COMPUTATIONAL COLOR CONSTANCY: TAKING THEORY INTO PRACTICE

by

Kobus Barnard

B.Sc. Computing Science

Simon Fraser University 1990

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in the School

of

Computing Science

SIMON FRASER UNIVERSITY

August 1995

APPROVAL

Name: Kobus Barnard

Degree: Master of Science

Title of thesis: Computational Color Constancy: Taking Theory into Practice

Examining Committee: Dr. Robert F. Hadley
Chair

Dr. Stella Atkins, External Examiner
Associate Professor, Computing Science, SFU

Dr. Brian V. Funt, Senior Supervisor
Professor, Computing Science, SFU

Dr. Ze-nian Li, Supervisor
Associate Professor, Computing Science, SFU

Date Approved: _____

Abstract

The light recorded by a camera is a function of the scene illumination, the reflective characteristics of the objects in the scene, and the camera sensors. The goal of color constancy is to separate the effect of the illumination from that of the reflectances. In this work, this takes the form of mapping images taken under an unknown light into images which are estimates of how the scene would appear under a fixed, known light.

The research into color constancy has yielded a number of disparate theoretical results, but testing on image data is rare. The thrust of this work is to move towards a comprehensive algorithm which is applicable to image data.

Necessary preparatory steps include measuring the illumination and reflectances expected in real scenes, and determining the camera response function. Next, a number of color constancy algorithms are implemented, with emphasis on the gamut mapping approach introduced by D. Forsyth and recently extended by G. Finlayson. These algorithms all assume that the color of the illumination does not vary across the scene. The results of these algorithms running on a variety of images, as well as on generated data, are

presented. In addition, the possibility of using sensor sharpening to improve algorithm performance is investigated.

The final part of this work deals with images of scenes where the illumination is not necessarily constant. A recent promising result from Finlayson, Funt, and Barnard demonstrates that if the illumination variation can be identified, it can be used as a powerful constraint. However, in its current form this algorithm requires human input and is limited to using a single such constraint. In this thesis the algorithm is first modified so that it provides conjunctive constraints with the other gamut mapping constraints and utilizes all available constraints due to illumination variation. Then a method to determine the variation in illumination from a properly segmented image is introduced. Finally the comprehensive algorithm is tested on simple images segmented with region growing. The results are very encouraging.

Acknowledgments

It would have been impossible to complete this work without the help of others. Certainly I thank my supervisor, Brian Funt, for encouraging me to study color computer vision several years ago, and for allowing me to follow my own interests once I started doing so. Thanks also go to all committee members for taking the time to read yet another thesis at a busy time of year.

Most of this work follows directly from that of Graham Finlayson, and I appreciate his patience in repeatedly explaining his contributions to me. Mark Drew also proved to be a good resource for the mathematical parts of computer vision. As excellent lab mates, Janet Dueck and Shubo Chatterjee helped in many ways.

Special thanks go to Emily Butler for being supportive of my work despite too many late night programming sessions, as well as for help with the proof-reading.

Finally, since this work was supported by NSERC, thanks go the Canadian tax payer.

Contents

List of Tables

List of Figures

Introduction

The purpose of a vision system, whether biological or man-made, is to determine properties of the world by sampling light. The nature of this signal is a function of both the viewed objects and the illumination. Color constancy exists as a research endeavor because illumination changes large enough to confound the process of world discovery are commonplace. As an example, consider an office with a window exposed to blue sky. Since blue sky is substantially more blue than typical indoor lighting, the sensors of a vision system will see the wall near the window as much bluer than the rest. Thus the simple question of the color of the wall becomes complex. In order to answer this question, it is necessary to model the interplay between the scene objects, the illumination, and the sampling apparatus. Thus the essence of color constancy research is the development of these models. This thesis deals with the applicability of current models to image data, with the eventual goal

being a completely automated color constancy algorithm which only fails when there is a provable paucity of information.

An example of robust color constancy processing is the human vision system. Although the exact nature of the processing is far from known (and it is beyond the scope of this work to discuss it in detail), a human observer would typically not notice the illumination variation described above. Although this thesis deals almost exclusively with machine vision, human color constancy processing provides a compelling hint that it is a desirable component of a vision system. Furthermore it adds interest to the computational problems inherent in color constancy.

This thesis builds on a large body of color constancy work which is outlined in chapter one. First, a few basic concepts used in this research area are introduced for future use. These include the imaging equation, small-dimensional models of spectra and reflectances, and diagonal models for color constancy. Then the gamut mapping approach to color constancy [For90, Fin95] is discussed in some detail, as a large part of this work involves this approach. In addition, algorithms used for comparison purposes, as well as several others included for a complete survey, are described. Next, the literature on the problem of color constancy for scenes with varying illumination is summarized to prepare the reader for the work introduced in chapter four. Finally, previous color constancy experiments using image data are reviewed.

In chapter two, measurements of illumination, surface reflectance, and camera properties, are presented. These parameters are important for the

algorithms used, and if these algorithms are to work on real image data, then data representative of the world is required.

