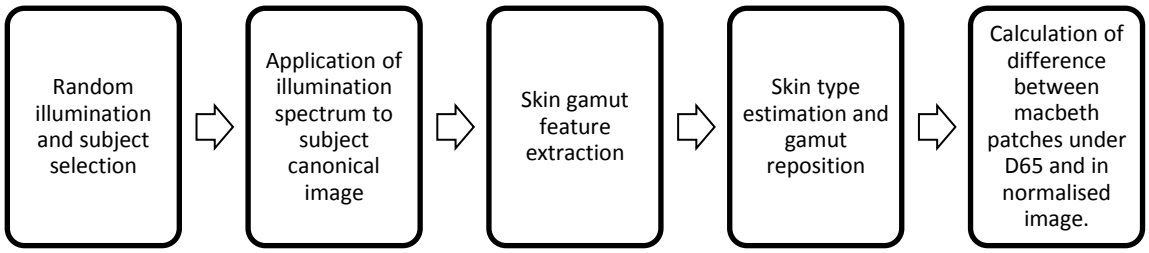


Figure 63 - Outline of testing methodology for the normalisation method

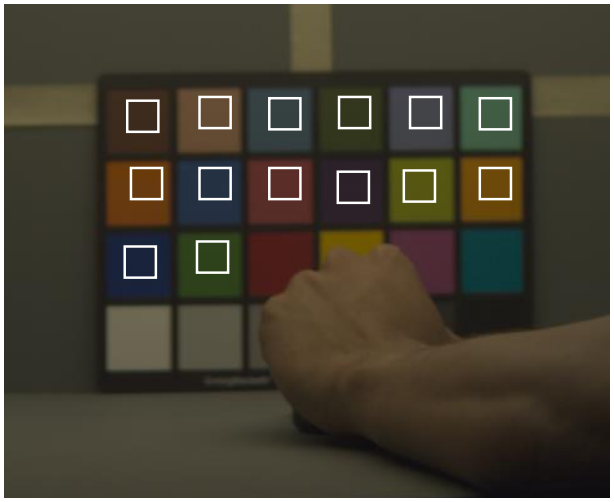


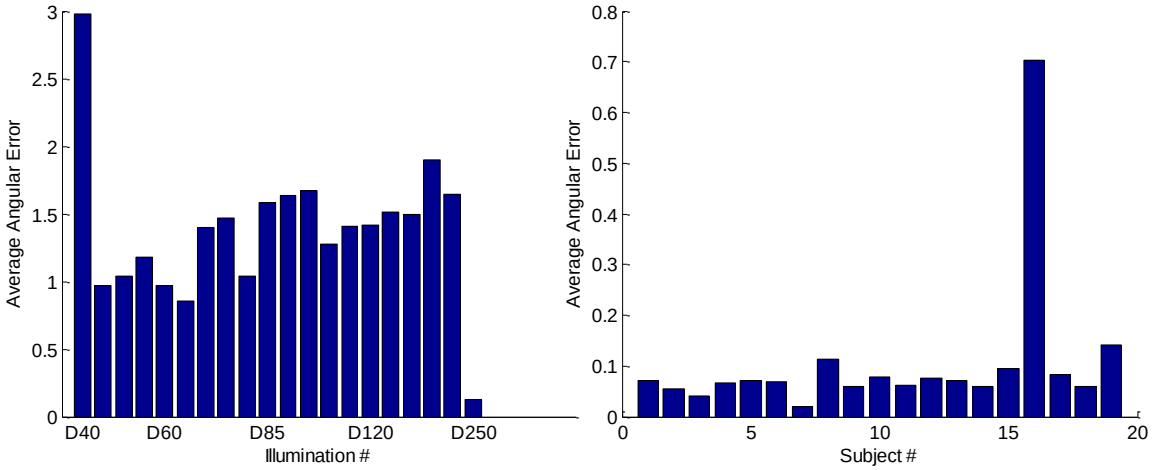
Figure 64 - Illustration of the colour patches used to gauge algorithm performance.

Due to the nature of displaying our hyperspectral images in RGB we must explain the conversion method. This will reinforce the validity of comparing patch values, on a subject-by-subject basis across the 20 daylight illuminations. For each subject image, after the application of a test illumination, the image is firstly normalised and then gamma ramped for display;

$$I_{sRGB} = \left(\frac{I}{\vee I}\right)^{\left(\frac{1}{2.2}\right)} \tag{38}$$

Where $\vee$ finds the maximum value in each channel.

As we utilise, on a subject-by-subject basis, the same canonical image prior to the application of a test illumination the maximal pixel value in each channel will be the same across illuminations. This means that across illuminations the normalisation and gamma ramping is consistent, and so patch values are directly comparable.



a) b)

Figure 65 - Viewing the average angular errors across test patches across a) illuminations, and b) subjects.

Algorithm	**Error Measure**	**Min.**	**10th perc**	**Median**	**Mean**	**90th perc**	**Max.**
Normalisation	Angular Error	0.09	0.35	0.91	0.88	1.32	2.29

Table 6 - Performance of the proposed normalisation method.

The performance of the method, in terms of RGB variance across the patches can be seen to be good in Figure 65 and Table 6, however it must be noted that this angular error performance is relative to that subject and skin type group only. Due to the method of using different skin types groups, as with the previous illumination estimation method, the visual results for two subjects under the same illumination may be visually different.

4.8 Performance of illumination estimation under chromatic illumination metamers

The illuminations utilised to create our dataset of human skin across illuminations were chromatic metamers with a high level of spectral accuracy when compared to the CIE standard daylight set of illuminations. These illuminations reproduce, to a certain degree, the natural illuminations encountered in the real world. However with the large increase in artificial lighting sources over the past 50 years the number of metameric light sources has increased massively. The overall emission

spectra of these light sources is largely related to their lighting group; such as fluorescent, halogen, incandescent and LED.

These light sources will produce chromatic metamers, however their spectra can be quite different. If the emission spectra of these sources are similarly broadband the resulting irradiances reaching the visual system will be consistent, and hence result in good CRI scores. However if we have instances of narrowband illumination metamers their CRI scores will obviously suffer.

Utilising the multiple channel LED lighting system outlined in [29] we can also create a number of different chromatic metamers for a given illumination chromaticity, both broad and narrowband. To a human observer, if we consider similarly broadband metamers, these illuminations will be indistinguishable from one another. However how would they affect the retrieved skin gamut features for our subjects? If human skin within cone-contrast space is a dependable cue for illumination estimation we would expect some level of invariance across similarly broadband illumination metamers. To investigate this, a number of D65 metamer spectra were generated, with a smaller subset of 10 being randomly chosen and captured using the hyperspectral imaging system. The spectra of these and our database D65 are compared below;

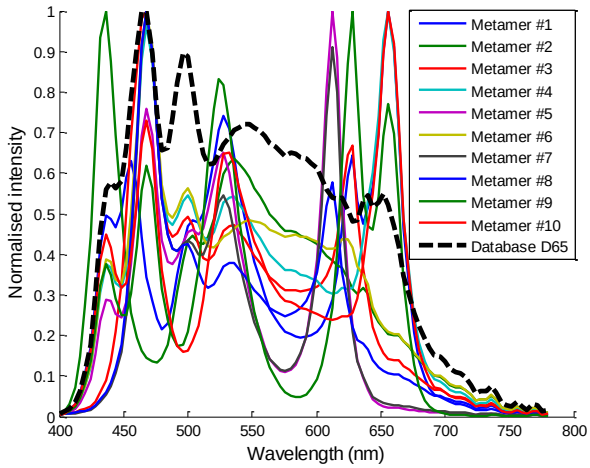


Figure 66 - Comparing the spectra of our database D65 and 10 chromatic metamers produced by the illuminator.

In order to gauge the performance of our method across these 10 metamers we tested our illumination estimation with the gamut features of these 10 D65 metamers. If our gamut features are solely related to CCT we expect a level of feature invariance in the skin gamut and can expect our method to work under all metameric light sources. However if our method is dependent upon the spectral distributions of the illuminations, which is expected, then our method will work only under the natural daylight spectra.

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Metamer	Spreading	.116	9	.013	1.695	.112
	Angle	203.161	9	22.573	.809	.610
	Area	.001	9	5.752E-005	1.023	.434
	MaxRG	.011	9	.001	49.147	.000
	MinRG	.001	9	9.383E-005	1.567	.149
	MaxBY	.008	9	.001	.745	.667
	MinBY	.076	9	.008	.734	.676
Skin Type	Spreading	.257	2	.128	16.923	.000
	Angle	1422.820	2	711.410	25.504	.000
	Area	.002	2	.001	16.074	.000
	MaxRG	.037	2	.018	744.666	.000
	MinRG	.004	2	.002	37.038	.000
	MaxBY	.135	2	.067	54.433	.000
	MinBY	.713	2	.356	30.913	.000
*Metamer * Skin Type*	Spreading	.001	18	4.220E-005	.006	1.000
	Angle	7.954	18	.442	.016	1.000
	Area	9.818E-005	18	5.454E-006	.097	1.000
	MaxRG	.002	18	.000	4.311	.000
	MinRG	.000	18	1.248E-005	.208	1.000
	MaxBY	.005	18	.000	.223	1.000
	MinBY	.008	18	.000	.041	1.000

Table 7 - Results of multivariate analysis on hand gamut features across the 10 D65 metamers.

Table 7 shows the results of multivariate analysis on the gamut features across the original group of 8 database subjects under our 10 D65 metamers and our database D65. What can be noted immediately is the maximum RG hull point varies significantly with the metamer [(*MaxRG*) $F(9,55)=9.147$, $p=0.000$]; there is also a significant interaction effect between metamer and skin type [$F(18,55)=4.311$, $p=0.000$].

With an end goal of illumination estimation and our approach we aim to differentiate different illuminations using skin gamut features, but by doing this we assume that chromatic illumination metamers will exhibit the same features. The statistics above show this not to be the case. It is nonetheless an empirical question whether, despite this variation in skin gamut features with specific metamer for a given illumination chromaticity, our method returns a reliable estimate of illumination chromaticity. To investigate this, the previous testing methodology was used once again to view the

known skin illumination estimation. This involved the random selection of a subject, from our overall set of 19, and the random application of one of the 10 D65 metamers or reference D65 illumination. Skin type was then first estimated, and illumination was then carried out

Algorithm	**Error Measure**	**Min.**	**10th perc**	**Median**	**Mean**	**90th perc**	**Max.**
Illumination estimation (known skin)	dE	0	6.01	38.46	29.79	41.72	62.14
Illumination estimation (known skin)	Angular Error	0°	1.62°	9.26°	6.69°	10.29°	12.54°

Table 8 - Viewing the performance of the algorithm with a set of D65 chromatic metamers

4.9 Summary

This chapter has illustrated a number of features of human skin when represented in cone-contrast space across a large range of illuminations. Firstly, that the gamut of human skin in cone-contrast space changes in a predictable manner across the daylight locus. Secondly that subject skin chromaticities can be easily grouped and estimated across the daylight range of illuminations. Due to the predictable nature of these gamut and skin group changes illumination estimates can be made to a good level of accuracy, with the greatest accuracy in the range of cooler daylight illuminations even with coarse skin type groupings and estimations. The results of this method are not easily comparable to other published results in other previous work.

Secondly findings have reinforced that the region within cone contrast space where average skin chromaticities exist under daylight illumination is relatively small, and as such allows for the conception of a skin gamut based normalisation algorithm. The performance of this algorithm can be seen to be very good, and in terms of complexity very simple. However the use of only 12 patches to convey the performance of the algorithm may lead to greater variation in other patches being missed.

Both of the developed algorithms have been based upon a great deal of information which has shown both novel and reinforced previous findings regarding the location and chromatic features of skin under daylight illuminations in cone contrast space. We have

shown that all skin chromaticity gamuts vary in a predictable manner across daylight illuminations in cone contrast space. Using this finding we have shown that an algorithm based upon using the image contents of human skin can be developed and achieve a good level of results. Further to this we have also reinforced the findings that human skin chromaticities exist within a small region of cone contrast space, as has been previously put forward in LMS [100] and CIELAB [85] spaces. The main aim of this chapter was to attempt to bridge human and machine colour vision by investigating the chromatic variation of known scene contents across illuminations. This has been achieved to a degree, and in terms of computational complexity is relatively simple.

However a drawback of our method is evident, in terms of the feasibility of implementation of the method in existing camera systems. The use of cone contrast space is complicated when considering the use of RGB camera sensors. XYZ calibrations for RGB sensors rely upon a known and constant illumination under which images will be taken, however the requirement of this knowledge is at direct odds to the challenge of colour constancy. As such methods such as [119], which also uses skin, but within RGB space, are easier to implement. However the importance of the use of the polychromatic nature of skin and other objects [15] has been overlooked in their conclusion with their consideration of single colour surfaces as areas of further interest, which has a number of associated problems [125, 126].

With this main deficiency in mind, future work will aim to implement aspects of the proposed methods within a more realistic and robust camera workflow. On top of this the hyperspectral data captured will also be investigated in order to uncover the most important wavelengths in skin reflectance functions that can give clues to the current illumination.

With all of the findings and contributions laid out above we must consider the original aims and objectives laid out within Chapter 2. The algorithms detailed within this chapter have succeeded in being designed with regards to the operation of the human visual system in terms of the utilised colour space. We have also shown that a polychromatic surface/ content can be utilised to base illumination judgements upon. By doing this we have also succeeded in bridging human and computer colour vision. The work laid out within this chapter also has another contribution; the development of a new large hyperspectral skin dataset in combination with the illumination spectra of 39 illuminations.

Chapter 5. Illumination Dependency of Colour Constancy

5.1 Introduction

Computer vision based colour constancy approaches aim to achieve the same degree of constancy across all illuminations. This is based upon the assumption that this is what colour constancy achieves within the human visual system. Recently however investigations into performance differences across different common and uncommon illuminations have become of interest. Achromatic adjustment based experiments have been used a number of times to investigate colour constancy within human subjects, but previous experiments have been carried out under a small number of illuminations, generally provided by placing filter in front of existing light sources, or using projectors as light sources. These methods, as they utilise sources with a relatively constrained spectrum, have only tested a small number of the commonly and uncommonly illuminations encountered by the human visual system on a day to day basis. As such our experiment firstly set out to test colour constancy across a range of these illuminations. Secondly the vast majority of current achromatic adjustment tasks are now performed through the use of emissive displays. These displays, natively, do not act like natural surfaces encountered by the human visual system. This can lead to experimental biases being brought into experiments if these devices are not incorporated in the correct manner. Further to this, their emissive nature means that reflectance and illumination are confounded in a manner different to within a natural scene. This is further compounded by the limited spectral reproductive ability of most monitors, and poses a challenge to prove colour constancy does indeed recover surface reflectances.

As such we set out to utilise, and verify the use of, a new type of display, the Mirasol reflective display. This has been designed to behave much more like a natural surface in a scene, and so should be easier to incorporate into a wider array of more realistic testing scenarios.

The main contributions of this chapter are to firstly show the applicability of reflective displays for human vision studies. The chapter will then move on to investigate the illumination and scene dependence of the overall human colour constancy process.

5.1.1 *Psychophysical experiment introduction*

Psychophysical testing of human subjects to prove that colour constancy exists, and to discover to what degree it exists and its limits have generally been undertaken using a small number of experiments.

The method of delivering achromatic adjustment experiments has changed over the years. Initial experiments were undertaken using physically changing stimuli, such as paper [3]. These methods were surpassed in time with the advent of computer display technology, starting with CRTs, then moving onto LCDs and now to OLED displays. Each of these display technologies have their own display and colour characteristics, but can all be grouped into the emissive display technology bracket. This definition relates to the way in which they display stimuli; however most natural objects seen in day to day life do not emit light to display their chromaticity. As is fundamental to this body of work, if we have an object in a scene it can be considered that at a single point on its surface it has a spectral reflectance. This combines with the ambient illumination to result in a spectral irradiance, which is infinitely variable with regards to changes in illumination and illumination intensity. However if we place a typical emissive display into the same location in the scene and try to replicate the chromaticity of the object, its emitted spectral function will not exactly match that of the original object, but will match in terms of chromaticity. This is due to metamerism. As such by using an emissive display in the same scene we have to accept that these characteristics of the display will also severely limit its use, firstly those of the dynamic range as illumination intensity varies, and secondly, linked to the first, the chromatic characteristics of the gamut at different intensities.

These display characteristics can cause problems if we do not constrain the illumination and test scene characteristics. For example, it is likely that, in terms of luminance, if there is a large difference between the luminance of the screen and the average luminance of the scene then the screen will not be perceived to be part of the scene. This can then lead to the subject ignoring the scene conditions and automatically achromatically adjusting the screen patch to the white point of the screen, rather than the illumination. Secondly the gamut of colours that the screen can show varies along with the luminance component. Thus if the screen has a small gamut across a small range of luminances then both the scene and illuminations are required to be constrained to maximise the chromatic choices available to the subject during adjustment. This can result in good matches being made, but within unrealistic illumination and scene

conditions when compared to the real world. These problems exist to differing degrees with existing display technologies, and also within different examples of these technologies. They will also affect, to different degrees, future OLED based displays and HDR LCDs. LCDs also have other problems relating to their, generally, non-black black levels and the non-uniformity of the back light across the display at this output.

Paper based methods, where the subject can view a number of different paper samples, in some manner, bypasses this problem. For example as illumination intensity varies the papers' radiance also varies. The use of a stimulus, which then also fits naturally within the scene, bypasses the other gamut intensity challenges mentioned previously.

Emissive displays have long been the norm of the display industry and recent advances have resulted in the development of reflective displays. Whereas emissive displays show a colour by emitting light, reflective displays show colour by replicating, to a simpler degree, the manner in which natural objects are seen to have a colour. This is by absorbing and reflecting different spectral regions, with regards to their static reflectance.

This manner of operation means that the display interacts much like a natural surface, such as a paper patch, and as such avoids the previously described challenges encountered when using emissive displays within a test environment, enabling use within much more realistic test conditions.

Due to the recent development of these displays and the novelty of their use within vision experiments we carried out a number of characterisation tests on the display. This was in order to reaffirm the hypothesis that this display type should, up to a certain degree, perform much like a natural surface within a test environment.

5.2 Reflective Display Characteristics

A number of characterisation tests were carried out to model the changes in gamut as both illumination intensity and spectral power distribution were varied. The changes in illumination intensity allowed us to see any changes in both overall chromatic and luminance characteristics of the display under a fixed illumination. If the display does perform, theoretically, as a natural surface we would expect a linear relationship between measured display luminance and illumination irradiance, and no change in the chromatic properties of the display (i.e. gamut size and shape).To test this, the screen was illuminated by the same illumination spectra at two different irradiance

levels, one at double the level of the other (at a relative 100% and 50% respectively) using the illumination system shown within [29]. Under these two irradiances the gamut of the Mirasol display was recorded at intervals of 30 in RGB space, for all possible RGB triplets at this resolution level;

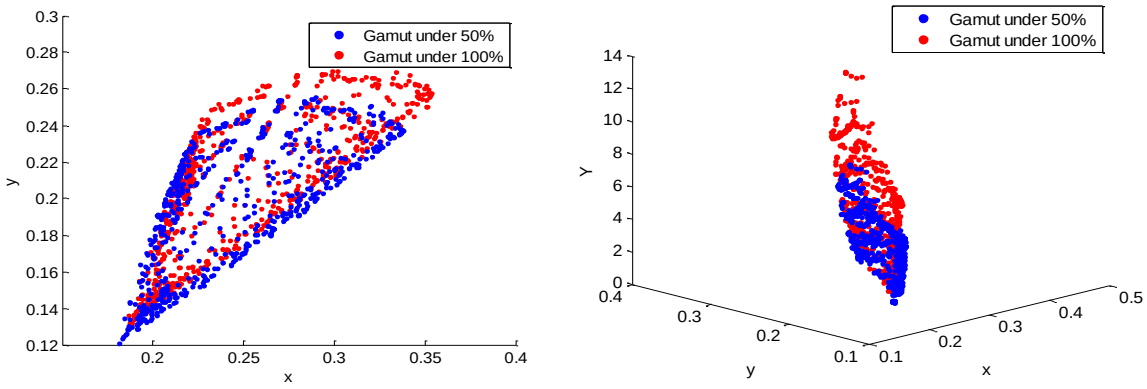


Figure 67 - 1931 CIE Yxy gamut of Mirasol display under illumination at 100% and 50%; a) xy plot, and b) Yxy plot

Figure 67 above shows the changes in the gamut of the Mirasol display, in CIE Yxy space, under two illuminations, created by the same weighted combination of LED channel outputs at two different absolute levels. Figure 67a shows that there are small changes to the maximal chromatic gamut from the display in the chromaticity plane. Figure 67b shows the variation in gamut in the luminance (Y) direction, which is much larger. It can be seen, and verified numerically, that the volume of the gamut has increased:

	100% Intensity (cd/m2)	**50% Intensity (cd/m2)**	**Scale factor (with respect to 50%)**
Illumination big Y	14.9700	8.6520	1.73
Gamut volume	0.0505	0.0257	1.97

Table 9 - Variation in Mirasol peak Y and gamut volume under two irradiances of a standard illumination.

The scale factor of the increases, in both peak luminance and gamut volume, follow the effective doubling of illumination irradiance. This shows that the Mirasol display, to a large degree, replicates the changes in object irradiance related to the ambient illumination conditions. This illustrates that the Mirasol display should easily be viewed as a natural part of a test scene under any illumination and its gamut, under a given illumination, will remain constant, even with scene irradiance changes.