Chapter three is concerned with the implementation and testing of a number of current color constancy algorithms. Results for tests on both a large generated data set and four images are provided. The algorithms are all variations on existing ones, although several of these variations have not been previously implemented. Also in this chapter, the use of sensor sharpening to improve color constancy performance is investigated.

The last chapter in the body of this work tackles the problem of color constancy in the case where the illumination in a scene varies. Recent work [FFB95] has indicated that varying illumination can be a useful constraint for color constancy. A major part of this research is extending this work to create a comprehensive algorithm applicable to image data. The first step is to modify the algorithm so that the constraints due to the varying illumination can be used in conjunction with the constraints used in the other gamut mapping algorithms. Then the problem of identifying the illumination variations is dealt with. Finally the very encouraging preliminary results on simple images are presented.

Chapter One

Color Constancy Overview

This chapter provides background for the following three chapters. By necessity, investigating the applicability of current color constancy ideas to image data relies heavily on the work of others. This work is outlined below, but prior to that a brief introduction is provided for some basic concepts used in color constancy.

1.1 Problem Definition

Denote the light energy reaching a surface as $E(\lambda)$. For a given scene and viewing geometry, the fraction of light reflected back defines the reflectivity of the surface, and is denoted by $S(\lambda)$. In the commonly assumed case of a Lambertion reflector, $S(\lambda)$ is independent of the angle of the surface to the camera, but it should be noted that the definition of reflectivity does not require this. A vision system samples image locations with one or more sensor types. In our case, the locations are simply image pixels, and the

sensor types are the red, green, and blue camera channels. The response of the i'th sensor, ρ_i, is often modeled by assuming sensor response functions $R_i(\lambda)$ such that:

$$\rho_i = \int_\lambda R_i(\lambda)S(\lambda)E(\lambda)d\lambda \tag{1.1}$$

This model is normally assumed for the human visual system (see for example [WS82]), and forms the basis for the CIE[1] colorimetry standard. In the case of a camera, the model can be verified (see §2.1). $R_i(\lambda)$ is a function of aperture and other optical parameters and thus care must be taken to set them consistently.

As mentioned above, the goal is to derive some characterization of the reflectances from the sensor responses. It is a difficult problem because it is severely under-constrained. Given a sensor response, there are many possible E(λ) and S(λ) that could account for it. In its degenerate form, the color constancy problem is analogous to determining two numbers from their product. More information is required to solve the problem, and therefore additional assumptions about the world are made in creating color constancy algorithms.

Essentially we wish to characterize S(λ) (and symmetrically E(λ)), from insufficient data from the sensors. A slightly different way of looking at the problem proved convenient for the expression of David Forsyth's innovative algorithm [For90]. Here the goal is expressed as transforming an image viewed under an unknown light into one showing how that scene would look under a known, canonical, light. This approach gets to the heart of the matter for the simple reason that color is only defined with respect to the sensors. A vision system cannot recognize an object by reflectance

[1]Commission International de l'Éclairage or International Commission on Illumination

properties outside the system sensitivity range. Furthermore, it is possible that an algorithm that is trying to solve for the light may lose accuracy in trying to minimize error in regions of low sensor response.[2] The mapping approach described above cleanly avoids such problems, and in general leads to a simpler expression of color constancy algorithms. In this thesis the mapping approach will be used exclusively.

1.2 Linear Models

One effective method to reduce the number of degrees of freedom inherent in the functions S(λ) and E(λ) is to use small dimensional linear models of these functions. Such models have been demonstrated to provide good fits for E(λ) in the case of daylight [JMW64, Dix71], as well as for a large number of surfaces [Coh64, PHJ89, MW92]. An expression of an N dimension linear model for surfaces is:

$$S(\lambda) \approx \sum_{i=0}^{N} \sigma_i S_i(\lambda) \tag{1.2}$$

Here $S_i(\lambda)$ are the basis functions and σ_i are the projections. Similarly, a linear model for illuminants is expressed as:

$$E(\lambda) \approx \sum_{i=0}^{N} \varepsilon_i E_i(\lambda) \tag{1.3}$$

The idea is useful because even a small number of basis functions yields excellent approximations. For example, in [PHJ89] the spectra of 1257 Munsell color chips were fit to 90% (absolute, averaged over the spectrum) with 4 basis functions, and to 98% with 8 basis functions. The number of

---

[2]For example, the Maloney-Wandell algorithm [MW86] implicitly has this problem with respect to the illumination.

basis functions required to fit daylight is even smaller. Dixon [Dix71] found that for a daylight data set taken at one location, three basis functions accounted for 99% of the variance, and for another data set, four functions accounted for 92% of the variance. (Note that the error measure here is not directly comparable to the one used above for the Munsell color chips.) The $S_i(\lambda)$ and $E_i(\lambda)$ are found using the mathematical technique of singular value decomposition, or on occasion, by principal component analysis.[3]