As was noted previously and has been stated a number of times [8, 127], different display technologies also have their own black level characteristics. The advent of LCD technology and its use within vision experiments was challenged by their original, relatively, poor black level performance, in terms of luminance level and uniformity across a panel when compared to that of CRT displays. OLED devices [128] have true black levels as they emit no light when displaying black. CRTs also have a closer to true black performance when compared to tube backlit LCDs, resulting in generally better contrast levels. In theory reflective displays such as the Mirasol display should also have a true black level under all illuminations, however the performance of this depends upon the characteristics of the grating over the panel. The variation of this black level in relation to ambient illumination was of great interest, as this would again confirm similarities between the manner in which natural objects and the Mirasol display work. We utilised the following illuminations for testing in conjunction to the two aforementioned;

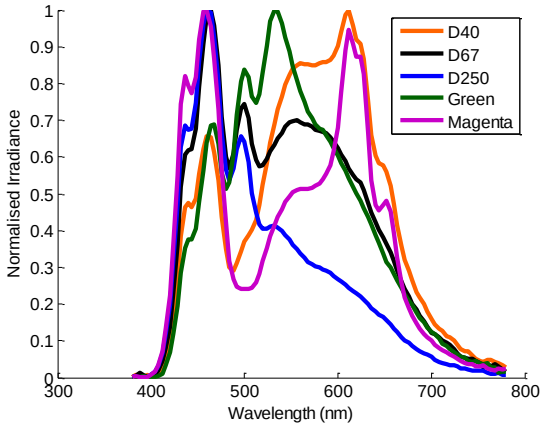


Figure 68 - Experiment test illumination spectra

We then proceeded to collect the spectral irradiance and colourimetric data for the display under these illuminations using a PR-650 spectrophotometer;

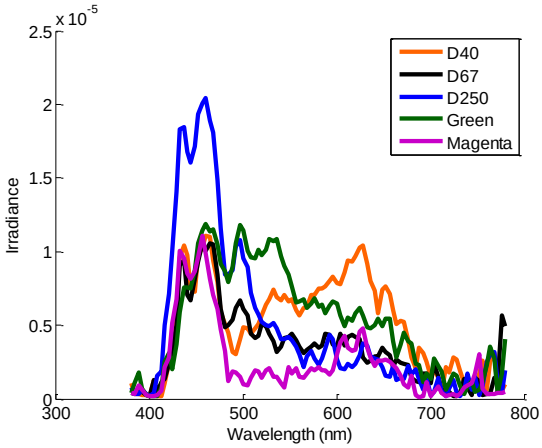
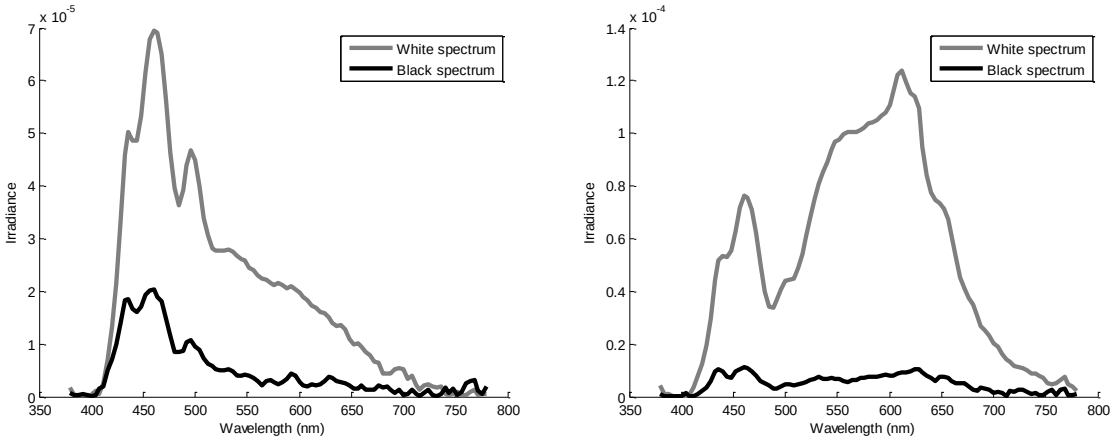


Figure 69 - Mirasol display black level across test illuminations

	Illumination			**Measured Screen Black**		
Illumination #	Y	x	y	Y	x	y
D40	16.8300	0.3902	0.3805	0.5029	0.3477	0.3085
D67	16.7700	0.3160	0.3355	0.3002	0.2723	0.2580
D250	16.9600	0.2452	0.2587	0.3172	0.2078	0.1776
Green	16.7200	0.3258	0.4135	0.5655	0.2870	0.3417
Magenta	17.1100	0.3469	0.2798	0.1583	0.2708	0.1726
All Channel (100%)	44.5600	0.2508	0.2261	1.4540	0.2053	0.1488
All Channel (50%)	21.8100	0.2508	0.2261	1.0430	0.2146	0.1626

Table 10 - Yxy of the Mirasol display black levels across test illuminations

As can be seen from Figure 69 and Table 10 the black level does vary in relation to ambient illumination. This immediately shows that the display is not consistent in terms of chromaticity, but is in terms of luminance levels. This consistency in luminance levels reduces the perceptual effect of changing black level chromaticities. Figure 69 shows the reason for this inconsistency in terms of irradiance spectrum. As can be seen, the black level irradiance spectrum across the 5 test illuminations varies mostly within the blue region of the visible spectrum. This means that in effect the blue primary absorption is not as uniform as across the other regions of the visible spectrum. As such under illuminations with a greater blue spectral component the black level will reflect this blue component, and be perceived to have a blue tint. This can immediately be noted when comparing the black levels for Illuminations 1 and 2, with respect to their irradiance spectra:



a) b)

Figure 70 - White and black level spectra of the Mirasol display under a) D250 and b) D40.

A further way in which to compare the performance of the Mirasol display in comparison to natural white surfaces is to investigate its white level. This is also an important consideration when utilising a reflective display in a psychophysical experiment. Traditional emissive displays have their own characteristic white point. However in theory reflective displays under an illumination will have a white point equal to that of the illumination. This has an effect upon the method of calculating quantitative colourimetric measures from experiments, and as such any shifts away from this assumption were required to be quantified. To do this we measured the irradiance of the screen when displaying white ([255 255 255]) under the five test illuminations and compared it to the illumination spectra:

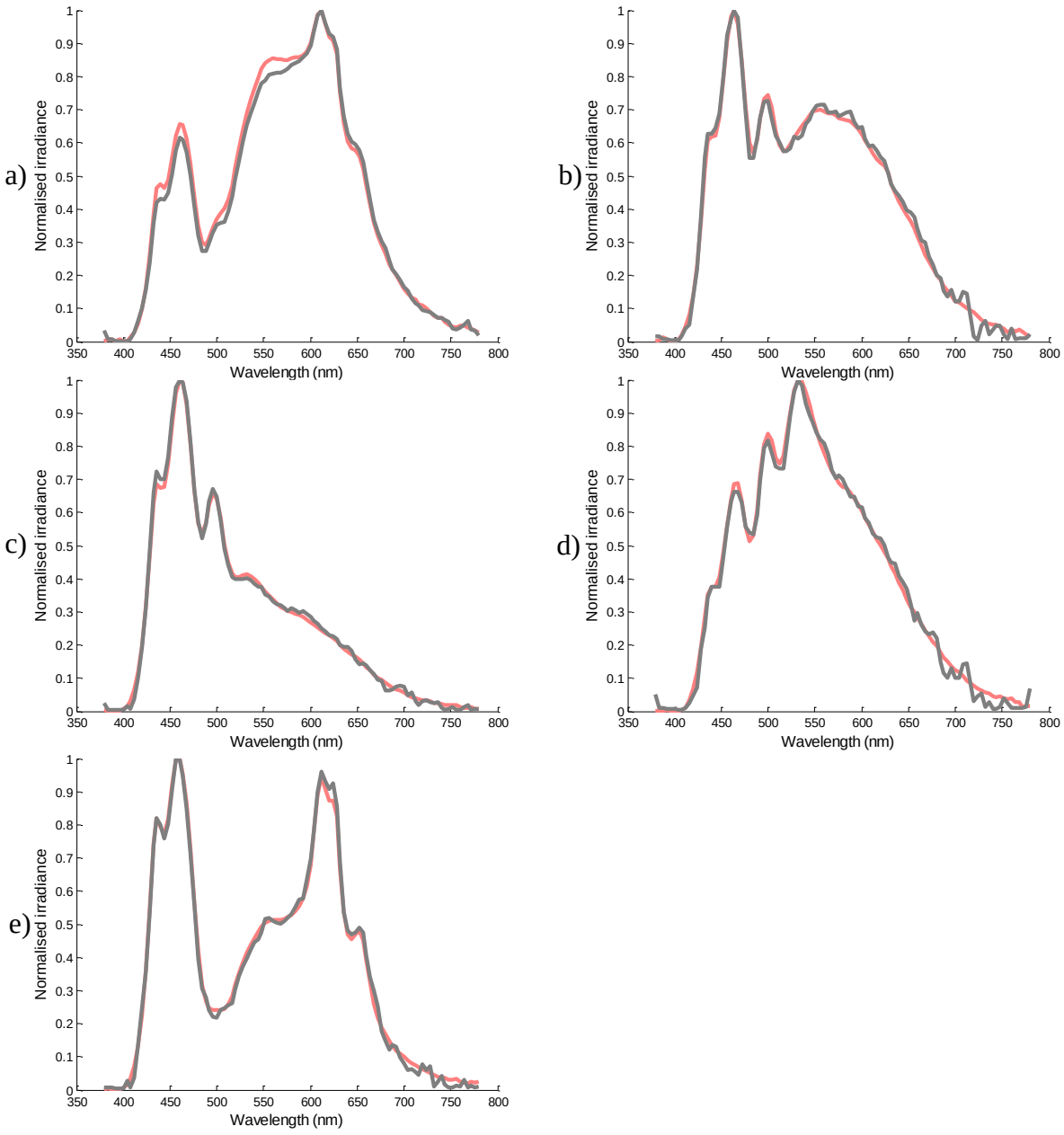


Figure 71 - Illumination spectrum and Mirasol display white level spectrum for test illumination a) D250, b) D40, c) Green, d) Magenta and e) D67

A perfectly white surface will, theoretically, reflect the ambient illumination. As such Figure 71 above shows that the irradiance of a white patch being displayed upon the Mirasol display closely mirrors that of the illumination for all illuminations. This again shows that the Mirasol display can be effectively considered to behave like a natural white surface. However we must also validate this hypothesis with colourimetric data:

	Illumination WP			Measured Screen WP		
Illumination #	Y	x	y	Y	x	y
D40	16.8300	0.3902	0.3805	3.0440	0.3954	0.3810
D67	16.7200	0.3258	0.4135	1.8080	0.3296	0.4149
D250	16.9600	0.2452	0.2587	1.8210	0.2466	0.2553
Green	17.1100	0.3469	0.2798	1.8770	0.3509	0.2796
Magenta	16.7700	0.3160	0.3355	1.8080	0.3184	0.3342
All Channel (100%)	44.5600	0.2508	0.2261	8.6520	0.2428	0.2113
All Channel (50%)	21.8100	0.2508	0.2261	4.4780	0.2455	0.2207

Table 11 - 1931 CIE Yxy differences between illumination and Mirasol display white points under test illuminations.

Table 11 reinforces the spectral findings from earlier, in that, the screen white point does not differ to a large degree from the illumination. However the differences that exist do mean that the best possible achromatic match by a subject cannot match the illumination. Therefore, in the following experiments we will use the screen whitepoint under the test illumination as the reference for the achromatic match under that illumination.

5.4 Psychophysical experimental design and implementation

This section will give an overview of the equipment and experimental design used for the investigation.

5.4.1 *Equipment*

The adjustable patch was placed within a box of dimensions H*W*D containing one of two possible linings, that of a uniform grey or a Mondrian. The patch achromatically adjusted by subjects was a 3cm*4cm section on the Mirasol based Kyobo e-reader. The Kyobo e-reader has a 5.7 inch (14.47cm) screen of resolution 1024*768. This patch was displayed using a developed Android based application and controlled over Wi-Fi using Matlab and an Eclipse IDE based Java server on a main computer. Subject inputs were entered using a mapped controller connected to the main computer as in [129].

5.4.2 *Experimental design*

The experiment utilised a staircase based design, in which subjects were given two possible axes of movement. These axes were along the red-green (RG) and blue-yellow (BY) directions in RGB space. RGB space was chosen and deemed acceptable for use due to the reflective nature of the display and its characteristics described previously. RGB values were calculated from their RG-BY counterparts using the following formula:

$$\boldsymbol{RGB} = \boldsymbol{A} \backslash \begin{bmatrix} \boldsymbol{RG} \\ \boldsymbol{BY} \\ \boldsymbol{I} \end{bmatrix} \tag{39}$$

Where *RG* and *BY* are the stimulus co-ordinates, of range -100≤ value≤100, I is the set patch intensity (to achieve iso-luminance), and A is defined as:

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{1} & \mathbf{-1} & \mathbf{0} \\ \mathbf{-1/2} & \mathbf{-1/2} & \mathbf{1} \\ \mathbf{1/3} & \mathbf{1/3} & \mathbf{1/3} \end{bmatrix} \tag{40}$$

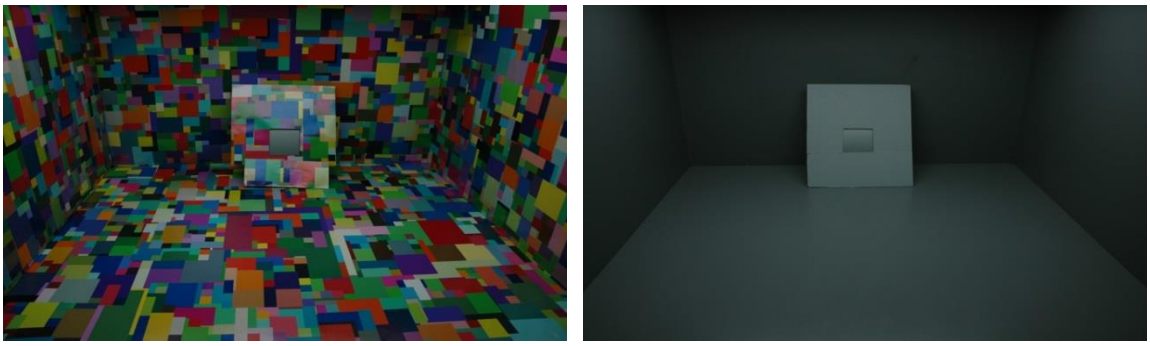
The staircase design incorporated 5 staircase step sizes, and the initial RG-BY of the patch was randomised at the start of each illumination trial. Each experimental session, for one box condition, incorporated 5 trials, in which each trial related to a different test illumination. Across subjects the order of test illuminations was randomised. For a given trial, subjects were allowed an unrestricted length of time for achromatic adjustment. Six reversals at a given staircase step size then changed the step size to the next smaller size, or ended the trial, if the smallest step size had already been reached. Three staircases were run concurrently for each trial.

5.4.3 Subjects

A total of 9 subjects were tested, 6 male and 3 female, all having tested colour vision normal on the Farnsworth-Munsell test. 8 of these participants were naïve, with the non-naïve subject being the author. Subject age ranged from 23 – 50.

5.4.4 *Experimental conditions*

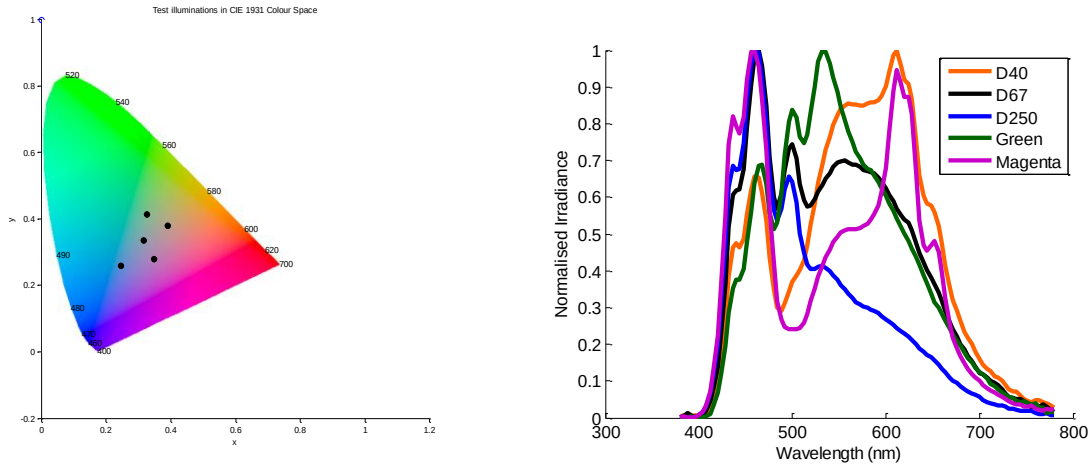
Two main box environment conditions were tested, in addition to 5 different illumination conditions. The two environment box conditions referred to different box linings; the first being a Mondrian style lining, and the other being a uniform grey box lining. These two conditions under standard D65 illumination, with the Mirasol display in test position, are shown below:



a) b)

Figure 72 - Test box under D65 with a) Mondrian-like, and b) grey conditions

Five different illuminations were also selected in order to test different hypotheses relating to the performance of colour constancy under common and less common illuminations. 3 of these illuminations lay upon the CIE daylight locus at the positions, and were broadband metamers, of D40, D67 and D250. This spread was chosen to see if human colour constancy performed to the same degree across a large spread of the daylight locus. The two remaining illuminations were chosen to be orthogonal, and thus similar in correlated colour temperature, to D67, whilst being illumination spectra that are not commonly encountered by the visual system. These were chosen to determine which illumination characteristics influence the extent of colour constancy: CCT, the commonness of occurrence or a combination of both. The locations of the five test illuminations and their spectra are shown below;



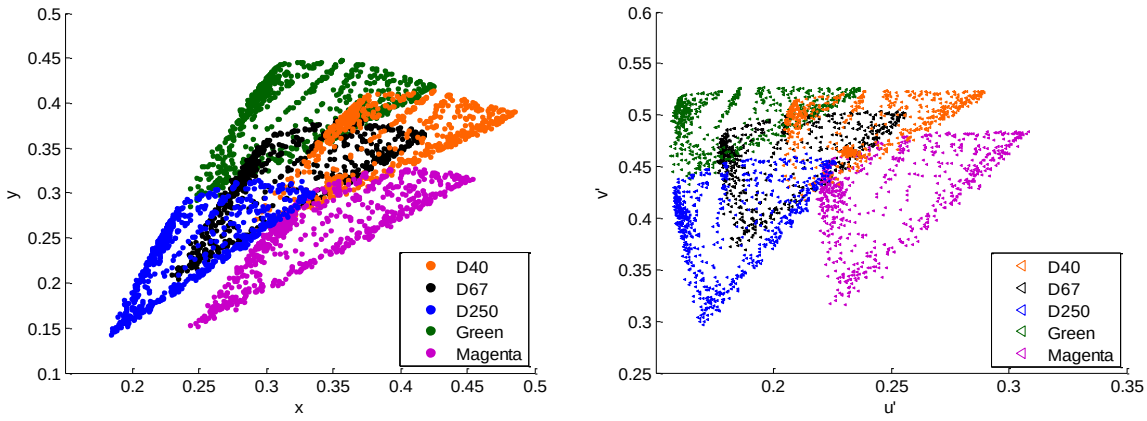
a) b)

Figure 73 - The illuminations used in the investigation, in terms of a) 1931 CIE Yxy location, and b) illumination spectrum.

Illumination	CCT (K)	x chromaticity	y chromaticity
D40	4000	0.3902	0.3805
D67	6700	0.3160	0.3355
D250	25000	0.2452	0.2587
Green	6700	0.3258	0.4135
Magenta	6700	0.3469	0.2798

Table 12 - Test illumination details

As previously stated the white point of the Mirasol display varies according to illumination whitepoint, the movement of the Mirasol gamut under the test illuminations in CIE 1931 xy space and Lu'v' are shown below;



a) b)

Figure 74 - Mirasol display gamut under the 5 test illuminations in a) 1931 CIE xy, and b) Lu'v' spaces

5.5 Results

A number of different metrics were utilised to provide a quantitative analysis of subject performance. This included the chromatic dE between subject achromatic choice and screen whitepoint for each trial. The chromatic, rather than full, dE metric was utilised due to the characteristics of the gamuts under the test illuminations and the resulting reduction in intensity resolution for the subject using the staircase method. In order to record colourimetric matches the irradiance of the RGB achromatic matches were recorded under their corresponding test illumination using a PhotoResearch PR-650 spectrophotometer.

Out of the maximum number of staircase results for each illumination under one of the two box conditions (3 staircases per trial * 9 subjects = 27) we had to disregard a small number of bad matches, where the staircase method did not converge. This resulted in there being 26, 24, 26, 27 and 25 convergent staircases for illuminations 1 – 5

respectively. These valid Yxy matches were then averaged for each subject and each illumination, and then averaged across subjects for both box and illumination conditions. The figures below compare the average subject match to the screen whitepoint under the relevant box condition. It can be noted that a small difference between the screen whitepoint under each condition does exist. Figure 75a also shows the large movement of subject matches between illuminations, showing a good level of adaptation to the scene and stimuli. The large movement of matches across illuminations, which in Figure 74a, have gamuts which overlap also reinforces this conclusion. It should be stated here that the whitepoints for each of the a*b* plots in Figure 74a and Figure 75b is that of CIE D65 for each test illumination.

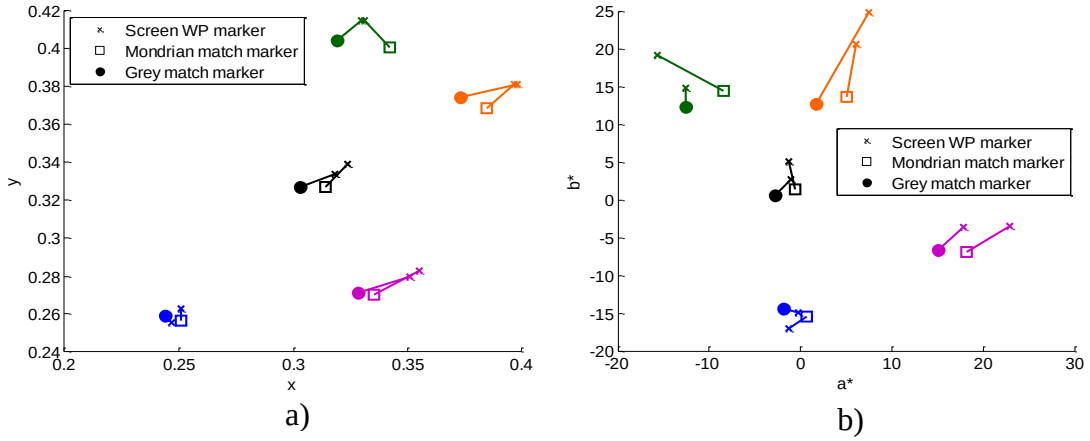


Figure 75 - Comparing the average subject matches under the 5 test illuminations and 2 test conditions to screen white point in a) 1931 CIE xy, and b) CIELAB spaces

The chromatic dE scores, relative to the screen whitepoint, for these average responses were then calculated using D65 xy co-ordinates for the colour space transform. The results are shown in Figure 76;

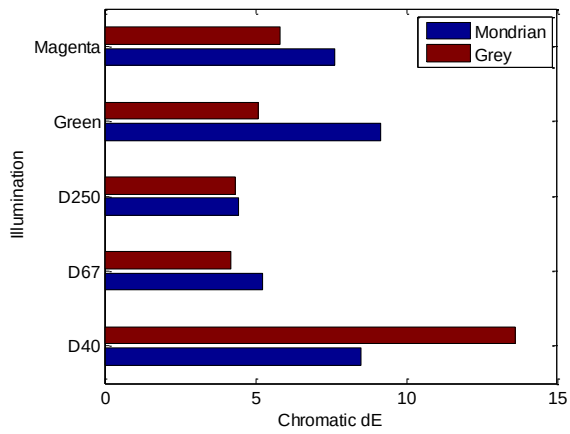


Figure 76 - Average ΔE values across test illuminations and conditions.

An ANOVA carried out on the average subject match chromatic dE scores revealed that there was the effect of condition on results was not significant ($F(1,8)=0.576$, $p=0.473$). However illumination had a significant main effect: $F(4,32) = 15.165$, $p<0.000$ A significant condition – illumination interaction was also uncovered: $F(4,32) = 10.106$, $p<0.000$.

However the dE metric is solely an absolute perceptual measure of the distance between the average achromatic match and the true whitepoint of the screen. It does not convey any general shift in direction of achromatic choice under each illumination, or box condition. This is useful in trying to uncover ways in which the human visual system may attempt to achieve colour constancy under different conditions. As such analysis on the angle in La*b* space, with the b* axis as the 0° reference, between average achromatic match and screen whitepoint was also carried out. These results are shown in Figure 77:

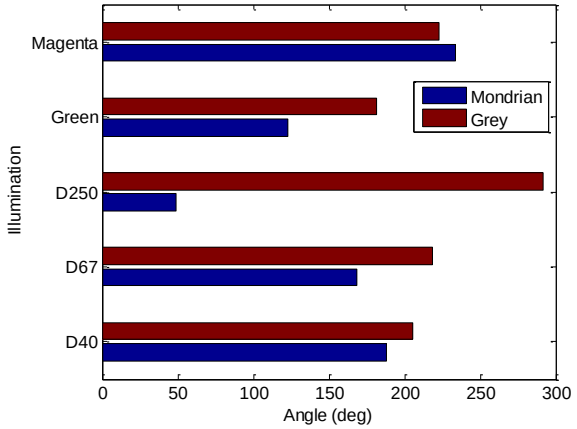


Figure 77 - Angle between average match and white point in CIELAB space across test illuminations and conditions.

ANOVA analysis on this a*b* angle data revealed a significant effect of condition on the angle between average subject-by-subject match and screen whitepoint. $F(1,8) = 14.649$; $p=0.005$). There was no significant main effect of illumination and no significant interaction effect of condition – illumination. However, in addition to the effect of illumination it can be seen that the match angles vary to a larger degree within the Mondrian condition. To investigate, a paired sample T test was run across the Mondrian angle results. This showed that illuminations 1 and 4, 2 and 3, 2 and 4, 2 and 5 along with 4 and 5 are significantly different.

To investigate the condition a*b* angle effect further, the underlying a* and b* differences between WP and match (screen value – match value) were investigated:

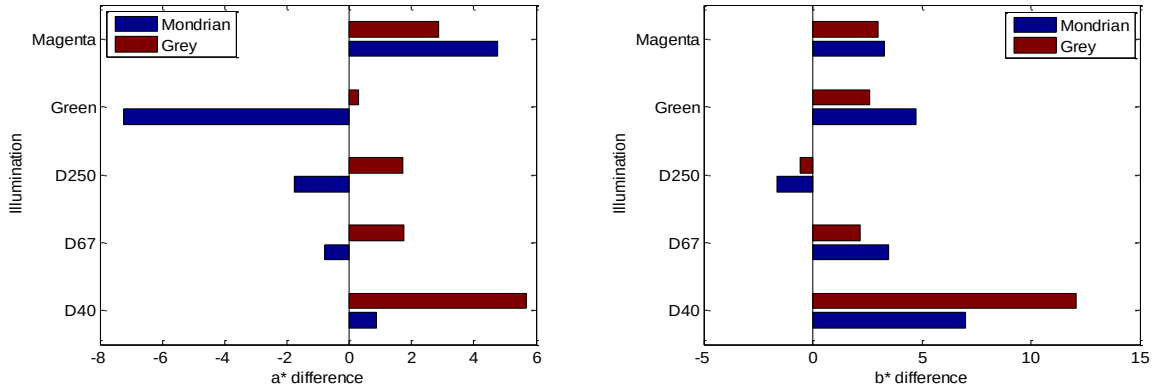


Figure 78 - Illustration of the CIELAB component distances between average match and screen white point across illuminations and conditions for a) a*, and b) b* components.

Figure 78 above shows the variation in a* and b* difference across the test illuminations. Analysis (ANOVA) on all subject average a* difference values showed a significant effect of condition ($F(1,8) = 35.830$, $p<0.000$), a significant main effect of illumination ($F(4,32) = 24.085$, $p<0.000$), and a significant condition – illumination interaction effect ($F(4,32) = 11.399$, $p<0.000$).

However the same analysis carried out for the b* differences showed no significant effect of condition. A significant effect of illumination ($F(4,32) = 47.556$, $p<0.000$), and a significant condition – illumination interactive effect ($F(4,32) = 5.967$, $p=0.001$) were found however.

Furthermore, the b* difference can be noted to have a decreasing value as CCT decreases across both conditions. The decrease in b* difference across the D type illuminations of D40 – D67 – D250 is roughly matched in value by the two illuminations orthogonal to D67, where these two illuminations have b* difference values within the same region as D67. This points towards a pattern of BY fine tuning in relation to illumination CCT.

In practical terms this means that, along the b* axis, matches for warmer CCT (yellow) illuminations (D40) are cooler (more blue) than the actual WP. Whilst for cooler (blue) illuminations (D250) the matches are warmer (more yellow) than the actual WP.

This pattern is also repeated when viewing the average a* differences across illuminations, except with differences between the two conditions. This difference may

point towards a difference in the cues within the scene that are utilised by the visual system.

5.6 Discussion

The use of a reflective Mirasol display with this experiment was a novel choice aimed at trying to incorporate and utilise a controllable display that had characteristics comparable to real reflective stimuli under illumination. Both investigations into the performance of the display, and its performance, including subject matches, under these illuminations showed the suitability of the display to use in vision experiments. The performance of the display under illuminations, in terms of gamut (Yxy), white point and spectrum was shown to be equivalent to natural surfaces, such as paper.

From experimental results, the variation of both illumination and box condition during enabled a number of different findings to be uncovered. In terms of illuminations, the inclusion of three daylight illuminations alongside two illuminations, orthogonal to D67 on the daylight locus, enabled human performance across these two types of illuminations to be gauged. The variation of dE performance across illuminations ($F(4,32) = 21.2934$, $p<0.000$), especially across the daylight locus was interesting. Performance along the daylight locus, across both conditions, was seen to decrease towards the warmer CCT end. This has been noted previously [22, 24, 25]. The results from subjects under the two uncommon non daylight illuminations were also of great interest. It can be reasonably imagined that illuminations along the daylight locus are more commonly encountered by people. As such an initial hypothesis was the expectation that dE performance would be better under the daylight illuminations in comparison. However, the performances under both orthogonal illuminations were different across conditions, and to each other. On average performance under the green illumination (illumination #3) was superior to under the red illumination (#4), and was noticeably better than performance under D40 (#2). This illustrates that colour constancy performance is not necessarily easy to estimate solely in terms of the commonality of illuminations being encountered.

Aside from investigating performance solely in dE terms, the difference between matches in terms of a* and b* chromaticity co-ordinates illustrated where performance varied. Greater significance was seen to exist with b* difference across the illuminations ($F(4,32) = 47.556$, $p<0.001$). This, in conjunction with the non-significance of a* difference, pointed towards a relationship between b* difference value and illumination

CCT. As shown by the similar b* difference values, across both conditions, for both D67 and the two orthogonal illuminations. This variation in b*, or the blue – yellow axis, illustrated that matches were made to more neutral points in the blue yellow axis than the illuminations/ whitepoints were located. This neutral point was further from the whitepoint in the case with the more chromatic daylight illuminations, suggesting the existence of a constrained illumination estimation ability within colour constancy along the BY axis. The greater, more methodical, movement in difference along this axis is linked to the colours of these illuminations.

Whereas the significance of illumination on colour constancy was seen, the effect and significance of box condition was not. However the significance of an illumination-condition interaction ($F(4,32) = 21.670$, $p<0.000$) was uncovered. These two box conditions were shown not to effect match accuracy on their own. However, the significant interaction between condition and illumination is illustrated by the dE under illumination 2 (D40) in the grey box condition. The error here is vastly different o the errors under each of the remaining conditions.

5.7 Summary

This chapter has shown the ability of reflective display technology, such as the Mirasol display, to be utilised within human vision experiments. This has been shown by the minimal chromatic differences between the illumination and screen whitepoints and irradiances under each of the test illuminations alongside the manner in which, in terms of luminance, the screen is absorbed into the scene. This ‘absorbance’ into the scene refers to the minimal luminance difference between the screen and other surfaces in the scene with respect to the illuminations along with the verification of this through the results. The stability of gamut shape, alongside near linear scaling with respect to luminance and shift in gamut location also prove that reflective displays such as the one tested are highly useable within vision research.

Further to this the chapter has also shown that colour constancy within the human visual system exhibits different levels of performance with regards to illumination, across both known and unknown illuminations. Overall dE performance showed higher achromatic match ability under the cooler daylight illuminations across both conditions, with the best performance being under D250, closely followed by D67. This matches results found previously within [22, 25].

Whilst no overall performance difference in terms of dE between the two conditions existed investigations into the directions of average matches along each CIELAB component axis (a*, b*) showed differences between the two condition alongside illustrating the performance differences between the different illuminations and along the daylight locus. The difference in a* match location points towards different clues being utilised by the visual system when experiencing differing visual scenes. However the similar nature of b* differences points towards the different manner in which different illuminations are processed. The shift of b* differences shows a tendency for the visual system to generally show an estimation bias within the blue region of the daylight locus located between D67 and D250. These three sets of performance and illumination estimation bias findings together point towards the human visual system being 1) pre disposed to illuminations being within a certain region of the daylight locus, and 2) utilising different cues in different scenes which enable the same level of performance.

With reference to the original challenges and aims laid out within Chapter 2 this work has contributed a number of findings. It has shown that human colour constancy has a bias towards cooler daylight illuminations. This bias deserves both further psychophysical and computer vision based investigation. Results also showed that test scene did not have any effect upon the colour constancy exhibited by the subjects. As such, this points towards computer vision based methods being required to be scene invariant to a certain degree, however the results do not tell us how to achieve this. These results cannot realistically be used as the basis for a computational algorithm due to a lack of knowledge with regards to how the mechanisms at play here operate. However the next chapter will show how findings from psychophysical tests, like with the work shown in Chapter 3, can be utilised in this way.

Chapter 6. Contrasting hyperspectral and consumer panel analysis of tomatoes under different storage techniques.

6.1 Introduction

The previous chapters within this thesis have shown the usefulness of using hyperspectral imaging system to create algorithms based upon bridging colour image processing and human perception of colour and object features. Research in the previous chapters has been broad in its potential applications, but we show, within this chapter, the importance of bridging the gap between human colour perception and colour capture with regards to tomato quality grading.

Hyperspectral imagery (HSI) has over the last few years been utilised in the analysis of agricultural [130] and livestock produce [131], across both the visual, near infra-red (NIR) and longer wavelength IR spectrum regions. Across the visual wavelength region the use of HIS has been increasing as of late, however RGB based imaging systems are still overwhelmingly utilised due to advantages, such as capture speed. However, RGB based systems can only be used to grade a small number of features, such as colour, size and texture and direct comparison to human colour perception is left open to inaccuracies under uncontrolled illumination.

6.2 Motivation

The importance of perceived food quality has long been an area of interest for both the food industry and psychologists. A great number of sensory studies have been undertaken to investigate the effects of different storage, cooking and other treatment methods upon the perception of food produce on end consumers. With specific regards to previous tomato studies these have involved the sensory testing of the tomatoes from different conditions. During this testing the tomatoes are presented not as a whole, but cut into individual sections. This style of presentation differs to the way in which produce, such as tomatoes are presented in a commercial sense. If a customer enters a store foodstuffs such as tomatoes are presented in their whole form. As such the previous method of enabling ripeness and colour judgements by test subjects on tomato slices may not translate into an accurate predictor of consumer behaviour when viewing produce in a shop setting. However this is not to remove the importance of other sensory perceptions such as taste upon the repeat buying nature of the consumer.

Whilst previous studies have encompassed basic colour judgements by consumers of food produce we have designed a more in depth colour and texture questionnaire to

investigate their overall visual perception of the tomato. Another novel aspect of this work is that in addition to this, and in order to ensure we can directly compare the results from both questionnaires, we present the same tomato for each subject for both the visual and sensory questionnaires.

Further to this we utilise hyperspectral imaging to measure the colour information of each individual tomato. Whilst this may not be novel in itself, the use of this information to form an accurate average colour descriptor of the tomato for comparison to the results of consumer testing is. Previous studies utilise generally 4 spot colourimetric measurements along the equatorial line of the tomato. However this method enables small area features, such as bruises or scars, to influence the overall colourimetric descriptor of the sample. The use of a larger tomato area minimises the influence of such small features. Whilst previous studies have utilised the hyperspectral imaging of an entire sample [132-134] our method, utilises the capture of half of the tomato sample which still enables our method to benefit from the previously mentioned advantages.

The use of hyperspectral imaging allows us to capture the spectral reflectance information of the tomato, enabling investigations into specific wavelengths as well as more accurate colourimetric conversions into different colour spaces, such as CIELAB. [134] discusses the different types of hyperspectral imaging; reflectance, interactance and transmittance. Here we utilise the interactance measurement technique, although for exterior measurements the reflectance technique would also have been valid.

6.3 Methodology

Images of each tomato were also taken using a spectro-radiometrically calibrated hyperspectral camera and mirror housing, using a Schneider 23mm 1.4 compact C-Mount lens in combination with a custom built LED illumination system [29]. This lighting system produced a diffuse, horizontally uniform illumination in the image capture box below it. The hyperspectral camera enabled the capture of the VNIR range of wavelengths (400–1000nm), at a 2.8nm resolution, for each individual pixel in the resulting image. An example of how data is captured for each pixel, on a wavelength-by-wavelength basis, with a hyperspectral imaging system is shown in [135]. For this experiment only the visual 400–780nm band of the VNIR spectrum was captured, due to our interest in colour and as the LED illuminator channels were only within this region. The previously mentioned spectro-radiometric calibration allowed the retrieval

of the true colourimetric values from captured images that are directly comparable to those from a calibrated PR-650 spectrophotometer. This involved the interpolation of wavelength resolution from the 2.8nm steps output from the Specim V10E to 4nm wavelength steps.

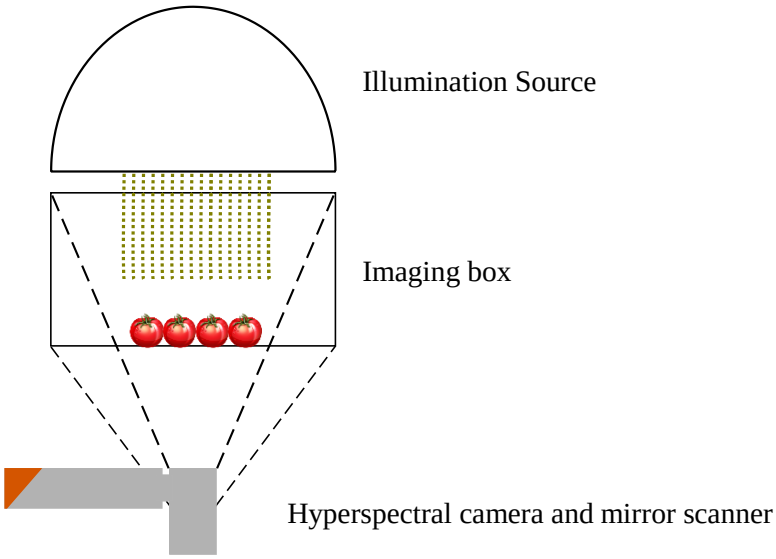


Figure 79 - Illustration of the hyperspectral imaging set up.

Tomato samples were placed under the LED light source in a grey Verivide lighting cabinet viewing box and measured under a metamer of CIE illuminant D65. From these captured images individual tomato regions were selected in post processing through the creation of image masks, with irradiance data then being retrieved for each pixel within each tomato mask. This irradiance spectrum was then normalised using the measured illumination spectrum to retrieve the reflectance spectrum for each pixel. The irradiance spectrum for each pixel was also converted into XYZ tri-stimulus values prior to colour space conversion into CIELAB. The number of sampled pixels for each tomato, when imaged from a set distance, varied depending upon the size of the tomato due to the use of individual tomato masks.

Each of these tomatoes was then shown to individual subjects who carried out two sets of tests. The first test incorporated the subjects completing a questionnaire based around their visual analysis of the exterior of the tomato. This enables the subjects to grade the following visual features; Percentage of tomato surface covered in red, green and orange colour; strength of the existing red, green and orange colour; lightness, glossiness, ripeness and freshness.

The subject was then asked to carry out a second set of tests on the same tomato, this time investigating the sensory perceptions of the tomato. Whilst for the previous test the subject is shown the exterior of the whole tomato, for this set of tests the tomato was presented in smaller pieces with the interior on show. Each subject rates three tomatoes at the same time, one from each of the three groups. As a side note, this sensory questionnaire was designed as part of another study, and is not the main focus of our investigations here. This means that the discussion of the results regarding from this questionnaire will be minimal. An overview of features that are rated is given below; colour, firmness, ripeness, aroma, crunchiness, sweetness, acidity, moistness, flavour and overall rating.

Subjects repeated these two tests across a group of three tomatoes, which had been stored under different conditions. Three initial tomato conditions were utilised, with tomatoes being stored under room temperature (RT), supply chain (SC) and intermediate (IT) conditions. Tomatoes were stored under these conditions for an initial 7 days. After this period each tomato condition group was split into two further groups for the remaining 8 days; under RT or refrigerated (F) conditions. These tests were repeated twice across two months; July and October. The testing across two months allowed any effect of harvest month to be extracted if it existed.

It must be made clear from this point that the intricacies of the different storage techniques will not be evaluated here as they form the basis of a connected study, but we will compare their effect upon the colour and the perceived sensory and colour nature of the tomatoes for human observers.

Initial colour data was captured at day 0 for all tomatoes, with visual-sensory tests, along with their accompanying colour measurements, being undertaken at days 7, 11 and 15.