Finite dimensional models allow image formation to be modeled compactly using matrices. For example, assuming three dimensional surfaces, we can define a lighting matrix for a given illuminant $E(\lambda)$ by:

$$\Lambda = \begin{vmatrix} \int E(\lambda)S_1(\lambda)R_1(\lambda) & \int E(\lambda)S_2(\lambda)R_1(\lambda) & \int E(\lambda)S_3(\lambda)R_1(\lambda) \\ \int E(\lambda)S_1(\lambda)R_2(\lambda) & \int E(\lambda)S_2(\lambda)R_2(\lambda) & \int E(\lambda)S_3(\lambda)R_2(\lambda) \\ \int E(\lambda)S_1(\lambda)R_3(\lambda) & \int E(\lambda)S_2(\lambda)R_3(\lambda) & \int E(\lambda)S_3(\lambda)R_3(\lambda) \end{vmatrix} \tag{1.4}$$

Then for a surface $\sigma = (\sigma_1, \sigma_2, \sigma_3)^{\mathbf{T}}$, the response $\rho = (\rho_1, \rho_2, \rho_3)^{\mathbf{T}}$ is given simply as:

$$\rho = \Lambda\sigma \tag{1.5}$$

Such linear models are intimately linked with many color constancy algorithms, and they appear repeatedly in the literature. This is not so much the case with respect to the algorithms central to this thesis, and thus linear models will be used explicitly in this document only occasionally.[4]

---

[3]The difference between SVD and PCA is slight. In PCA the mean of the sample set is subtracted before the basis functions are determined by SVD

[4]For example, equation (1.5) will be used in §3.2 to explain perfect sensor sharpening.

1.3 Chromaticity Spaces

Often we wish to deal with the RGB responses without magnitude, and for this we will use a two-component space where the magnitude has been factored out. For example, such a space can be obtained by dividing each component by the sum of all the components. Although this is perhaps the most common approach, in order to preserve some useful properties (as explained in §1.4.2), I will use a space obtained by dividing each component by the blue component. The third component is thus invariably one, and can be ignored. The choice of blue is arbitrary. For the remainder of this document, when the term *chromaticity space* is used without qualification, it will refer to this particular space.

1.4 Diagonal Transforms for Color Constancy

The simplest method to change the chromaticity of an image is to multiply each RGB value by the same triplet of scale factors. Thus each channel is adjusted independently, but each pixel is adjusted in the same way. This is essentially von Kries' adaptation, which dates back to the late 1800's, reinterpreted in the context of modern color constancy algorithms.[5] Many color constancy algorithms use such a "coefficient rule" including

[5]Some history of early color constancy is available in [WB82]. See also [BW81] and [For90].

Land's Retinex algorithm[6] [LM71, MMT75, Lan77, Lan86], Forsyth's CRULE algorithm [For90], and Finlayson's recent simplification and extension [Fin95]. All these algorithms assume that the illumination does not vary across the scene, except for Finlayson's which assumes that the chromaticity of the illumination does not vary. Since coefficient rules are restricted cases of linear transformations such that the corresponding post multiplication matrix is diagonal, coefficient rules are equally referred to as diagonal transforms.

It has been observed that the suitability of diagonal transforms is partly a function of the sensors. Overlap, linear independence, and most importantly, sharpness are considerations (see, for example, [BW81, WB82, For90]). Intuitively, if the sensors are delta functions, then the diagonal transform model follows directly from equation (1.1). Going further, Finlayson et. al. [FDF94a, FDF94b] have shown that by applying the appropriate linear transforms to the data, the coefficient rule can be made to work better; in essence, the sensors are "sharpened". This original work was based on published sensor functions which purport to model human cone responses.

To explain further the nature of the computation, let **T** be a matrix (to be determined) of the same dimension as the number of sensors. Let **U** be a matrix whose rows are the pixel RGB values of the image under the unknown illuminant. Then the transformed data is **UT**. (Using post

---

[6][MMT75] has an adjustment to the basic computation which makes Retinex a non-coefficient rule. This adjustment was added in response to measurements of the human visual system. It is not a reasonable way to implement Retinex in the case of linear camera.

multiplication for the mappings proves to be more natural). On the assumption that a diagonal model is now more applicable, we apply a coefficient rule algorithm with **UT** as input. The result of a coefficient rule algorithm is a diagonal matrix **D** (with the coefficients along the diagonal). Once we are done, our estimate **C**, for the pixel RGB values under the canonical illuminant is given by:

$$\mathbf{C} = \mathbf{UTDT}^{-1} \tag{1.6}$$

It should be noted that applying sharpening leads to an non-coefficient rule algorithm, since $\mathbf{TDT}^{-1}$ is not normally diagonal. Thus one way to look at sharpening is that it allows one to use a more powerful mapping function, but apply well motivated coefficient rule based algorithms. We must be able to do better using this approach, simply because the identity matrix is a possible candidate for **T**. Some of the available methods for calculating **T** are discussed in §3.3.

The majority of the algorithms explored in this thesis are coefficient rules. As part of this research I verified that the camera sensors are already quite sharp in the sense that the best possible diagonal result is within 10% of that which can be obtained with sharpening. Thus good results should be obtainable without sharpening. Nonetheless, sharpening the camera sensors was explored as part of this research. The initial goal was simply to maximize the performance of the algorithms, and the assumption was that sharpening sensors that were already quite sharp would either have negligible effect, or a small beneficial effect. Instead it was found that applying sharpening to this domain was complicated, with the effect being a function of the algorithms, the sharpening method, and the particular data used to derive the transform

(see §3.4.3). Hence the unambiguous success with the human cones did not immediately carry over to the camera sensors, and more work is required to reap any benefits that may be possible.