The measurement and testing regimen is illustrated in below;

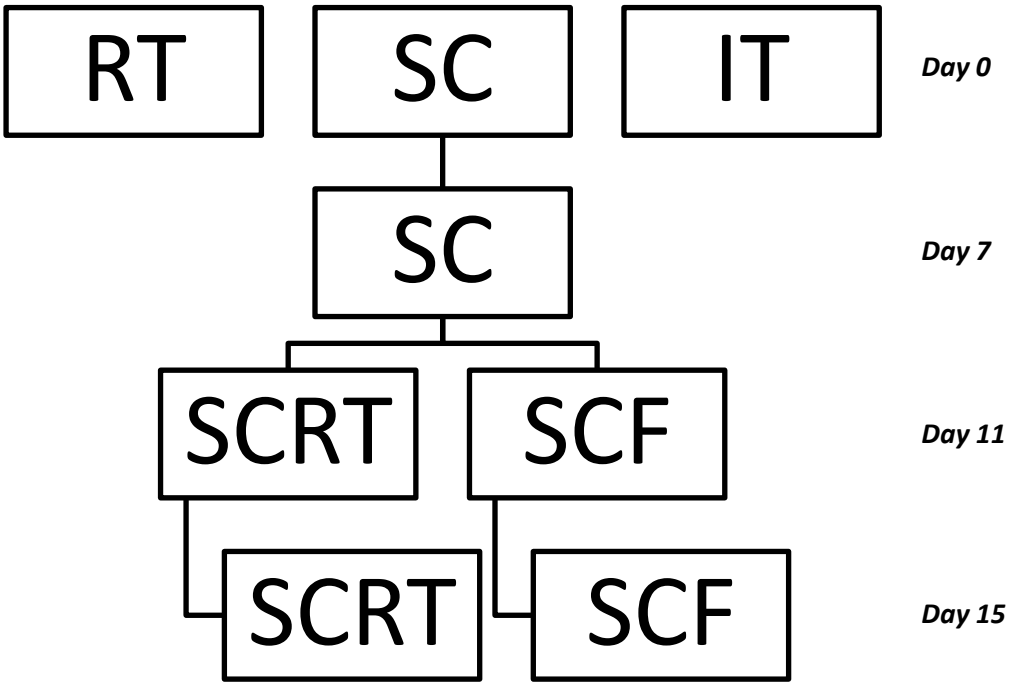


Figure 80 - An illustration of the colour measurement and observer testing regimen for a given test month, with Days 11 and 15 showing the resulting two sets each main tomato group was split into.

6.4 Results

From the overall study this chapter will focus upon the discussion of the results from the visual questionnaire and then move on to compare these results to those from the subsequent sensory questionnaire. This comparison between the two results will enable investigations into the effect of other sensory inputs upon judgments based solely upon the visual appearance of the tomato.

6.4.1 *Harvest month dependence*

The spread of testing across two different months enabled an initial investigation into any differences between tomatoes at these two stages of the U.K. growing season. Harvest month was not expected to have an effect upon the initial colour characteristics of the tomatoes as they had been harvested at the same period of maturity, at ripening stage 3. To investigate this, the average tomato chromaticities (CIELAB) at Day 0 for each condition across the two months for the initial tomatoes were compared. No significances related to month were found; chromaticity-month interaction $F(1,6)=2.222$, $p=0.187$, chromaticity-condition interaction $F(2,6)=0.877$, $p=0.463$, and for a chromaticity-month-condition interaction $F(2,6)=3.050$ with $p=0.122$.

We then investigated if any difference between the interior and exterior measured chromaticity of the tomatoes on day 0 existed. This was undertaken in order to investigate if the interior chromaticities differed across months and/ or days to prepare for the subsequent effects upon the exterior and interior subject judgements. We found no statistical differences between tomatoes from the same sets and their interior and exterior chromaticities. Differences in the significance of month were found only with the interior chromaticities on day 0. A chromaticity-month interaction of $F(2,12)=6.223$, $p=0.014$ was found. This difference in relationships for the exterior and interior tomato chromaticity shows that different parts of the tomato ripen at different rates during different parts of the growing season. This could be due to differences in growing conditions, such as rainfall, sun exposure and temperature.

6.4.2 *Whole tomato chromatic differences across conditions and test days*

Across the overall testing period of two months and the two sets 155 valid questionnaire responses were collected, a breakdown of the sets these responses were for are shown in Table 13 below;

	Day	**Set 1**	**Set 2**
July	7	18	15
	11	21	8
	15	17	15
October	7	12	14
	11	14	7
	15	14	
	Total	96	59

Table 13 - A breakdown of the valid questionnaire responses gained during the two test months across the two condition sets (Day 15 for October Set 2 was not run).

We first investigate the effect of tomato conditions on tomato chromaticity across the two months;

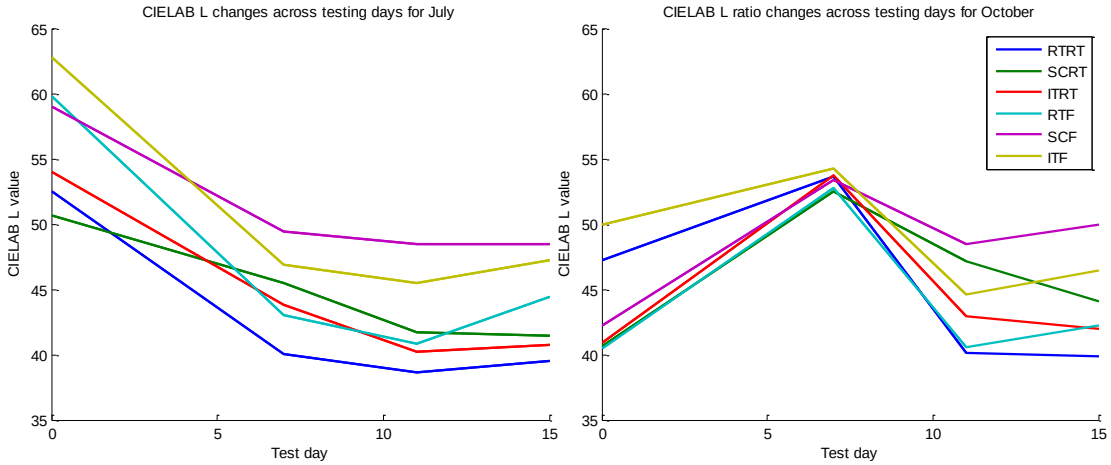


Figure 81 - Contrasting the change in mean L values across testing days for different storage conditions for a) July, and b) October.

The relative luminance of the tomato samples across the test days was investigated and shown in Figure 81. The two different tomato storage conditions, ~RT and ~F, exhibited different results. The greatest significances were found for tomatoes under storage conditions from day 7 onwards different to their original conditions for the first 7 days. This is illustrated that for tomatoes stored under RT condition onwards from day 7 significances were found, in relation to test day, for both the SC ($F(2,90)=24.449$ $p=0.000$) and IT ($F(2,90)=9.636$ $p=0.000$) groups. For the ~F group an effect for the RT tomatoes was seen related both with respect to day, ($F(2,54)=4.134$ $p=0.021$, as with the other two groups in the ~RT group. Along with a month-day interaction; $F(1,54)=5.765$ $p=0.020$.

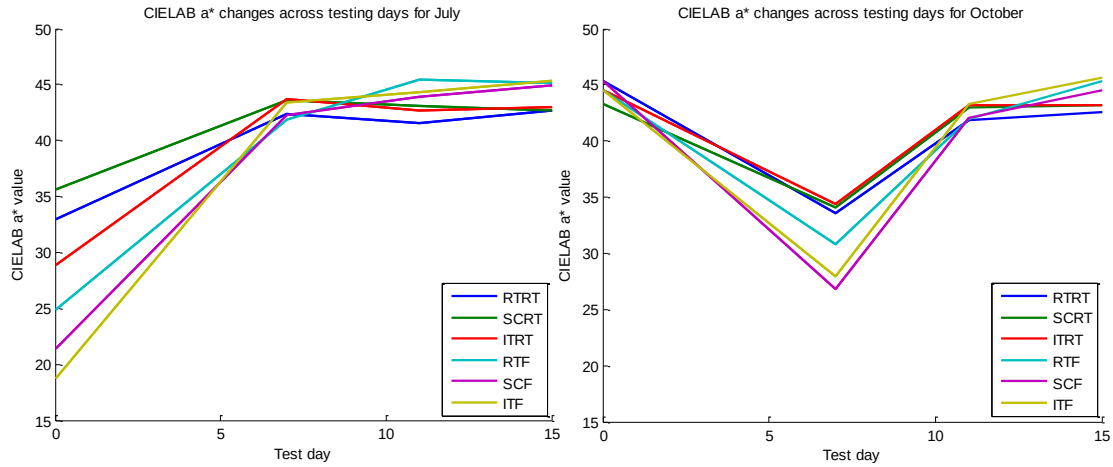


Figure 82 - Contrasting the change in mean a* values across testing days for different storage conditions for a) July, and b) October.

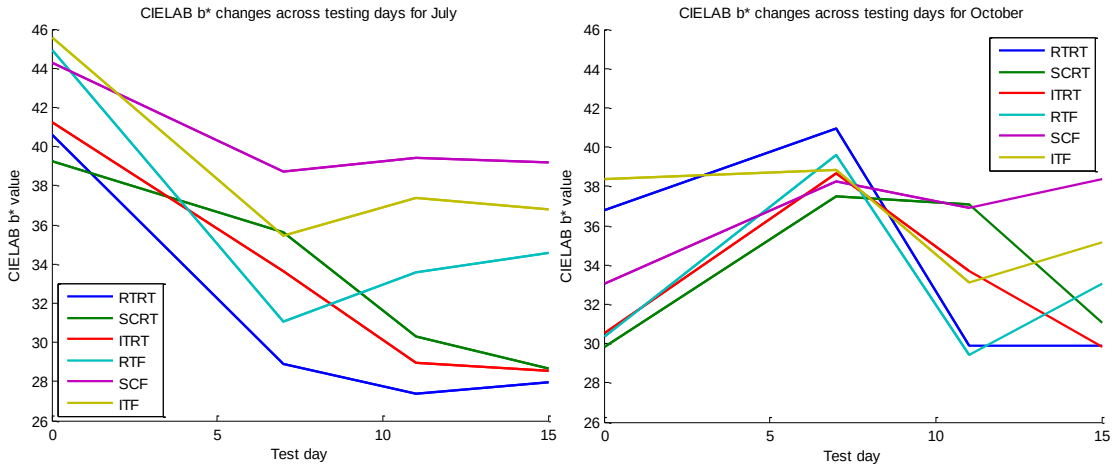


Figure 83 - Contrasting the change in mean b* values across testing days for different storage conditions for a) July, and b) October.

Comparing Figure 82 and Figure 83 shows that across the testing days a larger effect of storage condition can be noted for changes in b* value. This is reinforced through statistical analysis where a significant condition-day interaction for b* value was found $F(12,432)=8.951$, $p=0.000$. This significance was also found for both the relative luminance $F(12,432)=3.175$, $p=0.000$, and a*/b* ratio $F(12,432)=10.969$, $p=0.000$.

This exclusive relationship between the groups, their storage conditions and luminance changes shows that temperature over short term periods has a large effect upon the luminance of the tomato.

Investigations into changes for the average CIELAB a* value across conditions showed differences between the two storage condition sets of ~RT and ~F. The ~RT group exhibited a significant relationship for the SCRT tomato group with a month-day interaction; $F(2,90)=4.069$ $p=0.020$.

However the ~F group exhibited a statistically significant relationship for all tomato storage types in relation to test day (storage length). RTF; $F(2,54)=20.710$ $p=0.000$, SCF; $F(2,54)=11.607$ $p=0.000$, ITF; $F(2,54)=6.394$ $p=0.003$. This significance across all tomato groups illustrates a major effect of storing tomatoes under the ~F condition in comparison to the ~RT condition.

Investigations into the effect of the storage conditions upon the b* value of tomatoes also brought interesting relationships to the fore. For both the ~RT and ~F storage conditions significant relationships were found both in relation to month and test day. For the ~RT group RTRT; $F(1,90)=16.041$ $p=0.000$, SCRT; $F(1,90)=11.802$ $p=0.001$, ITRT; $F(1,90)=5.195$ $p=0.025$. In relation to day significant relationships were found for both SCRT ($F(2,90)=107.426$ $p=0.000$) and ITRT ($F(2,90)=46.480$ $p=0.000$).

The ~F storage group exhibited the same pattern of results. In relation to month both SCF ($F(1,54)=5.003$ $p=0.029$) and ITF ($F(1,54)=11.094$ $p=0.002$) exhibited significant relationships. On top of this day also was seen to be significant in its relationship with b* value for both the RTF ($F(2,54)=23.563$ $p=0.000$) and ITF ($F(2,54)=4.592$ $p=0.014$) groups.

The significant difference in b* value across months for both groups may illustrate a seasonal variation in the maturity of the tomato groups, even with the tomatoes being harvested at the same ripening stage by a human picker. With all tomato types effected to a significant degree by month for the ~RT group and two out of the three showing the same relationship for the ~F group. Previously it was mentioned that no month based differences were found between the tomatoes on day 0, which is expected if the tomatoes are harvested at the same ripeness level. However the differences in b* related to month above were across days 7, 11 and 15. As such we can hypothesise that the two months here have an effect upon the ripening process of the tomatoes.

As with the L and a* changes across test days b* values are also effected. With the greatest changes, as with luminance values, being with tomato groups being stored under a different secondary condition. For the ~RT storage condition SC ($F(2,90)=107.426$ $p=0.000$) and IT ($F(2,90)=46.480$ $p=0.000$) tomatoes were significantly affected, and for the ~F condition RT ($F(2,54)=23.563$ $p=0.000$) and IT ($F(2,54)=4.592$ $p=0.014$) groups were affected. This mutually exclusive relationship may point towards a stalled, or slowed ripening process for tomatoes stored under

conditions different to their original storage conditions. As such it could point towards methods of varying the speed at which ripening takes place.

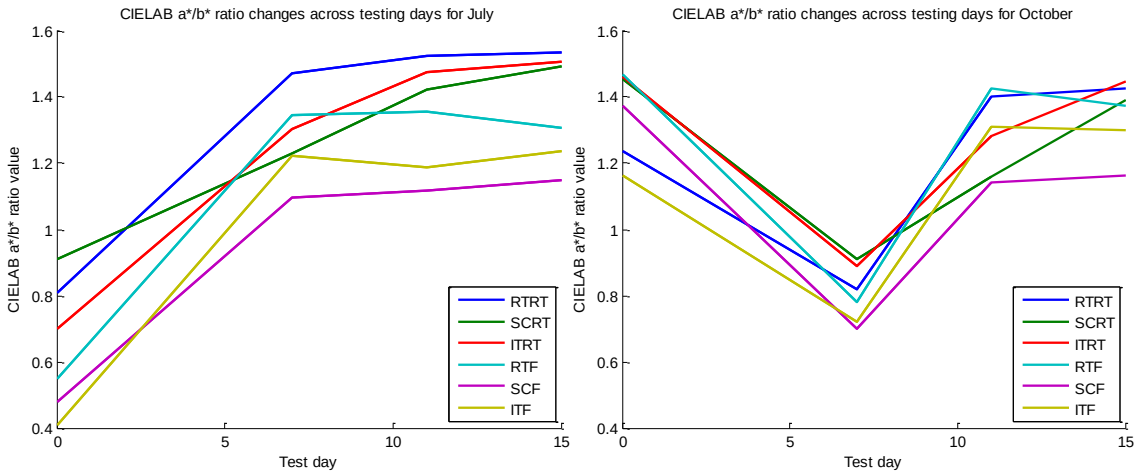


Figure 84 - Contrasting the change in mean a*/b* ratio values across testing days for different storage conditions for a) July, and b) October.

The ratio of CIELAB a*/b* has been proposed as a method of estimating the lycopene content, thus ripeness, of a tomato [136]. For all of the storage types Figure 84 shows that as storage length increases there is a general increase in the value of this ratio, which is to be expected. However for each tomato group and treatment it was found that month had a significant effect upon this value, this was to be expected after finding the month dependence of the a*b* tomato chromaticities at the beginning of the two testing months. Test month was seen to effect the ~RT group in the following manner; RTRT (RTRT $F(1,90)=30.104$ $p<0.001$), SCRT (SCRT $F(1,90)=11.413$ $p=0.001$), and ITRT (ITRT $F(1,90)=6.821$ $p=0.011$). The ~F treatment group was affected RTF ($F(1,54)=5.850$ $p=0.019$), SCF ($F(1,54)=6.044$ $p=0.017$), and ITF ($F(1,54)=25.968$ $p<0.001$).

6.4.3 *Chemometric analysis across tomato conditions*

Utilisation of a hyperspectral imaging technique also means that we have captured the spectral data for each tomato. This allows the use of chemometric methods to investigate the spectral reflectance changes across tomato groups, which a number of papers [132-134, 137, 138] have done. These papers utilise [132, 137, 138] methods such as partial least squares regression (PLSR) to analyse and uncover relationships between spectral and other empirical measurements. Rajkumar et al. [137] for instance used the PLSR method to uncover important NIR wavelengths from spectroscopy to aid moisture content in banana slices. As stated previously we utilised the interactance hyperspectral imaging method in comparison to the reflectance method used in the

aforementioned papers. Nevertheless we then utilise the PLSR method to investigate the variations in tomato reflectance spectra across the storage condition groups and across the storage length (test days).

Investigations into spectral variations for each tomato condition, across test days, were undertaken in order to illustrate the effect of each storage condition on different regions of the spectral reflectance function. Through testing and viewing the R^2 values, PLSR analysis was carried out;

Tomato Condition	**Sample Size**	**PLSR Components**	**R^2**	**Wavelengths of interest (nm)**
RTRT	96	6	0.9733	492, 496, 528, 576, 628, 692
SCRT	93	5	0.9752	552, 620, 640, 644, 680
ITRT	94	6	0.9557	456, 484, 556, 560, 592, 692
RTF	58	5	0.9993	480, 532, 556, 576, 648
SCF	57	4	0.9972	456, 484, 564, 652
ITF	58	4	0.9981	580, 596, 624, 652

Table 14 - Illustration of the R^2 error and important wavelengths resulting from PLSR for each tomato condition across the test months.

This shows that the storage length, for each of the storage conditions can be separated into their respective lengths. This means that we can tune and model the spectral reflectance changes that are exhibited by tomatoes under each storage condition. The fact that storage length can be correlated to specific storage lengths and thus ripening stage reinforces previous work [139], which put forward the idea of the use of hyperspectral imaging for ripeness detection. Whilst we do not specifically relate wavelengths to ripeness, we correlate them with the length of storage under a specific storage condition, which can also be of great use to the food processing industry.

All of this shows the effect of the different storage techniques upon the spectral and chromatic features of the tomatoes, however we must compare these differences to the consumer perception of the tomatoes in an attempt to model how consumer perception varies in relation to this measured data.

6.5 Questionnaire data analysis

6.5.1 *Visual data comparison*

In order to view the effect of the different storage conditions upon the perception of the tomatoes we will go through the visual judgements made by subjects. As we previously uncovered no difference between tomato chromaticity and changes across the two test months we will collapse the subject responses across them.

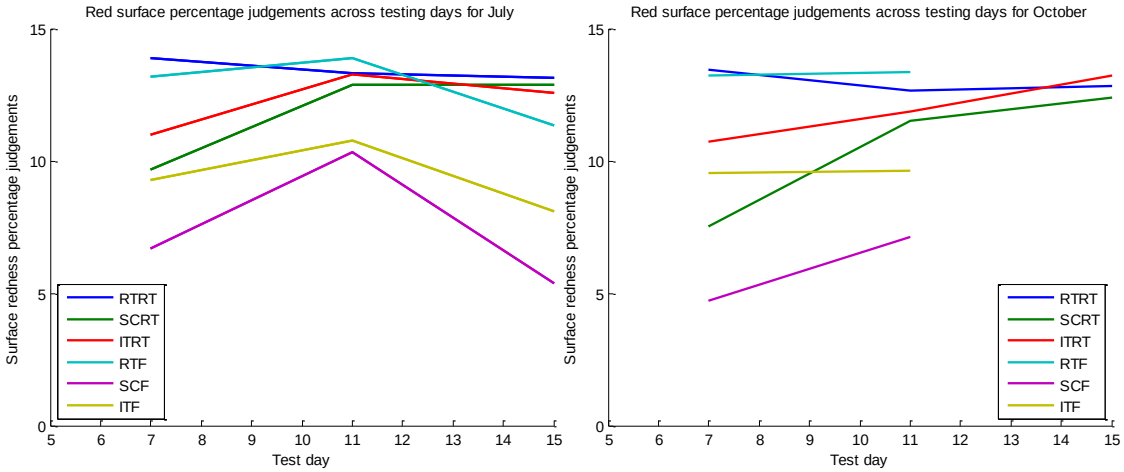


Figure 85 - Contrasting the change in mean red surface percentage values across testing days for different storage conditions for a) July, and b) October.

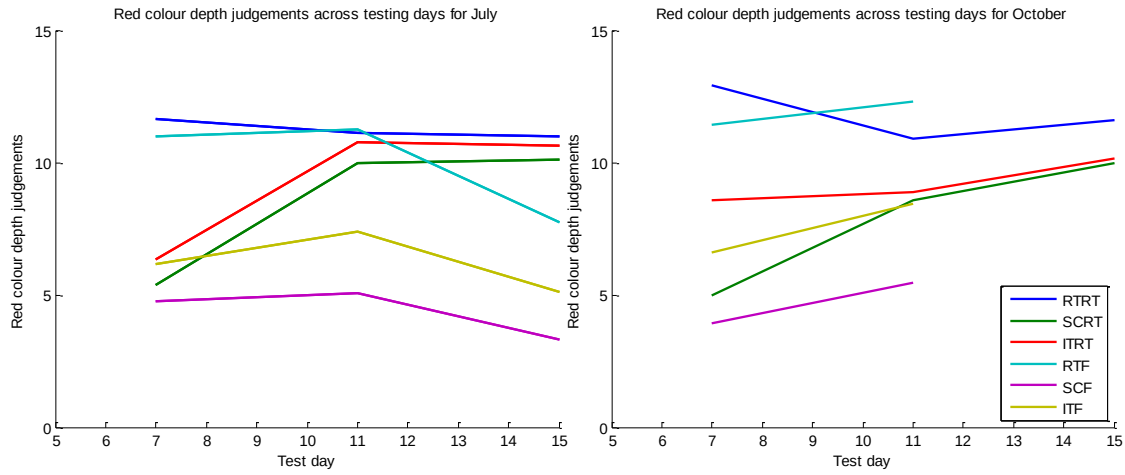


Figure 86 - Contrasting the change in mean red colour depth values across testing days for different storage conditions for a) July, and b) October.