1.4.1 Simple Coefficient Rule Algorithms

Several coefficient rule algorithms can be classified as simply normalizing the three channels by some method. The algorithm of Brill and West [BW81] normalizes by a patch determined to be white. The grey world algorithm assumes that the average reflectivity of the scene is that of middle grey. This is the hypothesis used by Buchsbaum [Buc80] in the context of linear models.[7] Assuming that diagonal models are adequate, the coefficients of each channel are normalized by twice the average of all the pixel values of that channel. Finally, if the Retinex algorithm is applied to a scene with no illumination variation, then each channel is normalized by the maximum in each channel. All these algorithms can easily be shown to be inadequate. They all fail, for example, if the scene is completely red. It could be argued that this is a bleak situation for any color constancy algorithm to face, but these algorithms simply do not recognize when it is better to do nothing, and thus can give anomalous results.

1.4.2 The Gamut Mapping Approach

The gamut mapping approach was introduced by Forsyth [For90], and has recently being modified and extended by Finlayson [Fin95]. The gamuts

---

[7]Buchsbaum was restricted to the same number of dimensions as sensors. The inaccuracies incurred doing this are larger than assuming a diagonal model, in the case of our camera sensors. In addition, the small dimensional model is assumed to hold for two consecutive calculations, increasing the error even more.

that are mapped are the set of all sensor responses that are physically possible under a given light. The gamut mapping approach views coefficient color constancy as saying that changes in illumination correspond to diagonal mappings between the gamuts. In other words, the sensor responses possible under an unknown light are transformed to the sensor responses possible under the canonical light by an independent scaling of each component. The sensor responses observed under the unknown light restrict the mappings which are possible. Since the sensor response for a light can be defined as the sensor response of a perfect white surface as seen under that light, the diagonal mappings between gamuts will equally be mappings between lights. Thus a compact statement of the idea is to constrain the possible mappings taking the unknown illuminant to the canonical illuminant. Once this mapping is determined, we simply apply it to the unknown image to produce the sensor response for the scene under the canonical light. A few more details are given below.

First, it is important that the gamuts are convex. A single pixel sensor may sample light from more than one surface. If we assume that the response is the sum of the responses of the two contributing pieces, and that the response due to each of these is proportional to their area, then it is possible to have any convex combination of the responses. Thus the gamut of all possible sensor responses to a given light must be convex.

Since the gamuts are convex, they will be represented by their convex hulls. Now consider the RGB's in the image taken under an unknown light. These must be in the entire gamut for the unknown illuminant. Since we are modeling illumination changes by diagonal transforms, each of these measured RGB's must be mapped into the canonical gamut by the

illumination change diagonal transform. Specifically, a proposed solution must map each hull point of the measured gamut into the canonical gamut. Figure 1.1 illustrates the situation using two-dimensional triangular sets for explanatory purposes. Here triangle "abc" represents the convex hull of the measured RGB's. A proposed solution must map it into the canonical gamut represented by triangle "ABC". Reiterating the above, a proposed solution must map "a" into the canonical gamut (and similarly "b" and "c").

Now the set of maps which take a given point (e.g. "a") into some point in the canonical gamut is determined by the maps that take that point