Figure 85 – Figure 92 illustrate the differences between subject responses across the different sets an months, this inter subject variation is to be expected. As previously, a tomato colour difference across months for the exterior of the tomatoes was found to not be significant. Figure 85 illustrates the variation in the percentage of the tomato surface perceived to be red across the test months and days. Month was discovered to

have no statistical significance in relation to the perceived surface covered in red. However length of storage under each condition was seen to have an effect. With the tomatoes kept under ~RT conditions the greatest effect of this storage condition across days was seen to be with the SC ($F(2,90)=13.570$, $p=0.000$) and IT ($F(2,90)=4.945$, $p=0.009$) tomatoes, with an increase in red percentage from day 7 onwards. For those stored under ~F conditions the greatest effect was seen upon the RT ($F(2,54)=3.360$, $p=0.042$) and SC (SCF $F(2,54)=3.889$, $p=0.026$) tomatoes. The pattern for the red percentage here shows a peak in red percentage value at day 11, which differs greatly from the ~RT group, both in value and pattern.

The depth of the red colour was also judged and can be seen in Figure 86. Analysis showed that for the ~RT group significance was seen with respect to both day and a month – day interaction for certain basic tomato groups. With relation to storage length this was seen with both SC ($F(2,90)=26.496$ $p=0.000$) and IT ($F(2,90)=9.469$ $p=0.000$) groups. A significant month – day interaction was also seen with the IT group $F(2,90)=4.470$, $p=0.014$. However with the ~F group significance was only noted in relation to storage length for the RT group $F(2,54)=8.536$, $p=0.001$. It can immediately be seen that the patterns exhibited in both the red percentage and redness responses are the same.

This illustrates a difference in the depth of the red colour in relation to storage length. If we compare this to ripeness and a*b* values this change matches the increase in a*/b* ratio with time.

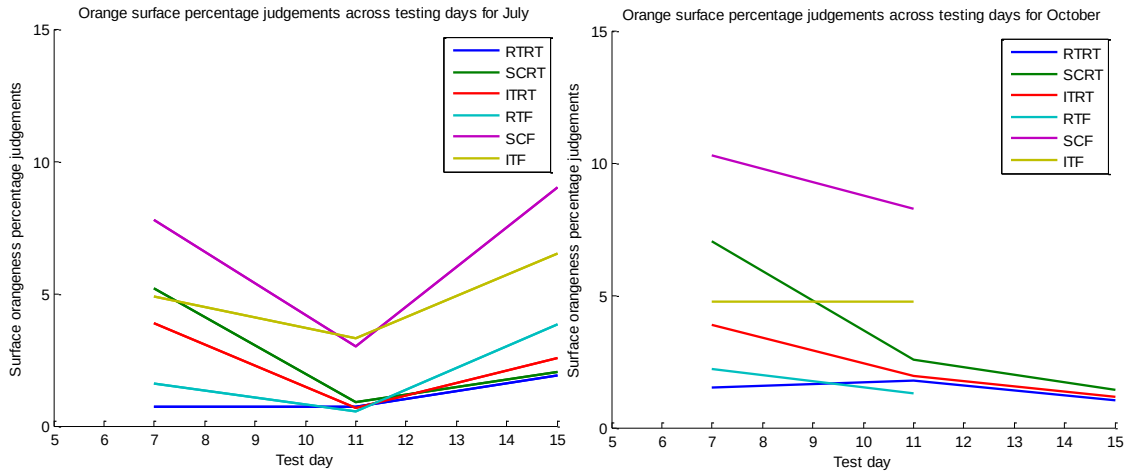


Figure 87 - Contrasting the change in mean orange surface percentage values across testing days for different storage conditions for a) July, and b) October.

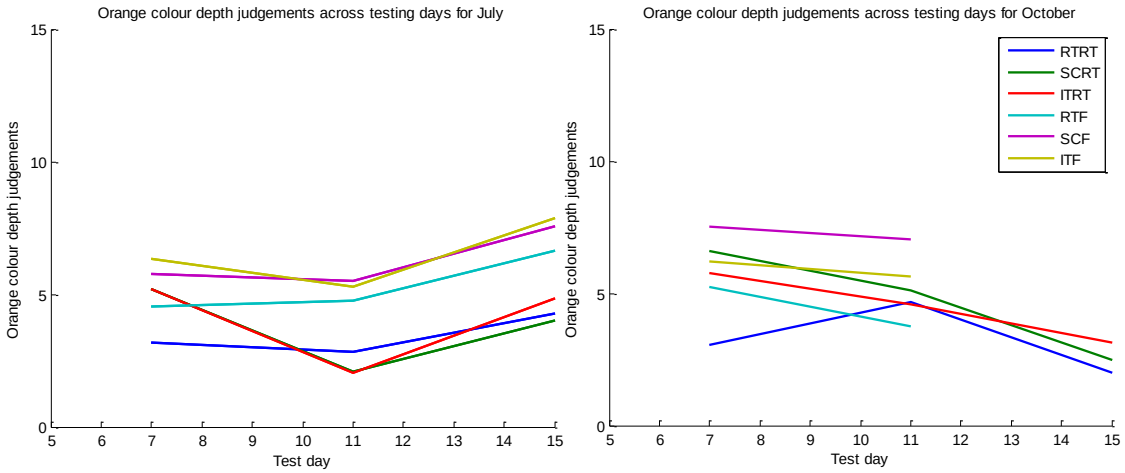


Figure 88 - Contrasting the change in mean orange colour depth values across testing days for different storage conditions for a) July, and b) October.

Figure 87 illustrates the variation in the percentage of the tomato considered to be orange. Again significant variation was seen with respect only to test day, and thus storage length. The tomato groups significantly affected for both the ~RT and ~F groups were the same as previously seen with the surface red coverage. For the ~RT group SC ($F(2,90)=16.855$, $p=0.000$) and IT ($F(2,90)=6.195$ $p=0.003$) tomatoes were affected. For the ~F group the RT ($F(2,54)=5.208$ $p=0.009$) and SC ($F(2,54)=5.058$ $p=0.010$) tomatoes were affected.

On top of this it can be noted that the ~F group is considered to have a greater orange surface area than the ~RT group tomatoes. The dip in orange percentage at day 11 for the ~F corresponds directly to the previously mentioned peak at the same time in red surface percentage. The depth of the existing orange colour on the tomato was also seen to have less significant relationships across storage length. Only the SC tomato group under the ~RT condition ($F(2,90)=4.060$ $p=0.021$) showed a relationship to the depth of the perceived orange surface colour.

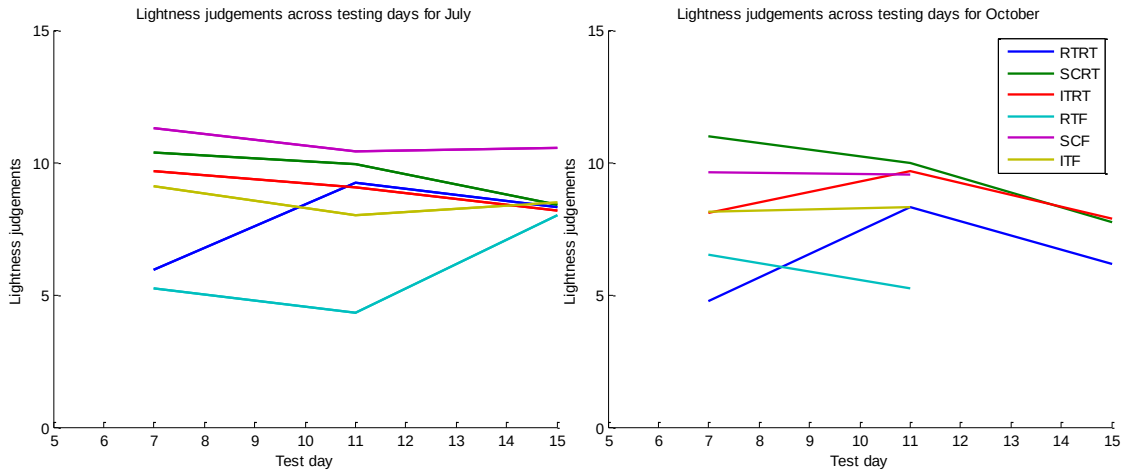


Figure 89 - Contrasting the change in mean lightness values across testing days for different storage conditions for a) July, and b) October.

Perceived surface lightness was also seen to vary in relation to storage length for both treatment types. With the ~RT treatment having an effect upon both the RT ($F(2,90)=7.191$, $p=0.001$) and SC (SCRT $F(2,90)=8.702$, $p=0.000$) tomatoes. Within the ~F treatment groups only the RT tomatoes were effected to a significant degree; $F(2,54)=4.148$ $p=0.021$.

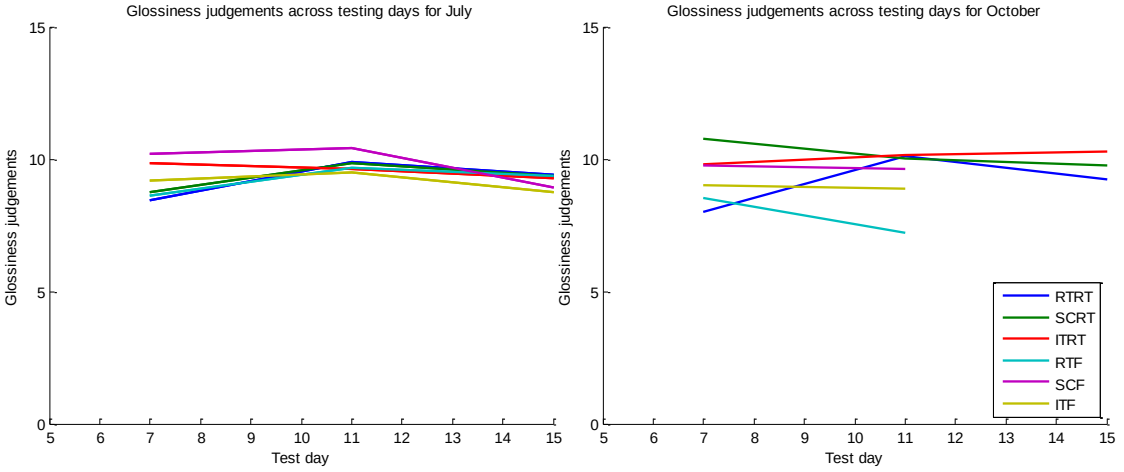


Figure 90 - Contrasting the change in mean glossiness values across testing days for different storage conditions for a) July, and b) October.

Perceived glossiness, as shown in Figure 90, was seen to have no significant relationship with respect to month or day.

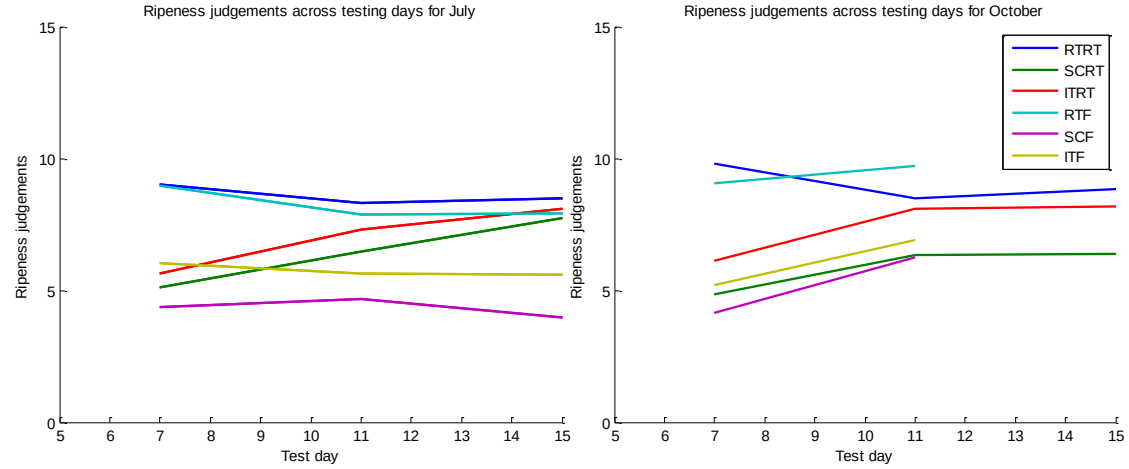


Figure 91 - Contrasting the change in mean ripeness values across testing days for different storage conditions for a) July, and b) October.

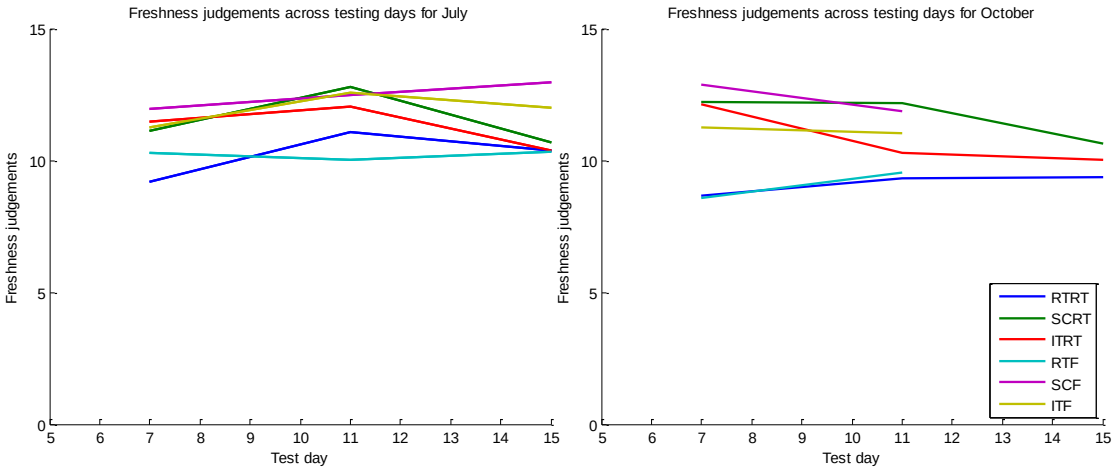


Figure 92 - Contrasting the change in mean freshness values across testing days for different storage conditions for a) July, and b) October.

For the consumer perception of raw vegetable or crop foodstuffs the most important indicators of the effect of different storage treatments is upon the perceived ripeness and freshness of the produce. By considering perceived ripeness it was seen that the ~RT treatment had the greatest effect upon the tomatoes. The ~F condition showed no significance across any of the original tomato groups. Within the ~RT treatment significant effect upon ripeness with respect to storage length for both the SC ($F(2,90)=7.729$, $p=0.001$) and IT ($F(2,90)=6.749$, $p=0.002$) tomatoes. Figure 91 shows that the RT tomatoes under both treatments are affected to a smaller degree. What this means is that perceptually storage of tomatoes under the ~F condition minimises changes in ripeness across the test period for all tomato groups, both amongst themselves and when compared to the other groups.

On top of this consideration of the perceived freshness can also aid comparisons between the different tomato groups and treatments. For freshness a significant effect was only noted for the SCRT tomatoes; SCRT $F(2,90)=5.283$ $p=0.007$.

As we have considered the perceived ripeness and freshness of the tomato groups, it makes sense to investigate if a correlation between them exists. A correlation between the two does exist with similar values across all tomato groups RT ($r(155)=-0.230$, $p=0.004$), SC ($r(155)=-0.217$, $p=0.007$) and IT ($r(155)=-0.356$, $p=0.000$).

We previously utilised a chemometric method of investigation to find wavelengths of importance in relation to storage length across each tomato condition. However methods to uncover regions of the spectral reflectance functions that correlate to the visual scores given by human observers have been found to not be achievable with an appropriate number of components. This can be attributed to the highly variable nature of the visual scores with regards to our untrained subjects.

6.5.2 *Contrasting the visual and sensory questionnaire responses*

Whilst the findings from the visual and sensory questionnaires on their own are useful to compare the consumer perception of the different tomato storage groups the manner in which subjects compared the same tomatoes across the two questionnaires is also useful.

With the design of the experiment incorporating this it enables the response of a subject using firstly exterior visual judgments to be compared to some similar judgments after viewing the interior and taste tests. This information in conjunction with the exterior colour data from each tomato, and group average interior colour information, enables statistical investigations into the importance of different tomato features.

The first comparison between the results of the two questionnaires regarding the correlation between visual and sensory ripeness showed correlation for each tomato group. Comparing the two ripeness ratings is of interest here to visualise the importance of both overall visual and sensory ratings upon the ripeness rating. For the ~RT group each tomato condition exhibited a significant correlation, RTRT ($r(66) =0.429$, $p=0.000$), SCRT ($r(66) =0.324$, $p=0.008$) and ITRT ($r(66)=0.336$, $p=0.006$). The ~F group also showed significance; RTF ($r(48) =0.372$, $p=0.009$), SCF ($r(48) =0.5$, $p=0.000$) and ITF ($r(48)=0.342$, $p=0.017$).

However the most important comparison to view in an attempt to uncover any validity in using the visual ripeness rating to estimate the overall rating for the tomato by a consumer after sensory testing. This uncovered a relationship for only three out of the six tomato groups; SCRT ($r(66)=0.374$, $p=0.002$), SCF ($r(48)=0.5$, $p=0.000$) and ITF ($r(48)=0.404$, $p=0.004$). The insignificant correlations exhibited the following RTRT ($r(66)=0.000$, $p=1.000$), ITRT ($r(66)=0.05$, $p=0.686$), RTF ($r(48)=0.372$, $p=0.009$) .

If we view the correlations between sensory ripeness and overall ratings we see a slightly greater level of correlation, both in terms of tomato groups and their correlation values. Here we find significance for both the SC and IT tomato groups under the ~RT and ~F conditions. SCRT ($r(66)=0.535$, $p =0.000$), ITRT ($r(66)=0.442$, $p=0.000$), SCF ($r(48)=0.550$, $p=0.000$), and ITF ($r(48)=0.593$, $p=0.000$). The insignificant RT group statistics were RTRT($r(66)=-0.017$, $p=0.891$) and RTF ($r(48)=0.259$, $p=0.076$) . The consistency and size of these relationships across the same two tomato groups across the two main storage conditions means that for these two groups the sensory ratings outweigh the visual results for consumers.

Overall we can see a strong correlation in ripeness rating between the visual and sensory questionnaires across all tomato groups, reinforcing the previously noted correlation between interior and exterior tomato colour and the importance of colour to ripeness perception in comparison to other metrics.

6.5.3 *Chemometric analysis of reflectance spectra to measured empirical non visual tomato data*

Further to the visual and sensory testing carried out by untrained human observers hyperspectral images were also taken of a number of other tomato samples from each test group across test days across 3 different test months, including the two previously mentioned. This means the wavelengths of interest below have been uncovered using tomatoes across the growing season. These tomatoes were the subject to testing for a number of different empirical measurements covering; pH, Total soluble solids (TSS, °Brix), Sugar content (g/100g), Titratable acidity (TA, ml), Citric acid content (g/100g), Vitamin C titre (ml), Ascorbic acid in sample (mg), Ascorbic acid in volumetric (mg/50ml) and Vitamin C (mg/100g) .

Tomato Feature	Components	R^2	Wavelengths of interest (nm)
pH	11	0.7171	456, 468, 484, 524, 544, 572, 600, 624, 632, 660, 672
Total Soluble Solids (TSS) (°Brix)	7	0.7756	456, 484, 496, 620, 640, 680, 684
Sugar g/100g	7	0.7756	456, 484, 496, 620, 640, 680, 684
Titratable Acidity (TA) (ml)	10	0.6509	456, 468, 476, 560, 592, 620, 632, 672, 680, 684
Citric acid g/100g	10	0.6509	456, 468, 476, 560, 592, 620, 632, 672, 680, 684
Vitamin C in titre (ml)	17	0.8595	456, 488, 504, 516, 520, 532, 592, 608, 624, 628, 632, 640, 656, 672, 680, 684, 692
Ascorbic acid in sample (mg)	16	0.8294	456, 488, 496, 504, 516, 520, 580, 592, 608, 620, 624, 628, 632, 640, 672, 680, 684, 692.
Ascorbic acid in volumetric (mg/50ml)	19	0.8294	456 , 488, 496, 504, 516, 520, 580, 592, 608, 620, 624, 628, 632, 640, 672, 680, 684, 692
Vitamin C mg/100g	13	0.8294	456, 460, 504, 520, 592, 608, 620, 624, 632, 640, 672, 680, 684

Table 15 - Overview of key wavelengths connected to different tomato features and the r^2 value using the PLSR method using a total of 83 tomatoes.