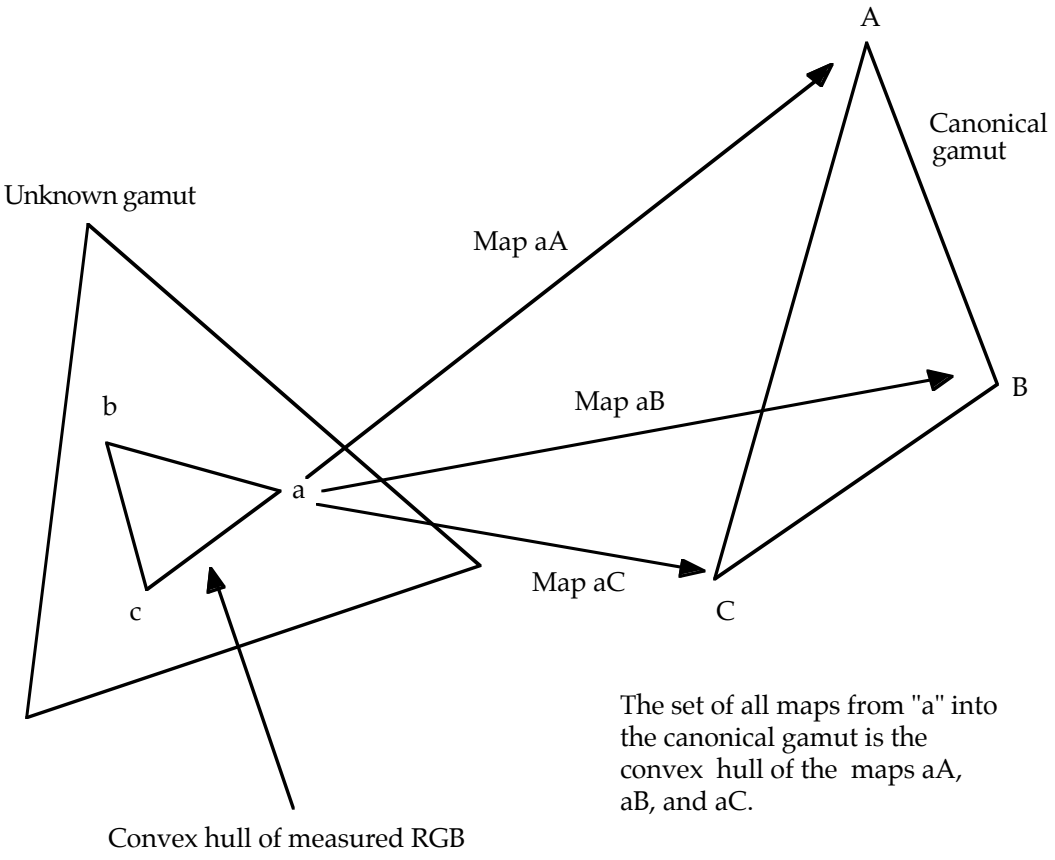

Figure 1.1 Visualization of the first part of the gamut mapping procedure.

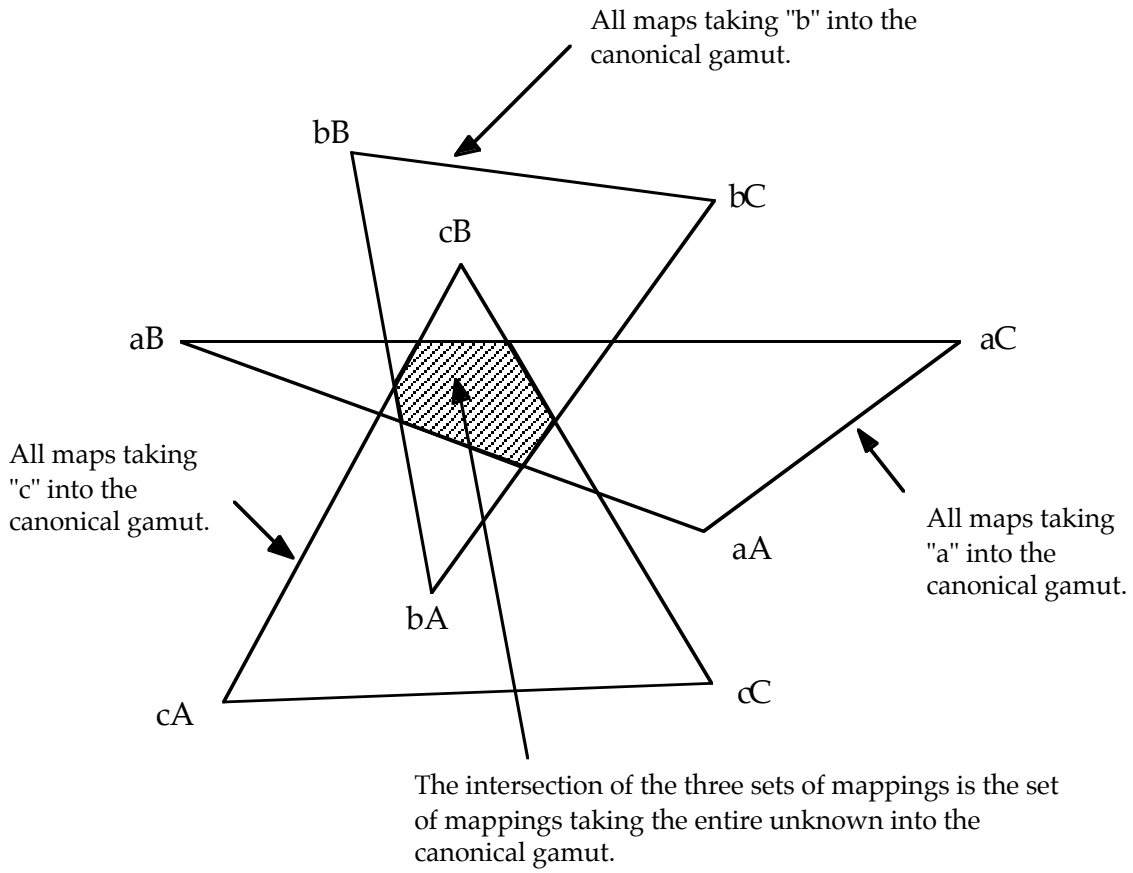

Figure 1.2 Visualization of the second part of the gamut mapping procedure.

into the hull points of the canonical gamut. If we use vectors to represent the mappings from the given point to the various canonical hull points, then we seek the convex hull of these vectors. It is critical to realize that we have introduced a level of abstraction here. We are now dealing with geometric properties of the mappings, not the gamuts. It is easy to verify that it is sufficient to consider the mappings to the hull points (as opposed to the entire set), by showing that any convex combination of the *maps* takes a given point into a similar convex combination of the canonical hull *points*.