Table 15 above illustrates the wavelengths found to vary the most with respect to the different measured features of the tomatoes. When comparing to previous works summarised within [140], we can see that our results in terms number of components and r^2 values are well within accepted complexity and accuracy limits. This shows that these different features can be modelled and estimated using, non-destructive, hyperspectral imaging of the exterior of tomatoes. These findings are novel in themselves as although some of the aforementioned features have been investigated with other food products, such as meats and fish [132-134], fruits [137, 141] and vegetables alongside ripeness for tomatoes [139] this work has not been undertaken to relate these biochemical features with tomatoes. Our results show good findings and a very good level of accuracy, however the low number of total tomatoes utilised to create these models also needs to be taken into consideration. Previous investigations into

linking biochemical and wavelength reflectances across the visual and NIR wavelength regions vary greatly in their number of samples used to build models. With [132] utilising 22 salmon fillets to build their mode, and [137] utilising 360 banana samples. In our experiments here for this model a total of 83 tomatoes were utilised.

We show low levels of RMSE error with regards to pH, TSS, sugar content and ascorbic acid/ vitamin C content combined with a large reduction in the required wavelengths from 96 to 7 in some cases. Previous methods for investigation into other foodstuffs have generally focused upon the NIR region of the spectrum. The spectrum utilised here is purely the visible wavelength region, due to the output of our light source. Other regions may also prove to be useful for the prediction of the same features, however the cheaper nature of imaging and lighting equipment in the visible range of wavelengths is a major advantage for our findings and methodology.

6.6 Summary

We have discovered that the six different tomato storage conditions have very different chromatic and sensory effects upon the tomatoes.

Storage condition was noted to have the largest chromatic effect upon b* value ($F(12,432)=8.951$, $p<0.001$) in comparison to a*, with a smaller effect also present for relative luminance value ($F(12,432)=3.175$, $p<0.001$). This effect upon b* value then has a subsequent effect upon a*/b* ratio. This effect is noted across both ~RT and ~F tomato storage groups. The relative luminance value of the tomatoes was also seen to vary the most after the introduction of different storage techniques from day 7 onwards. With the most significant variation exhibited when temperature conditions we drastically changed from the condition used for the initial 7 days. This shows that the luminance of the tomatoes can be altered over a shorter time scale than chromaticity values with rapid changes in storage conditions.

However a difference between the a* chromatic effect of the post 7^{th} day storage type (~RT or ~F) was noted. It was seen that all tomatoes stored under the ~F condition exhibited a relationship between storage length and a* value; RTF ($F(2,54)=20.710$, $p<0.001$), SCF ($F(2,54)=11.607$, $p<0.001$), and ITF ($F(2,54)=6.394$, $p=0.003$). These differing effects upon chromaticity values by different storage techniques were then reinforced by the different visual ratings by human observers, with greater changes in visual ratings when greater absolute colour changes were exhibited. Post 7^{th} day storage

conditions were also seen to have an effect upon both the ripeness and freshness consumer ratings.

Comparisons between the visual and sensory ratings also showed some level of correlation. Whilst visual ratings cannot be relied upon, understandably so, as a sole guide for overall rating prediction, visual and sensory ripeness ratings were seen to be correlated.

The application of hyperspectral imaging allowed the use of chemometric analysis techniques. This analysis uncovered a number of findings. Firstly for each tomato storage technique storage length could be estimated using, at most, 6 wavelengths. This could feasibly allow a real time system to sort and classify different tomatoes under a single storage technique. The differences in these wavelengths for each group also illustrates the effect of each technique upon the spectral reflectance of a tomato.

Further chemometric analysis comparing surface reflectance to different biochemical measurements also uncovered interesting relationships. It was discovered that a number of different biochemical features within the tomato can be estimated, to a reasonable degree of accuracy, using wavelengths within the visible wavelength region. These features range from pH level, TSS, citric acid to vitamin C contents.

Previous work [137, 140, 141] has largely focused upon chemometric investigations within the NIR region of the electromagnetic spectrum, however our research has shown that cheaper devices within the visible range can also be used. However as future work investigations into comparing the use of these two different spectral regions, whether exclusively, or in combination, could be carried out to further lower the error of the regression model.

With regards to our original aims and objectives investigations into linking human and computer visual perception has succeeded in terms of correlations when considering chromatic changes. However we could find no way in which to correlate our chromatic measurements, or spectral measurements, with the visual findings by subjects. This means that we have successfully bridged human and computer vision to a certain degree, and have also shown how a future imaging system can be designed to take advantage of these relationships. Future work could involve investigation into the utilisation of the VNIR spectrum range to carry out the same tests. This would however take the work further away from our original aims and objectives.

Chapter 7. Conclusion & Future Work

This chapter will summarise the achievements of the work within this thesis and compare these achievements to the original aims and objectives laid out within the Chapter 2. Using this comparison, possibilities for future work extending on from the findings of this thesis are then outlined.

7.1 Conclusion

The work across each chapter within this thesis attempted to meet the aims and objectives laid out within the literature review within Chapter 2. This aims and objectives were developed from the findings of the literature survey and are directly linked to the task of developing novel approaches for, and investigations into colour constancy. These aims and objectives stemmed from the research approach that computational colour constancy solutions have not previously attempted to replicate the human visual system in a number of ways. This had led to approaches, that while do achieve good levels of performance in RGB imagery, have no connection to any comparable proven processes in the visual system. Some algorithms do this by assuming certain scene statistics to always be constant, however these statistical assumptions do not match up with scenes in which colour constancy operates. Whilst developed computational colour constancy approaches are generally connected to the physiological basis of chromatic adaptation this is generally the limit of any connection to any process in the human visual system. This has historically been due to imaging technology and the separate development of the computer vision industry. The work within this thesis has attempted to bridge computational approaches to colour constancy with a deeper perceptual and physiological basis.

The developed luminance skewness monotonic transform within Chapter 3 showed that even for images using the non-perceptually uniform RGB space, psychophysical findings can improve the colour normalisation performance. The utilisation of psychophysical findings related to the perception of gloss connected to luminance. Whilst RGB imagery cannot produce true luminance images, but rather luma images, these findings were shown to still be applicable. The addition of this perceptual finding improved the results of the original monotonic transform with images exhibiting highly skewed luma histograms. This algorithm also managed to preserve the most important aspects of the existing monotonic transform by both utilising and preserving perceptual scene contents. This also enabled the work to successfully meet another of the initial

objectives. However performance was seen to falter with certain images, notably those which do not exhibit skewness in combination with a large number of existing channel intensities.

Whilst Chapter 3 achieved the development of an algorithm with a psychophysical basis, it acknowledged the restricting nature of RGB imaging technology in relation to the physiological and psychophysical objectives of this thesis. However hyperspectral imagery allows the bypassing of this problem. Chapter 4 introduced a novel calibrated hyperspectral dataset of human skin under 39 illuminations. It utilised this to develop two colour constancy methods, assuming daylight illuminations, based upon the effect of these illuminations upon human skin gamuts. Both of these methods were developed using the physiological cone-contrast space, alongside inspiration from psychophysical findings on the effect of colour memory and polychromatic objects upon colour constancy. The approach was also inspired through the consideration of the content of most photographs or scenarios in which colour constancy algorithms are required. This chapter also showed a further bridging of the two colour constancy fields and utilised findings from both to develop novel methods. However the true ability of the methods cannot be evaluated until they are fully tested within more real world images.

Having developed two novel approaches for colour constancy within the previous two chapters Chapter 5 then showed a novel investigation into colour constancy within the human visual system. With Chapter 4 having developed and tested colour constancy approaches across daylight illuminations, and performance variations were noted, human performance was then investigated and compared across these. Whilst change in colour constancy performance had been seen previously across the daylight locus, results here showed this trend to a much greater degree and also in terms of how this is achieved. On top of this global scene content was seen not to exhibit any effect upon the human performance. These findings should help form the basis of future computational approaches. Furthermore the investigation should, for the first time, the applicability of reflective type computer displays in vision experiments.

The bridging of human colour perception and imaging was the carried out to a further degree within the study carried out in Chapter 6. Here it was shown that information captured by hyperspectral imaging devices, whilst they could not be utilised to predict human perception of non-colour traits, could be used to differentiate and estimate the different tomato storage conditions and storage length respectively alongside other

biochemical features. This is a novel finding in itself, which could aid the accuracy of ripeness estimation in computer vision systems. This ability shows the advantage of the use of hyperspectral imaging techniques, and subsequently computational illumination techniques, within machine vision systems.

7.2 Future Work

The bridging of human and computer vision is a research approach within which a large scope of future potential exists. It is felt that Chapter 4 offers a large remit for future work within the scope of this thesis with a greater focus upon end applications. The method of utilising human skin as a cue for illumination estimation stemmed from the hypothesis that the human visual system may utilise known objects as references for colour constancy. Findings from Chapter 5 neither support or work against this hypothesis but also confirm the complex nature of this visual perception mechanism, which continues to work under a variety of conditions. It also showed performance deviations in colour constancy across the daylight locus under two different scene conditions. However these scene conditions do not accurately reflect scenes in which colour constancy performs unnoticeably on a day-to-day basis, and under which we expect image processing algorithms to do the same. The process of aiming to utilise a single image, upon which to achieve perfect colour constancy may also differ from how the human visual system operates. Therefore one avenue to work will be that of testing colour constancy under more realistic conditions, whether this be through scene variables, or methods of image capture.

Appendix A

Database	Objects	Illumination	Scenes	Image Types	Given camera(s) characteristics	Image and scene characteristics	Image capture and scene settings	Data Types
Amsterdam Library of Object Images (ALOI) [103, 111]	• 1000 objects • Varied size, colour and texture • Simple object description	• 12 CCTs (2175 – 3175K). • Single light • Uniform • 24 directions under 1 illuminant	Indoor Black backdrop White balanced reference	RGB	None	768*576 pixels & ½ and ¼ resolutions Compressed JPEG	Scene layout. Explicit camera settings	24bit & GS
Tominaga [111]	• Daylight scenes of people and natural environments	• Daylight/ evening illumination • Multiple lights • Non uniform	Outdoor	RGB	Camera sensitivity under tungsten Weights for daylight - tungsten provided	Spectral power distribution of ambient light	Spectral sensitivity of camera Tungsten mode	24bit RGB
Simon Fraser University (A dataset for colour constancy) [104]	• Fluorescent, metallic, mondrian and specular objects • 51 objects	• 11 illuminants/ object • Single and multiple lights • Non uniform	Canonical illuminant reference Colour checker & white patch reference	RGB		Illuminant spectrum & RGB	Camera calibration provided	24bit RGB

Simon Fraser University (A large image database for colour constancy research) [105]	• Natural scenes	• Daylight and artificial light	Reference grey sphere	RGB		Illuminant spectrum & RGB	Camera calibration provided	24bit RGB
Simon Fraser University (A dataset for colour research) [104]	• 20 objects • 11 illuminants • Change in object pose	• 11 illuminants/ object • Single and multiple lights • Non uniform	Canonical reference image	RGB		Illuminant spectrum & RGB	Camera calibration provided	24bit RGB
Tarrlab (psychophysical image database)		• Alteration of colour of objects for perceptual testing		RGB				24bit RGB
Gehler image database [108]	• 568 images • Natural environment and people subjects	• Indoor and outdoor illumination • Single and multiple • Non uniform	Colour chart reference	RGB		JPG and RAW files provided.	Co-ordinates of colour chart under canonical provided	24bit RGB
Hong Kong Polytechnic	• 25 subjects • 1 illuminant	• Uniform single illumination	None reference	HSI	Hyperspectral image.	None	None	HSC

hyperspectral face database [106]					33 Bands (10nm step) 400 – 720nm			
McGill Image Database [109]	• 9 image types • 850 images	• Indoor/ Outdoor	None	RGB		None	Calibration provided	24bit RGB
David Foster Database [107]	• 8 scenes • Outdoor	• Multiple • Non uniform	Grey patch reference	RGB & HSI	33 bands. Average 10nm steps 1344*1024	Grey patch in each scene with photometer reading	Camera settings and set up provided	HSC
Nascimento et al. [107]	• 30 scenes • Indoor and outdoor	• Indoor and outdoor • Non uniform	None	HSI	31 bands Average 11nm steps 1024*1024 (10 bit), only central 820*820 recorded	None	Set up provided	HSC
Barcelona Image Database [110]	• 7 scene types • 49 images per type	• Indoor and outdoor • Non uniform	Grey sphere and camera calibration	RGB			Camera calibration	XYZ, 24bit RGB

Appendix B. Calibration Methodology

B.1 Introduction

Chapters 4 - 6 within utilise results gained from the utilisation of a novel LED illumination system [29] in conjunction with a calibrated hyperspectral imaging system. This chapter will detail the specific of this illumination system and how it was utilised in order to spectro-radiometrically calibrate a hyperspectral imaging system (Specim V10E).

B.2 Custom built LED Illumination system

The novel LED illumination system detailed in [29] contains 6 illumination lamps, comprised of 10 physical LED sources, outputting 9 different channels of illumination. The disparity between these two numbers is due to one illumination channel being comprised of two identical LED light sources. Figure 93 illustrates the spectrum of each of these channels, when at maximum intensity. The disparity in maximum luminance between each channel can easily be noted.

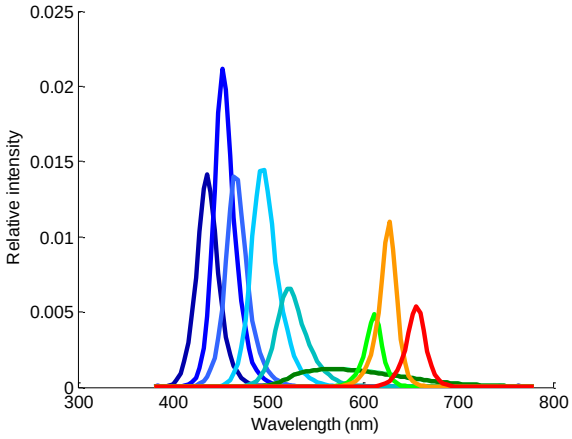


Figure 93 - The channels of the illuminator

B.2.1 *Controlling the output of the illuminator*

Each channel in each lamp can be varied in intensity using 16 bit resolution (values from 0-65535). This allows rudimentary control of the output of the lamps, however precise spectral control was also required. This led to a full characterisation of the illumination system to allow the creation of a metamer toolbox. This toolbox allowed the creation of a number of colourimetric metamers for a given illumination chromaticity.

B.3 Spectro-radiometric calibration of a hyperspectral camera

Within colour and imaging research spectro-radiometers have been utilised to capture accurate spectral and colourimetric information. These devices, whilst accurate, are not applicable for use when information across large spatial regions is required to be captured. However with the advent of affordable hyperspectral imaging cameras these have been utilised increasingly to capture this scene irradiance data.

A number of current hyperspectral datasets exist, as detailed within Chapter 2, however none of these datasets are able to provide the colourimetric data for each image, only spectral. This is because, whilst these imaging systems capture spectral information, they cannot immediately provide the same output as a calibrated spectrophotometer. This is because their outputs are 1) on different scales, and 2) may have some slight wavelength disparities. As such, in order to be able to retrieve this information a calibration must be undertaken. Of course from this colourimetric calibration only surfaces which are in focus can be assumed to be spectrally and colourimetrically accurate.

As such the purpose of the calibration is to allow equivalent spectra and colourimetric data to be gained from the hyperspectral imaging system as from the calibrated reference spectro-photometer.

The differences in the normalised recorded spectra between these two systems can be seen in below in Figure 94.

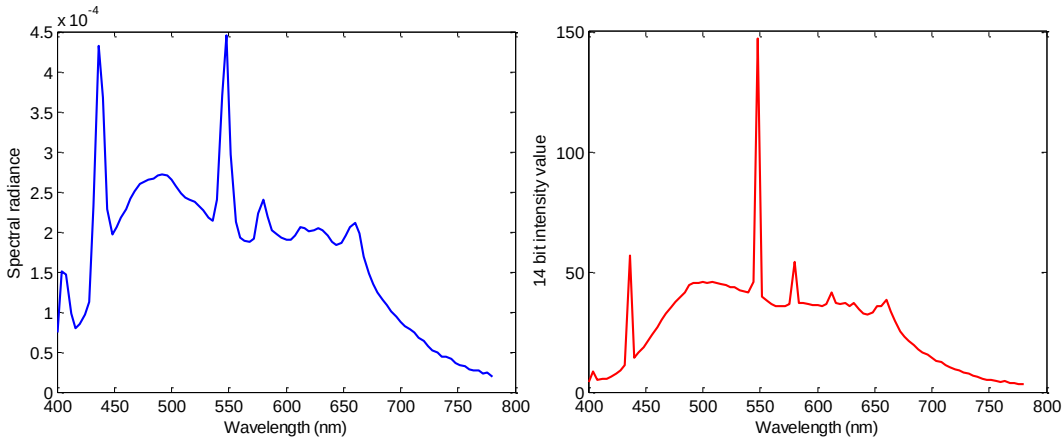
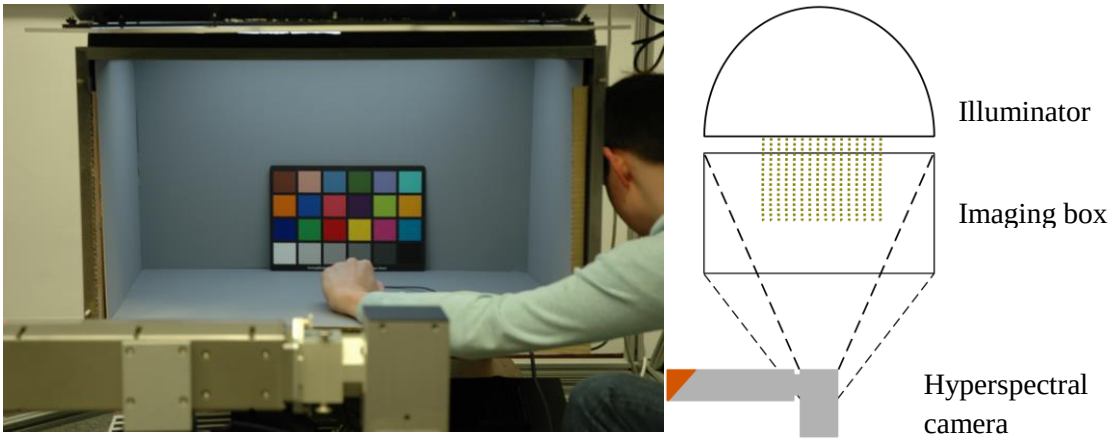


Figure 94 - A comparison of the recorded spectra of the black patch on the MacBeth Colour Checker Chart from a) a spectro-photometer, and b) the Specim V10E.

Our hyperspectral imaging setup is illustrated below;



a) b)

Figure 95 - Specim V10E imaging set up, a) real world, b) schematic.

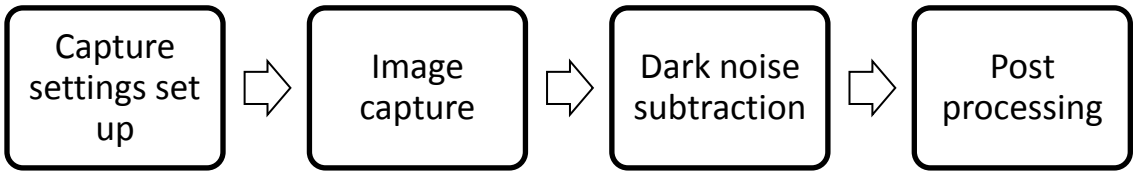


Figure 96 - Workflow of image capture devices.

Whereas for the spectral calibration of a light source required solely the characterisation of the light sources at different output levels, a hyperspectral camera offers a much greater number of variables for images capture. Each of these variables may affect spectrum capture, and as such had to be investigated. Table 16 describes the possible variables during hyperspectral image capture.

Variable	Variable type	Explanation
Lens	Mechanical	Different lenses may have different optical characteristics, such as transmission.
Aperture	Mechanical	Aperture size will affect the intensity of light reaching the sensor.
Exposure length	Capture setting	Exposure time for each line frame during capture.
Frame rate	Capture setting	The number of frames (lines) captured per second
Spectral binning	Capture setting	Spectral resolution of data captured for each pixel (1, 2, 4 and 8nm increments).
Spatial binning	Capture setting	Spectral resolution of data captured for each pixel. This occurs over square neighbourhood regions of 1, 2, 4 and 8 pixel increments.
Mirror Speed	Capture setting	The speed at which the translation mirror ($°s^{-1}$) moves across the scene. This in conjunction with spatial binning and frame rate determines the physical size represented by one pixel at a given imaging distance.
Incident illumination	Environmental	The calibration should perform well under any illumination. As such any resulting calibration error should be uniform across the entire visible spectrum.

Table 16 - Image capture variables with the Specim V10E hyperspectral camera.

Each of these variables were required to be investigated. Further to this, as a hyperspectral camera is a digital imaging device, different methods of dark noise cancellation were also investigated to ensure best performance.

B.3.1 *Basic experimental methodology*

In order to investigate the effect of each of these factors upon the spectral recordings of the hyperspectral imaging system hyperspectral images of a 24 patch MacBeth Colour Checker chart under a constant illumination from the L.E.D. lighting system [29] was utilised. The illumination used is shown below;

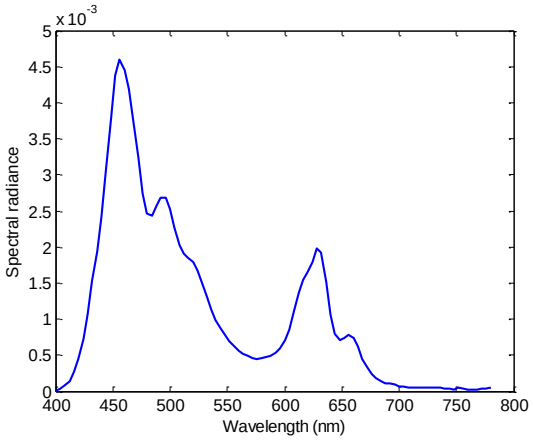


Figure 97 - The test illumination used for the calibration process.

The spectral irradiances for each of these patches were also captured using a calibrated PR-650 spectrophotometer and then directly compared to those captured using the hyperspectral imaging sensor. Two different metrics were then also utilised to investigate the effect of the imaging variables and to calibrate the system. Further to this due to the spatial intensity variation of illumination underneath the illumination system a white calibration tile, larger than the colour checker itself was also utilised. This allowed the capture of the illumination across each captured colour checker pixel. This then allowed investigations into the effect of this spatial illumination variable upon the resulting calibration accuracy to be undertaken.

B.4 Experimental results

B.4.1 *Wavelength alignment and spectral resolution*

The Specim V10E allows the capture of spectral data under 4 different spectral resolution settings; 1, 2, 4, and 8. This controlled the binning of the recorded spectra for each pixel, with samples being spaced 1, 2, 4 or 8nm apart respectively. However from testing different spectral offsets were noted from the Specim equipment under the 4 different spectral resolution settings, and was uniform across other capture setting variations. An example of this is shown below in Figure 97.

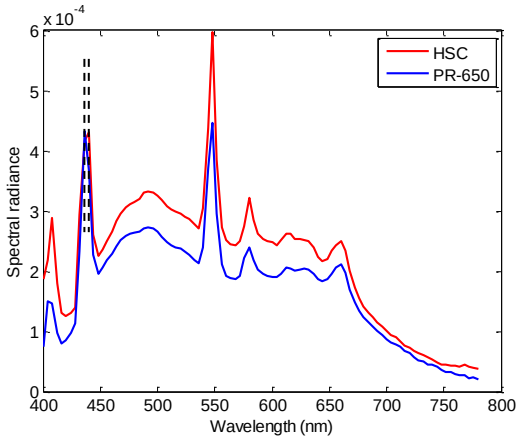


Figure 98 - Illustration of the wavelength alignment problems encountered under different spectral resolution capture settings.

From testing wavelength shift constants were calculated for each of the spectral resolutions, these are shown in Table 17 below;

Spectral Resolution	**Wavelength shift (nm)**
1	-0.30
2	-0.20
4	-1.10
8	-3.80

Table 17 - Wavelength shift for each spectral resolution

The effect of incorporating this wavelength shift was then tested across a number of different image capture settings, with the results compared utilising both the mean patch dE values (across the 24 patches of the MacBeth Colour Checker) and mean spectral errors. Table 18 contrasts the effect upon spectral error and mean patch dE of utilising the calculated wavelength shifts. Exposure below is defined as;

$$\boldsymbol{Exposure = Spectral * Spatial * Shutter} \tag{41}$$

Where spectral error (S.E.) is calculated, on a patch-by-patch basis, by firstly calculating the gain factor (r) for each measured wavelength component (in our case 96, spread from 400:780 in 4nm increments as represented by i) to normalise the amplitude of the hyperspectral camera spectra (S_{HSC}) to that of the spectrophotometer (S_{PR650});

$$\boldsymbol{r_i = \frac{S_{HSC_i}}{S_{PR650_i}}} \tag{42}$$

This gain factor is then used to factor the S_{HSC} spectrum to the same scale as the PR650 measurements, and subtract it from that of S_{PR650}. This vector is then normalised to give ΔS;

$$\Delta \boldsymbol{S} = \left(\widehat{\boldsymbol{S}_{\boldsymbol{PR650}_{\boldsymbol{\iota}}} - \left(\boldsymbol{r}_{\boldsymbol{\iota}} * \boldsymbol{S}_{\boldsymbol{HSC}_{\boldsymbol{\iota}}}\right)}\right) \quad (43)$$

The reference PR650 spectrum is then also normalised to give $\widehat{S_{PR650}}$. This allows the calculation of S.E.;

$$\boldsymbol{S.E.} = \left(\left(\frac{\Delta \boldsymbol{S}}{\widehat{\boldsymbol{S}_{\boldsymbol{PR650}}}}\right) * \mathbf{100}\right) \quad (44)$$

dE values were calculated by integrating spectra with the CIE XYZ 1964 standard colourimetric observer. Mean XYZ values were then converted into CIELAB space to allow the calculation of dE values.

					Original		Wavelength shift	
Aperture (AP)	**Spectral**	**Spatial**	**Shutter (ms)**	**Exposure**	**dE**	**S.E.**	**dE**	**S.E.**
2.8	1	1	125	125	6.88	9.94	6.75	10.42
2.8	1	2	60	120	7.19	11.13	6.95	11.50
2.8	2	1	65	130	6.96	9.61	6.69	9.73
2.8	2	2	30	120	6.63	9.71	6.45	9.83
4.2	1	1	210	210	7.05	7.32	6.78	7.74
4.2	1	2	105	210	7.08	7.57	6.78	8.04
4.2	2	1	110	220	6.89	7.20	6.62	7.37
4.2	2	2	50	200	6.94	7.69	6.64	7.84
5.6	1	1	332	332	6.97	7.94	6.82	8.47
5.6	1	2	220	440	7.04	8.17	6.88	8.66
5.6	2	1	300	600	6.86	7.73	6.57	7.90
5.6	2	2	160	640	6.88	7.81	6.61	8.06

Table 18 - The spectral error and mean dE effect of incorporating the appropriate wavelength shift.

As Table 18 shows the wavelength shift operation increases the mean spectral error, however it reduces our mean dE value. As the aim of our calibration is to gain accurate colourimetric data from hyperspectral images it can be seen that this wavelength shift aids us in reaching this goal.

B.4.2 *Dark noise counteracting methods*

Three methods of dark noise (DN) cancellation were investigated; mean value DN subtraction on a line-by-line basis, mean value (across multiple frames (100<)) DN subtraction (*meanDN*), and maximum value (across multiple frames (100<)) DN subtraction (*maxDN*). In order to compare these three DN counteracting methods a number of hyperspectral images were captured of a MacBeth Colour Checker chart. The average radiance for each patch was then retrieved, with no other gain or exposure

linearity processing except for the wavelength shift, using different DN cancellation techniques. The normalised average hyperspectral radiance for each patch was then compared to the normalised spectral radiance measured by the PR-650. For comparison we utilise two different metrics with the PR-650 measurement as the standard. Here we utilise only the metric of mean patch dE due to our concentration upon the colourimetric accuracy of our output.

					Patch dE	
Aperture (AP)	**Spectral**	**Spatial**	**Shutter (ms)**	**Exposure**	**maxDN**	**meanDN**
2.8	1	2	60	120	7.618464	6.949719
2.8	2	1	65	130	7.153473	6.692037
2.8	2	2	30	120	6.574065	6.44558
4.2	1	2	105	210	8.017262	6.783716
4.2	2	1	110	220	7.0752	6.619562
4.2	2	2	50	200	8.010856	6.638639
5.6	1	2	220	440	7.81383	6.878133
5.6	2	1	300	600	6.654976	6.567564
5.6	2	2	160	640	6.925581	6.610571

Table 19 - Effect of different DN removal methods upon error metrics.

Table 19 illustrates the effect upon our two error metrics of the two DN removal methods. As can be noted, on average, the mean DN frame method outperforms the removal of the frame maximum value. This points towards noise sources, such as those exhibiting shot noise, having a lesser effect than other more consistent sources of noise upon our imaging equipment in our test environment.

B.4.3 *Exposure linearity*

A sensor with a linear response would exhibit a linear response with regards to integration, and thus, exposure time. Exposure, when stated here, once again relates to the combination of spectral, spatial and shutter variables as previously described. Figure 99b shows how, if exposure is constant, sensor response linear across changes in either spectral or spatial resolution or shutter speed.

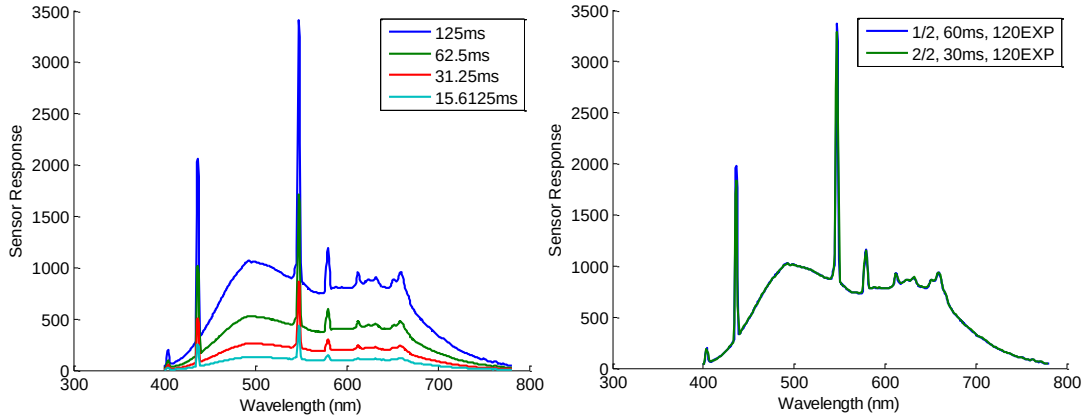


Figure 99 - Showing the linearity of recording as a) shutter speed is varied under 1/1 capture settings, and b) as exposure is varied through variation of spectral binning and shutter speed whilst preserving exposure.

This shows that no adjustments are required to account for changes in each of these constituent factors.

B.4.4 *Aperture linearity*

Whilst exposure may be linear, the effect of aperture size upon sensor response also needs to be investigated. In order to do this a number of images at different apertures sizes were captured, as detailed within both Table 18 and Table 19 . These were aperture sizes of 2.8, 4.2 and 5.6. Further aperture sizes, whose range is lens dependent, were also tested, but the results will not be illustrated here for brevity.

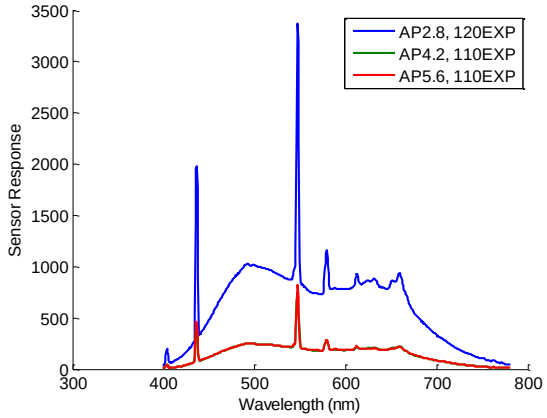


Figure 100 - Effect upon sensor response as lens aperture is varied.

Whilst the sensor response is consistent for aperture sizes of 4.2 and 5.6, the effect of the largest aperture size of AP2.8 can easily be noted. As such different calibration models needed to be created for each aperture setting on the lens.

B.4.5 *Lens dependency*

The different compositions of lenses, such as aperture shape, focal length, number of individual lenses and focal length, means that even at the same aperture and F-stop setting different levels of light pass through the lenses to the sensor. This physical characteristic is well known, and has been accounted for using the T-stop metric. As such the calibration method needs to build models for each lens being utilised.

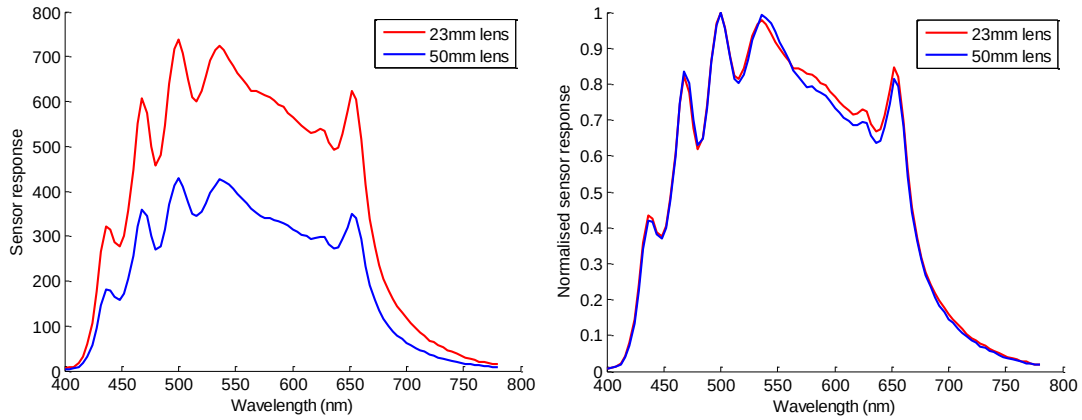


Figure 101 - Illustration of the different performance of lenses, in terms of a) light transmission, and b) wavelength transmission.

B.5 Calibration methodology

The previous sections have shown the need to create a calibration method and model for each lens, and each possible capture setting. On top of the aforementioned wavelength shift and dark noise correction, a wavelength-by-wavelength gain function dependent upon lens and aperture setting was required. In order to build this model for each possible setting a number of images, across settings such as those documented in Table 18, of a white calibration tile were collected. The measurements of this tile under the capture settings and those of the reference PR-650 measurements were then compared. This is after having interpolated the hyperspectral measurements down to the 4nm reading interval of the reference PR-650.

Using this comparison, aperture and lens specific average gain functions across the different exposures were modelled (wavelength-by-wavelength) and utilised to calibrate measurements from the hyperspectral camera. For each wavelength the calculation of this gain factor can be summarised in the following equation;

$$\boldsymbol{G_\lambda = \frac{PR650_\lambda}{HSC_\lambda}} \tag{45}$$

The overall calibration process applied immediately after image capture, altered from Figure 95, is outlined below;

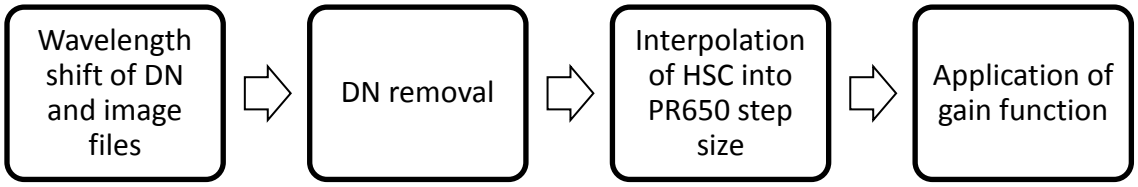


Figure 102 - Overview of the calibration process.

B.6 Calibration results

In order to test this calibration methodology, the calibration was applied to new images detailed below in Table 20 at an aperture of 2.8, and the average dE across the 24 patches of the MacBeth Colour Checker was calculated;

Spectral	Spatial	Shutter (ms)	Exposure	Average patch dE
1	1	332	332	1.50
1	2	166	332	1.60
1	2	332	664	0.97
2	1	166	332	1.58
2	1	332	664	1.23
2	2	83	332	1.71
2	2	166	664	1.22
2	2	332	1328	1.16
2	4	83	664	1.52
2	4	166	1328	1.17
2	4	332	2656	1.15
4	2	83	664	1.01
4	2	166	1328	0.93
4	2	332	2656	0.87
8	2	41.5	664	1.05
8	2	83	1328	1.03
8	2	166	2656	0.92
8	4	20.75	664	1.53
8	4	41.5	1328	1.08
8	4	83	2656	0.98

Table 20 - Viewing the effect of the wavelength calibration method upon average dE.

As can be seen, although the images listed within Table 18, Table 19 and Table 20 are not directly comparable the large reduction in average dE values across the colour checker patches can be noted. This proves that the calibration process does indeed work.

B.7 Summary

This appendix has given an overview of the considerations required when utilising a hyperspectral camera to make colourimetric measurements. The factors affecting hyperspectral measurement have also been broken down and investigated individually, allowing the validation of the proposed calibration process.

References

[1] A. Hurlbert, "Colour vision: Is colour constancy real?," *Current Biology,* vol. 9, pp. R558-R561, Jul-Aug 1999.

[2] D. H. Foster, "Does colour constancy exist?," *Trends in Cognitive Sciences,* vol. 7, pp. 439-443, Oct 2003.

[3] E. H. Land and J. J. Mccann, "Lightness and Retinex Theory," *Journal of the Optical Society of America,* vol. 61, pp. 1-&, 1971.

[4] J. D. Mollon, "“Tho'she kneel'd in that place where they grew…” The uses and origins of primate colour vision," *Journal of Experimental Biology,* vol. 146, pp. 21-38, 1989.

[5] M. D. Fairchild, *Color appearance models*: John Wiley & Sons, 2013.

[6] E. Kanematsu and D. H. Brainard, "No measured effect of a familiar contextual object on color constancy," *Color Research & Application,* pp. n/a-n/a, 2013.

[7] M. Olkkonen, T. Hansen, and K. R. Gegenfurtner, "Color appearance of familiar objects: Effects of object shape, texture, and illumination changes," *Journal of Vision,* vol. 8, p. 13, 2008.

[8] W. H. Ehrenstein and A. Ehrenstein, "Psychophysical methods," in *Modern techniques in neuroscience research*, ed: Springer, 1999, pp. 1211-1241.

[9] L. T. Maloney, "Illuminant estimation as cue combination," *Journal of Vision,* vol. 2, 2002.

[10] A. Hurlbert, "Computational models of color constancy," in *Perceptual Constancy: Why Things Look As They Do*, V. Walsh and J. Kulikowski, Eds., ed Cambridge: Cambridge University Press, 1998, pp. 283-322.

[11] J. N. Yang and S. K. Shevell, "Surface color perception under two illuminants: The second illuminant reduces color constancy," *Journal of Vision,* vol. 3, pp. 369-379, 2003.

[12] D. H. Foster and S. M. C. Nascimento, "Relational Color Constancy from Invariant Cone-Excitation Ratios," *Proceedings of the Royal Society B-Biological Sciences,* vol. 257, pp. 115-121, Aug 22 1994.

[13] M. G. Bloj and A. C. Hurlbert, "Does Mutual Illumination Improve Human Color Constancy," *Investigative Ophthalmology & Visual Science,* vol. 36, pp. S569-S569, Mar 15 1995.

[14] B. Xiao and D. H. Brainard, "Surface gloss and color perception of 3D objects," *Visual Neuroscience,* vol. 25, pp. 371-385, May-Jun 2008.

[15] A. Hurlbert, M. Vurro, and Y. Ling, "Colour constancy of polychromatic surfaces," *Journal of Vision,* vol. 8, pp. 1101-1101, 2008.

[16] Y. Ling, M. Vurro, and A. Hurlbert, "Surface chromaticity distributions of natural objects under changing illumination," in *Conference on Colour in Graphics, Imaging, and Vision*, 2008, pp. 263-267.

[17] A. C. Hurlbert and Y. Ling, "If it's a banana, it must be yellow: The role of memory colors in color constancy," *Journal of Vision,* vol. 5, pp. 787-787, 2005.

[18] K. Uchikawa, I. Kuriki, and Y. Tone, "Measurement of color constancy by color memory matching," *Optical Review,* vol. 5, pp. 59-63, Jan-Feb 1998.

[19] J. M. Kraft and D. H. Brainard, "Mechanisms of color constancy under nearly natural viewing," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 96, pp. 307-312, Jan 5 1999.

[20] G. Buchsbaum, "A Spatial Processor Model for Object Color-Perception," *Journal of the Franklin Institute-Engineering and Applied Mathematics,* vol. 310, pp. 1-26, 1980.

[21] E. H. Land, "Retinex Theory of Color-Vision," *Scientific American,* vol. 237, pp. 108-&, 1977.