The final piece of the logical structure is straightforward. Based on a given point ("a" in our example), we know that the mapping we seek is in a specific convex set. The other points lead to similar constraints. Thus we intersect the sets to obtain a final constraint set for the mappings. Figure 1.2 illustrates the process.

Recently Finlayson proposed using the gamut mapping approach in chromaticity space, reducing the dimensional complexity of the problem from three to two in the case of trichromats [Fin95]. Not all chromaticity spaces will work. However, Finlayson showed that if the chromaticity space was obtained by dividing each of two sensor responses by a third, as in the case of (red/blue, green/blue), then convexity is maintained where required. One advantage to working in a chromaticity space is that the algorithm is immediately robust with respect to illumination intensity variation. Such variation is present in almost every image, as it originates from the ubiquitous effects of shading and nearby, extended light sources. Furthermore, highlights due to specular reflection do not present trouble in chromaticity space, because here they behave as very white surfaces (this assumes that they do not saturate the camera sensors).

In addition to using chromaticity space, Finlayson added an important new constraint. Not all theoretically possible lights are commonly encountered. From this observation, Finlayson introduced a constraint on the illumination. The convex hull of the chromaticities of the expected lights makes up an illumination gamut. Unfortunately, the corresponding set of allowable mappings from the unknown gamut to the canonical gamut is not convex (it is obtained from taking the component-wise reciprocals of

the points in the above convex set). Nonetheless, Finlayson was able to apply the constraints in the two dimensional case. In the work that follows the convex hull of the non-convex set was simply taken, as it was found to be a satisfactory approximation both for the two-dimensional and three-dimensional case.

1.5 Other Algorithms

A few approaches to color constancy only peripherally related to this work should be at least mentioned to provide some balance. Foremost, at least historically, is the Maloney-Wandell algorithm [MW86, Wan87]. This approach is based on the small dimensional linear models defined in §1.2. They found that one could solve for the components of the light vector provided that the surface dimensionality is less than the number of the sensors. In the case of three sensors, this forces us to assume that the reflectances are two dimensional. Assuming two dimensional reflectances, the sensor responses under a fixed, unknown light will fall in a plane. The orientation of this plane indicates the illumination. Despite its significance, the Maloney-Wandell algorithm does not work very well [FF94, FFB95]. The first reason is simple: the dimensionality of surfaces is greater than two. The algorithm is also not very robust with insufficient data. For example, if there is essentially one color in the scene, then the sought after plane is entirely free in one of its two degrees of freedom (the plane is anchored at the origin).

Gershon et. al. [GJT86] were able to use an additional dimension for surfaces by incorporating an average scene reflectance assumption similar to the grey world assumption. Here the average reflectance of the scene is

determined by averaging the RGB's of distinct regions. This overcomes the weakness of the grey world assumption when applied to a scene predominantly one (non-grey) color, provided that other colors are available in smaller proportions. The expected scene reflectivity is determined by averaging published reflectance data. This algorithm relies on several optimistic assumptions. First, in the usual trichromatic case, it is assumed that surfaces can be described by three basis functions. Second, the average reflectance of the scene is that given by averaging the published data set. Although this is a step beyond the simple grey world hypothesis, no mention is made of any attempts to verify this assumption.

Another interesting approach involves modeling the specular and diffuse parts of scene reflectance [Sha85]. For most practical situations, this cannot be used as one's sole method of color constancy processing. However, this idea will likely be very useful for a complete color constancy system. For now it has to be left as a tantalizing possibility for future work.

Finally, in the course of this research, I studied the probabilistic approaches initiated by Vhrel and Trussel [VT91], and researched extensively by Brainard and Freeman [BF94, FB95]. The thrust of the work is estimating the reflectance and illumination factors from their observed product, in the face of measurement error, with access to a priori estimates of reflectance and illumination likelihoods. The problem of obtaining these likelihoods is not addressed, and therefore a large part of the color constancy problem is ignored. They provide the results of simulations where they compare their algorithm to others, but the data is generated with a method favorable to their algorithm. Their model assumes that the components of a small

dimensional model for surfaces and illumination will follow a Gaussian distribution, and they test the algorithm against such input. Although it is common to generate data for simulations, it is a problem here because there is no additional justification for the key assumption. Nonetheless, it would be interesting to see this algorithm tested against real data.

1.6 Varying Illumination Algorithms

The preceding algorithms assume that the illumination, or at least its chromaticity, is constant over the scene. Only a few algorithms try to deal with the reality that this is often not the case. An important part of this thesis deals with the varying illumination problem, and the relevant algorithms are introduced here to provide some context.