[22] S. Crichton, B. Pearce, M. Mackiewicz, G. Finlayson, and A. Hurlbert, "The illumination correction bias of the human visual system," *Journal of Vision,* vol. 12, pp. 64-64, 2012.
[23] M. Mackiewicz, B. Pearce, S. Crichton, G. Finlayson, and A. Hurlbert, "Achromatic adjustment outdoors using MEMS reflective display," in *PERCEPTION*, 2012, pp. 1522-1522.
[24] B. Pearce, S. Crichton, M. Mackiewicz, G. Finlayson, and A. Hurlbert, "Colour Constancy by Illumination Matching in Real World Scenes," *Journal of Vision,* vol. 12, pp. 57-57, 2012.
[25] B. Pearce, S. Crichton, M. Mackiewicz, G. D. Finlayson, and A. Hurlbert, "Chromatic Illumination Discrimination Ability Reveals that Human Colour Constancy Is Optimised for Blue Daylight Illuminations," *PLOS ONE,* vol. 9, p. e87989, 2014.
[26] H.-C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," *JOSA A,* vol. 3, pp. 1694-1699, 1986.
[27] D. H. Foster, S. M. C. Nascimento, B. J. Craven, K. J. Linnell, F. W. Cornelissen, and E. Brenner, "Four issues concerning colour constancy and relational colour constancy," *Vision Research,* vol. 37, pp. 1341-1345, May 1997.
[28] M. R. Pointer, "Measuring Colour Rendering–A New Approach II," 2004.
[29] M. Mackiewicz, S. Crichton, S. Newsome, R. Gazerro, G. D. Finlayson, and A. Hurlbert, "Spectrally tunable LED illuminator for vision research," *Conference on Colour in Graphics, Imaging, and Vision,* vol. 2012, pp. 372-377, 2012.
[30] P. B. Delahunt and D. H. Brainard, "Does human color constancy incorporate statistical regularity of natural daylight?," *Journal of Vision,* vol. 4, pp. 57-81, 2004.
[31] J. Hernandez-Andres, J. L. Nieves, E. M. Valero, and J. Romero, "Spectral-daylight recovery by use of only a few sensors," *Journal of the Optical Society of America a-Optics Image Science and Vision,* vol. 21, pp. 13-23, Jan 2004.
[32] J. V. V. McCann, "Accurate Information vs. Looks Good: Scientific vs. Preferred Rendering," in *CGIV 2012*, 2012.
[33] D. B. Judd, D. L. MacAdam, G. Wyszecki, H. W. Budde, H. R. Condit, S. T. Henderson, *et al.*, "Spectral distribution of typical daylight as a function of correlated color temperature," *JOSA,* vol. 54, pp. 1031-1040, 1964.
[34] G. Finlayson, "Spectral Sharpening: What is it and why is it important?," *Cgiv'2002: First European Conference on Colour in Graphics, Imaging, and Vision, Conference Proceedings,* pp. 230-235, 2002.
[35] G. D. Finlayson, M. S. Drew, and B. V. Funt, "Spectral Sharpening - Sensor Transformations for Improved Color Constancy," *Journal of the Optical Society of America a-Optics Image Science and Vision,* vol. 11, pp. 1553-1563, May 1994.
[36] B. Bastani and B. Funt, "Spectral gamut mapping and gamut concavity," *Fifteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications, Final Program and Proceedings,* pp. 218-221, 2007.
[37] Y. D. Chen, R. S. Berns, and L. A. Taplin, "Improving the color constancy of prints by ink design," *Thirteenth Color Imaging Conference, Final Program and Proceedings,* pp. 159-164, 2005.
[38] P. M. Hubel, "Foveon technology and the changing landscape of digital cameras," *Thirteenth Color Imaging Conference, Final Program and Proceedings,* pp. 314-317, 2005.

[39] Specim. Available: http://www.specim.fi/index.php/products/research/imaging-spectrographs
[40] D. H. Brainard, "Cone contrast and opponent modulation color spaces," *Human color vision*, vol. 2, pp. 563-579, 1996.
[41] S. Tominaga, A. Ishida, and B. A. Wandell, "Illuminant estimation of natural scene using the sensor correlation method," *Aic: 9th Congress of the International Colour Association*, vol. 4421, pp. 918-921, 2002.
[42] S. Tominaga and B. A. Wandell, "Natural scene-illuminant estimation using the sensor correlation," *Proceedings of the Ieee*, vol. 90, pp. 42-56, Jan 2002.
[43] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern recognition*, vol. 40, pp. 1106-1122, 2007.
[44] M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes, "Proposal for a standard default color space for the Internet - sRGB," *Fourth Color Imaging Conference: Color Science, Systems and Applications*, pp. 238-246, 1996.
[45] I. E. Commission, "Multimedia systems and equipment - Colour measurement and management - Part 2-2: Colour management - Extended RGB colour space - scRGB ", ed, 2003.
[46] T. Matsumoto, Y. Shimpuku, T. Nakatsue, S. Haga, H. Eto, Y. Akiyama, *et al.*, "19.2: xvYCC: A New Standard for Video Systems using Extended-Gamut YCC Color Space," *SID Symposium Digest of Technical Papers*, vol. 37, pp. 1130-1133, 2006.
[47] S. A. Shafer, "Using Color to Separate Reflection Components," *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 1, pp. 1248-1248, 1984.
[48] H. Y. Chong, S. J. Gortler, and T. Zickler, "The von Kries hypothesis and a basis for color constancy," *2007 Ieee 11th International Conference on Computer Vision, Vols 1-6*, pp. 2143-2150, 2007.
[49] D. L. Macadam, "Chromatic adaptation," *J Opt Soc Am*, vol. 46, pp. 500-13, Jul 1956.
[50] G. Wyszecki, "A graphical interpretation of a three-components theory of chromatic adaptation in terms of the CIE chromaticity diagram," *J Opt Soc Am*, vol. 44, pp. 787-92, Oct 1954.
[51] E. H. Land, "Recent Advances in Retinex Theory and Some Implications for Cortical Computations - Color-Vision and the Natural Image," *Proceedings of the National Academy of Sciences of the United States of America-Physical Sciences*, vol. 80, pp. 5163-5169, 1983.
[52] 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 Sciences of the United States of America*, vol. 83, pp. 3078-3080, May 1986.
[53] B. Funt and L. L. Shi, "The Effect of Exposure on MaxRGB Color Constancy," *Human Vision and Electronic Imaging Xv*, vol. 7527, 2010.
[54] B. Funt and L. L. Shi, "MaxRGB Reconsidered," *Journal of Imaging Science and Technology*, vol. 56, Mar-Apr 2012.
[55] S. M. Wuerger, "Color vision: from genes to perception," *Perception*, vol. 29, pp. 629-630, 2000.
[56] D. H. Brainard and B. A. Wandell, "Analysis of the Retinex Theory of Color-Vision," *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 3, pp. 1651-1661, Oct 1986.
[57] J. van de Weijer and T. Gevers, "Color constancy based on the grey-edge hypothesis," *2005 International Conference on Image Processing (ICIP), Vols 1-5*, pp. 2069-2072, 2005.

[58] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Research & Application,* vol. 27, pp. 147-151, 2002.
[59] R. C. W. Gonzalez, R.E., *Digital Image processing*: Prentice Hall, 2002.
[60] Q. Wang and R. K. Ward, "Fast image/video contrast enhancement based on weighted thresholded histogram equalization," *Ieee Transactions on Consumer Electronics,* vol. 53, pp. 757-764, May 2007.
[61] S. D. Chen and R. Ramli, "Minimum mean brightness error bi-histogram equalization in contrast enhancement," *Ieee Transactions on Consumer Electronics,* vol. 49, pp. 1310-1319, Nov 2003.
[62] Y. T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *Ieee Transactions on Consumer Electronics,* vol. 43, pp. 1-8, Feb 1997.
[63] R. Hummel, "Image-Enhancement by Histogram Transformation," *Computer Graphics and Image Processing,* vol. 6, pp. 184-195, 1977.
[64] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *Ieee Transactions on Image Processing,* vol. 9, pp. 889-896, May 2000.
[65] V. Caselles, J. L. Lisani, J. M. Morel, and G. Sapiro, "Shape preserving local histogram modification," *Ieee Transactions on Image Processing,* vol. 8, pp. 220-230, Feb 1999.
[66] N. e. a. Limare, "Simplest Color Balance," *Image Processing On Line,* 2011.
[67] J.-M. Morel, A. B. Petro, and C. Sbert, "Toward a Natural Local Color Image Enhancement," *Conference on Colour in Graphics, Imaging, and Vision,* vol. 2012, pp. 211-216, 2012.
[68] J. H. Han, S. Yang, and B. U. Lee, "A Novel 3-D Color Histogram Equalization Method With Uniform 1-D Gray Scale Histogram," *Ieee Transactions on Image Processing,* vol. 20, pp. 506-512, Feb 2011.
[69] G. Finlayson, S. Hordley, G. Schaefer, and G. Y. Tian, "Illuminant and device invariant colour using histogram equalisation," *Pattern Recognition,* vol. 38, pp. 179-190, Feb 2005.
[70] D. A. Forsyth, "A novel method for color constancy," *International Journal of Computer Vision,* vol. 5, pp. 5-36, 1990.
[71] D. A. Forsyth, "A Novel Algorithm for Color Constancy," *International Journal of Computer Vision,* vol. 5, pp. 5-36, Aug 1990.
[72] G. Finlayson and S. Hordley, "A theory of selection for gamut mapping colour constancy," *1998 Ieee Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings,* pp. 60-65, 1998.
[73] G. Finlayson and S. Hordley, "Selection for gamut mapping colour constancy," *Image and Vision Computing,* vol. 17, pp. 597-604, Jun 1999.
[74] H. R. V. Joze and M. S. Drew, "White Patch Gamut Mapping Colour Constancy," *2012 Ieee International Conference on Image Processing (Icip 2012),* pp. 801-804, 2012.
[75] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Color by correlation: A simple, unifying framework for color constancy," *Ieee Transactions on Pattern Analysis and Machine Intelligence,* vol. 23, pp. 1209-1221, Nov 2001.
[76] G. D. Finlayson, P. Morovic, and S. D. Hordley, "The chromagenic colour camera and illuminant estimation," *Thirteenth Color Imaging Conference, Final Program and Proceedings,* pp. 20-24, 2005.
[77] R. Shrestha and J. Y. Hardeberg, "Computational color constancy using chromagenic filters in color filter arrays," *Sensors, Cameras, and Systems for Industrial and Scientific Applications Xiii,* vol. 8298, 2012.

[78] C. Fredembach and G. D. Finlayson, "Bright chromagenic algorithm for illuminant estimation," *Journal of Imaging Science and Technology,* vol. 52, Jul-Aug 2008.
[79] N. Wang, B. Funt, C. Y. Lang, and D. Xu, "Video-Based Illumination Estimation," *Computational Color Imaging,* vol. 6626, pp. 188-198, 2011.
[80] S. Susstrunk, J. Holm, and G. D. Finlayson, "Chromatic adaptation performance of different RGB sensors," *Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts Vi,* vol. 4300, pp. 172-183, 2001.
[81] P. L. Vora, J. E. Farrell, J. D. Tietz, and D. H. Brainard, *Digital Color Cameras: Spectral Response*: Citeseer, 1997.
[82] G. D. Finlayson, S. D. Hordley, and R. Xu, "Convex programming colour constancy with a diagonal-offset model," *2005 International Conference on Image Processing (ICIP), Vols 1-5,* pp. 2617-2620, 2005.
[83] K. Barnard and B. Funt, "Color constancy with specular and non-specular surfaces," *Seventh Color Imaging Conference: Color Science, Systems and Applications,* pp. 114-119, 1999.
[84] G. Finlayson and G. Schaefer, "Single surface colour constancy," *Seventh Color Imaging Conference: Color Science, Systems and Applications,* pp. 106-113, 1999.
[85] L. Shi and B. Funt, "Skin colour imaging that is insensitive to lighting conditions," *Proc. AIC 2008 Color for Science and Industry,* pp. 15-18.
[86] B. Funt and M. Mosny, "Color calibration via natural food colors."
[87] S. Battiato, A. Bosco, A. Castorina, and G. Messina, "Automatic image enhancement by content dependent exposure correction," *Eurasip Journal on Applied Signal Processing,* vol. 2004, pp. 1849-1860, Sep 15 2004.
[88] B. Martinkauppi, M. Soriano, and M. Pietikainen, "Detection of skin color under changing illumination: a comparative study," in *Image Analysis and Processing, 2003. Proceedings. 12th International Conference on*, 2003, pp. 652-657.
[89] M. Störring, T. Kočka, H. J. Andersen, and E. Granum, "Tracking regions of human skin through illumination changes," *Pattern Recognition Letters,* vol. 24, pp. 1715-1723, 2003.
[90] M. Soriano, B. Martinkauppi, S. Huovinen, and M. Laaksonen, "Skin detection in video under changing illumination conditions," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, 2000, pp. 839-842.
[91] M. Storring, H. J. Andersen, and E. Granum, "A multispectral approach to robust human skin detection," in *Conference on Colour in Graphics, Imaging, and Vision*, 2004, pp. 110-115.
[92] N. Sebe, I. Cohen, T. S. Huang, and T. Gevers, "Skin detection: A Bayesian network approach," *Proceedings of the 17th International Conference on Pattern Recognition, Vol 2,* pp. 903-906, 2004.
[93] R. Khan, A. Hanbury, J. Stöttinger, and A. Bais, "Color based skin classification," *Pattern Recognition Letters,* vol. 33, pp. 157-163, 2012.
[94] M. Storring and E. Granum, "Adapting a statistical skin colour model to illumination changes," in *Conference on Colour in Graphics, Imaging, and Vision*, 2002, pp. 16-21.
[95] M. Störring, E. Granum, and H. J. Andersen, "Estimation of the illuminant colour using highlights from human skin," in *First International Conference on Color in Graphics and Image Processing*, 2000, pp. 45-50.
[96] M. Storring, H. J. Andersen, and E. Granum, "Estimation of the illuminant colour from human skin colour," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, 2000, pp. 64-69.

[97] J. B. Martinkauppi and G. D. Finlayson, "Designing a simple 3-channel camera for skin detection," in *Color and Imaging Conference*, 2004, pp. 150-156.
[98] M. Störring, H. J. Andersen, and E. Granum, "Physics-based modelling of human skin colour under mixed illuminants," *Robotics and Autonomous Systems*, vol. 35, pp. 131-142, 2001.
[99] O. Y. Kwon and S. I. Chien, "Estimation of Illuminant Chromaticity Based on Highlight Detection for Face Images with Varying Illumination," *Color Research and Application*, vol. 39, pp. 188-199, Apr 2014.
[100] S. D. O. Cotton and E. Claridge, "Do all human skin colours lie on a defined surface within LMS space?," *SCHOOL OF COMPUTER SCIENCE RESEARCH REPORTS-UNIVERSITY OF BIRMINGHAM CSR*, 1996.
[101] J. Ruiz-del-Solar and R. Verschae, "Robust skin segmentation using neighborhood information," *Icip: 2004 International Conference on Image Processing, Vols 1- 5*, pp. 207-210, 2004.
[102] K. Barnard, L. Martin, A. Coath, and B. Funt, "A comparison of computational color constancy algorithms - Part II: Experiments with image data," *Ieee Transactions on Image Processing*, vol. 11, pp. 985-996, Sep 2002.
[103] J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The Amsterdam Library of Object Images," *International Journal of Computer Vision*, vol. 61, pp. 103-112, Jan 2005.
[104] K. Barnard, K. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Research and Application*, vol. 27, pp. 147-151, Jun 2002.
[105] F. Ciurea and B. Funt, "A large image database for color constancy research," *Eleventh Color Imaging Conference: Color Science and Engineering - Systems, Technologies, Applications*, pp. 160-164, 2003.
[106] W. Di, L. Zhang, D. Zhang, and Q. A. Pan, "Studies on Hyperspectral Face Recognition in Visible Spectrum With Feature Band Selection," *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans*, vol. 40, pp. 1354-1361, Nov 2010.
[107] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 23, pp. 2359-2372, Oct 2006.
[108] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, "Bayesian color constancy revisited," *2008 Ieee Conference on Computer Vision and Pattern Recognition, Vols 1-12*, pp. 3291-3298, 2008.
[109] A. Olmos and F. A. A. Kingdom, "A biologically inspired algorithm for the recovery of shading and reflectance images," *Perception*, vol. 33, pp. 1463-1473, 2004.
[110] C. Parraga, J. Vazquez-Corral, and M. Vanrell, "A new cone activation-based natural images dataset," *Perception*, vol. 38, pp. 180-180, 2009.
[111] S. Tominaga, "Natural image database and its use for scene illuminant estimation," *Journal of Electronic Imaging*, vol. 11, pp. 434-444, 2002.
[112] B. L. Anderson and J. Kim, "Image statistics do not explain the perception of gloss and lightness," *Journal of Vision*, vol. 9, p. 10, 2009.
[113] L. Sharan, Y. Li, I. Motoyoshi, S. y. Nishida, and E. H. Adelson, "Image statistics for surface reflectance perception," *JOSA A*, vol. 25, pp. 846-865, 2008.
[114] I. Motoyoshi, S. y. Nishida, L. Sharan, and E. H. Adelson, "Image statistics and the perception of surface qualities," *Nature*, vol. 447, pp. 206-209, 2007.
[115] M. F. Hossain, M. R. Alsharif, and K. Yamashita, "An Approach to Color Image Enhancement Using Minimum Mean Brightness Error Dynamic Histogram Equalization," *International Journal of Innovative Computing Information and Control*, vol. 7, pp. 827-840, Feb 2011.

[116] A. Hore and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 2366-2369.
[117] J. Kim, P. Marlow, and B. L. Anderson, "The perception of gloss depends on highlight congruence with surface shading," *Journal of Vision*, vol. 11, p. 4, 2011.
[118] J. L. Dannemiller, "Rank orderings of photoreceptor photon catches from natural objects are nearly illuminant-invariant," *Vision research*, vol. 33, pp. 131-140, 1993.
[119] S. Bianco and R. Schettini, "Adaptive Color Constancy Using Faces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PP, pp. 1-1, 2014.
[120] J. Roca-Vila, C. A. Parraga, and M. Vanrell, "Chromatic settings and the structural color constancy index," *Journal of Vision*, vol. 13, 2013.
[121] S. Crichton, J. Pichat, M. Mackiewicz, G.-Y. Tian, and A. Hurlbert, "Skin chromaticity gamuts for illumination recovery," *Conference on Colour in Graphics, Imaging, and Vision*, vol. 2012, pp. 266-271, 2012.
[122] R. T. Eskew, McLellan, and F. Giulianini, "Chromatic detection and discrimination," in *Color vision: from genes to perception*, K. R. Gegenfurtner and L. T. Sharpe, Eds., ed Cambridge: Cambridge University Press, 1999, pp. 345-368.
[123] D. H. Foster, K. Amano, and S. M. C. Nascimento, "Relational colour constancy under rapid changes of illuminant on natural scenes," *Perception*, vol. 29, pp. 44-44, 2000.
[124] H. S. Chen, T. M. Wang, S. H. Chen, and J. S. Liu, "Skin-Color Correction Method Based on Hue Template Mapping for Wide Color Gamut Liquid Crystal Display Devices," *Color Research and Application*, vol. 36, pp. 335-348, Oct 2011.
[125] F. Ciurea and B. Funt, "Failure of luminance-redness correlation for illuminant estimation," *12th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*, pp. 42-46, 2004.
[126] J. Golz and D. I. A. MacLeod, "Influence of scene statistics on colour constancy," *Nature*, vol. 415, pp. 637-640, Feb 7 2002.
[127] R. S. Berns, S. R. Fernandez, and L. Taplin, "Estimating black-level emissions of computer-controlled displays," *Color Research & Application*, vol. 28, pp. 379-383, 2003.
[128] E. A. Cooper, H. M. Jiang, V. Vildavski, J. E. Farrell, and A. M. Norcia, "Assessment of OLED displays for vision research," *Journal of Vision*, vol. 13, 2013.
[129] M. Mackiewicz, B. Pearce, S. Crichton, A. Hurlbert, and G. Finlayson, "Achromatic adjustment outdoors and indoors using the Mirasol reflective display," in *AIC 2013*, Newcastle-upon-Tyne, 2013, pp. 295-298.
[130] H. Lee, M. S. Kim, D. Jeong, K. Chao, B. K. Cho, and S. R. Delwiche, "Hyperspectral Near-infrared Reflectance Imaging for Detection of Defect Tomatoes," *Sensing for Agriculture and Food Quality and Safety Iii*, vol. 8027, 2011.
[131] I. Kim, C. Z. Xu, and M. S. Kim, "Poultry skin tumor detection in hyperspectral images using radial basis probabilistic neural network," *Advances in Neural Networks - Isnn 2006, Pt 3, Proceedings*, vol. 3973, pp. 770-776, 2006.
[132] D. Wu and D.-W. Sun, "Potential of time series-hyperspectral imaging (TS-HSI) for non-invasive determination of microbial spoilage of salmon flesh," *Talanta*, vol. 111, pp. 39-46, 2013.

[133] D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review—Part II: Applications," *Innovative Food Science & Emerging Technologies,* vol. 19, pp. 15-28, 2013.
[134] D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review—Part I: Fundamentals," *Innovative Food Science & Emerging Technologies,* vol. 19, pp. 1-14, 2013.
[135] A. A. Gowen, C. O'Donnell, P. J. Cullen, G. Downey, and J. M. Frias, "Hyperspectral imaging–an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology,* vol. 18, pp. 590-598, 2007.
[136] R. Arias, T.-C. Lee, L. Logendra, and H. Janes, "Correlation of Lycopene Measured by HPLC with the L*, a*, b* Color Readings of a Hydroponic Tomato and the Relationship of Maturity with Color and Lycopene Content," *Journal of Agricultural and Food Chemistry,* vol. 48, pp. 1697-1702, 2000/05/01 2000.
[137] P. Rajkumar, N. Wang, G. Elmasry, G. S. V. Raghavan, and Y. Gariepy, "Studies on banana fruit quality and maturity stages using hyperspectral imaging," *Journal of Food Engineering,* vol. 108, pp. 194-200, Jan 2012.
[138] P. Talens, L. Mora, N. Morsy, D. F. Barbin, G. ElMasry, and D.-W. Sun, "Prediction of water and protein contents and quality classification of Spanish cooked ham using NIR hyperspectral imaging," *Journal of Food Engineering,* vol. 117, pp. 272-280, 2013.
[139] G. Polder, G. W. A. M. Van der Heijden, and I. T. Young, "Hyperspectral image analysis for measuring ripeness of tomatoes," presented at the 2000 ASAE International Meeting, Wisconsin, U.S.A., 2000.
[140] D. Liu, D. W. Sun, and X. A. Zeng, "Recent Advances in Wavelength Selection Techniques for Hyperspectral Image Processing in the Food Industry," *Food and Bioprocess Technology,* vol. 7, pp. 307-323, Feb 2014.
[141] G. ElMasry, N. Wang, A. ElSayed, and M. Ngadi, "Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry," *Journal of Food Engineering,* vol. 81, pp. 98-107, 2007.