1.6.1 The Retinex Algorithm

In theory, Land's Retinex algorithm [LM71, MMT75, Lan77, Lan86] can deal with varying illumination. The Retinex algorithm emerged from work on the human visual system, and treating it as a computational color constancy algorithm takes some care. On the basis of psychophysical experiments showing that perceived lightness is influenced by edges, Land and his colleagues proposed a simple scheme in which illumination variation is removed from lightness computation. Thus the motivation was not to extend an algorithm to deal with an additional challenge, but to model human vision. This may explain the anomaly that despite the implicit promise of dealing with varying illumination, the test of Retinex which is most comparable to tests of computational color constancy

algorithms was done on scenes where the illumination variation was carefully minimized.[8]

In the Retinex algorithm, the basic idea is that changes in illumination can be distinguished from changes in surface reflectance by the assumption that reflectance changes are spatially abrupt, whereas illumination changes will occur gradually. Thus reflectance changes can be determined simply by identifying jumps larger than some threshold value. In Retinex the reflectivity of a given location is determined relative to a bright spot by taking a random path from the location in question. This path must not intersect itself. With luck (or with a complex enough path), a patch for each channel which is close to the maximum brightness possible for that channel will be crossed. The results for a number of these paths are averaged to reduce the error.

An algorithm based on these random walks is arguably inelegant. For one, the results are irreproducible. Horn [Hor74] realized that the essence of the matter was differentiation (to identify the jumps), followed by thresholding (to separate reflectance from illumination), followed by integration (to recover lightness). It should be noted that this method uses the logarithm of the pixel values. This can also be implemented by using . The difficulty with this method is that the two-dimensional integral does not necessarily exist. Funt et. al. [FDB92] developed a method to insure the

---

[8]This is best explained in [MMT75]. The same experiment is also referred to in [Lan77], but some details relevant to this work are omitted. The scenes were illuminated indirectly through the walls of an integrating cube. Using this method, the illumination variation would be far less than that in the real scenes used in this thesis.

existence of the integral in a slightly different context. The problem can also be approach using homomorphic filtering. [GW87] Another approach to the problem will be introduced as part of this work.

1.6.2 Gamut Mapping for Varying Illumination

Recently Finlayson et. al. [FFB95] introduced a new approach to the problem of varying illumination. Rather than accepting illumination variation as a hindrance to be removed, it was found that it could provide a powerful constraint. Certainly the information provided by seeing the same surface under different illumination has been studied, but primarily only in the context of seeing the entire scene under two different illuminants at different times; the illumination for each such view is assumed uniform (see [TO90, ZI93, Fin94]). One method available for using varying illumination without multiple views is in the case of chromaticity shifts at a shadow boundary [FF94]. This method requires shadows in the scene, and more critically, some external method to identify them.

In [FFB95] it is observed that in chromaticity space the mappings from unknown illumination to a canonical illuminant fall nearly on a straight line. This corresponds to the illuminant chromaticities lying on a curve, since they are inverted to produce the mappings. This is congruent with the observation the chromaticities of the 10 illuminants used lie approximately on the Planckian locus. This set of mappings produces a set of chromaticities defining possibilities for the sensor responses of the surface under the canonical illuminant. Since the computation is simply a scaling, the mapped set is also a line. A second view produces a second line. Provided the illuminations are indeed different in chromaticity, and the system's sensors

are adequate, these lines will intersect, providing a single estimate of the surface chromaticity under the canonical light.

As promising as this method is, there is still much to be done. First, the problem of identifying the appropriate surface is not dealt with. Second, since we are concerned with scenes where the illumination varies substantially, it is not safe to assume that the solution for one appropriate surface (a wall, for example), propagates without modification to other parts of the scene (a bookshelf, for example). Also it is quite optimistic to assume that every part of a scene will be close enough to an appropriate surface. On a similar track, if there is little variation in the illumination chromaticity, then the algorithm will fail. Third, the lines obtained are justified on the assumption that the chromaticities of the lights lie roughly on a particular curve. This leaves open the question of how to handle a wider illumination set. In short, we have a kernel of a solution that must be integrated into a more comprehensive system.

Chapter four of this thesis deals with this integration. The algorithm is modified to work together with the gamut mapping algorithms based on surface and illumination constraints described in §1.4.2. In this extension, the entire illumination gamut is used, overcoming the commitment to one approximation for the expected illumination. Then the problem of identifying the illumination variation is solved in the case of easy to segment images. The results obtained are very encouraging.

1.7 Color Constancy on Image Data

The eventual goal of computational color constancy is effective and robust color constancy processing on image data. Experiments on generated data are necessary to evaluate ideas, but the eventual goal implies that testing should be done on real data as soon as possible. If this is not done, algorithms suitable only for generated data may be favored. Furthermore, it seems natural to point the camera at some part of the world and test one's algorithms on that input, and it is a little surprising how few results of this sort are available. This may be due, in part, to the brittleness of computer vision algorithms when run on arbitrary data.

Although they do not use a camera, some of the work of Land and his associates qualifies as being tested on "real" data[9]. There is human interaction, but all the steps that are not automated could easily be automated. The system must deal with noise and other varieties of bad data, as well as non-conformance to theoretical models. It should be noted, however, that all the work is with well behaved images with lots of color (very carefully illuminated patchworks of rectangular pieces of matte paper dubbed "Mondrians").

David Forsyth tested his algorithm on a number of color Mondrians [For90]. He also writes about the absence of tests on real data, coming up with a few instances of single tests (page 16). Since the algorithms he

---

[9]Since their research focus is on human vision, they estimate cone responses for the image locations and use this for input to a computer program [MMT75].

implemented will work using each pixel as a surface patch, I assume this was done. He does not, however, provide this information.

Finally, Tominaga reports results from image data from a camera with six sensors implemented by using a monochrome CCD in conjunction with six narrow band filters [Tom94]. In addition, the dynamic range was extended by taking pictures at various shutter speeds. A combination of the Maloney-Wandell approach with the dichromatic modeling of Shafer was implemented. The results reported were for cylinders covered with colored paper as well as plastic ones.

If we accept the philosophy put forth at the beginning of this section, then the current lack of carefully controlled results on image data suggests that such results should be desirable. In this thesis, results on image data are provided for four scenes (included one which can be described as arbitrary) viewed under three different illuminants. The algorithms tested under the same conditions include two simple coefficient rules and eight variations of gamut mapping algorithms. Results are also provided for image data under varying illumination, but here the results are presented graphically and visually, since algorithms that could compete under the same conditions have not been implemented.[10] In conclusion, the results presented in this thesis are a healthy contribution to the embarrassingly small set of careful tests on image data.

---

[10]The only candidate would be the Retinex algorithm.

Chapter Two

Preliminary Measurements

In order to address the issues of color constancy on real images, the nature of the input to the RGB values as defined by (1.1) must be investigated. To this end, real world lighting and surface reflectances, as well as the camera sensor functions, were measured.

2.1 Camera Calibration

In theory it is possible to implement all the algorithms tested as part of this work without knowing the camera response functions (assuming that they are modeled by (1.1)), but in practice it is exceedingly convenient if they are known. Once the model has been verified, and the sensor functions determined, it is possible to generate the canonical gamuts and test data with far less effort than doing so directly. This is due in part to the fact that it is easier to obtain high quality measurements with a modern spectraradiometer than with the camera. In a sense, camera calibration is a process where the

camera is used very carefully once, and then the results are used to predict the responses.

If we let $C(\lambda) = S(\lambda)E(\lambda)$, and restrict our consideration to only one channel, then (1.1) becomes:

$$\rho = \int_{\lambda} R(\lambda)C(\lambda)d\lambda \tag{2.1}$$

The goal is to determine $R(\lambda)$ from a number of ρ and $C(\lambda)$ pairs. If we could produce nicely spaced color signals which were very sharp, then the response function would be sampled exactly at the peak locations, and a smooth curve could be fit through the result. This approach is not used because it is difficult to produce such color signals with enough energy, and such that the signal is uniform over a sufficient number of pixels.[1] Instead a method inspired by Sharma and Trussel's approach [ST94] is used.

To begin, we represent the continuous functions by vectors whose components are the sampling of the functions at equally spaced intervals. Since the spectraradiometer measures from 380nm to 780nm in steps of 4nm, these vectors will have 101 components, with the first component being the value at 380nm, the second component being the value at 384nm, and so on. Using **R** for the reflectance vector, and **C** for the color signal vector, equation (2.1) becomes:

$$\rho = \mathbf{R} \bullet \mathbf{C} \tag{2.2}$$

---

[1]If this method is to be used, then the best approach is to use a set of interference filters. This was used by Tominaga [Tom94] with reasonable success. Nonetheless, a better use of such filters would be to use them in conjunction with the method explained shortly.

The strategy in camera calibration is to probe the camera with a number of different **C**, to get a number of different responses, and to use this to estimate **R**. Thus we have a number of equations:

$$\rho^{(k)} = \mathbf{R} \bullet \mathbf{C}^{(k)} \tag{2.3}$$

Due to the large dimensionality of the vectors (101), and the small dimensionality of the signals, the system of equations in (2.3) is severely under-constrained. It is part of the challenge of calibrating the camera to produce unnatural illuminants to increase the dimensionality of the space of the $\mathbf{C}^{(k)}$. Even so, the set of equations is expected to be under-constrained. Sharma and Trussel introduce additional constraints. First they insist that the response functions are non-negative. Second, they introduce smoothness constraints. These take the form of bounds on a discrete estimation of the second derivative:

$$\left|2*\mathbf{R}_i - \mathbf{R}_{i-1} - \mathbf{R}_{i+1}\right| \le \delta \quad \text{(Except for first and last } \mathbf{R}_i\text{)} \tag{2.4}$$

Next they constrain the maximum error:

$$\left|\rho^{(k)} - \mathbf{R} \bullet \mathbf{C}^{(k)}\right| \le \zeta \tag{2.5}$$

Finally, they constrain the sum of squares error:

$$\sum_k \left(\rho^{(k)} - \mathbf{R} \bullet \mathbf{C}^{(k)}\right)^2 \le \varepsilon \tag{2.6}$$

Sharma and Trussel then observe that the constraint sets are all convex, and propose that the method of projection onto convex sets (POCS) be used to characterize the result. They use their method to obtain a good estimate for the sensitivity of a color scanner.

However, we can do a little better with less effort, given that implementing POCS would be an involved process, as an external program for this method could not be found. First, it is convenient to rewrite (2.5) as:

$$\begin{aligned} \rho^{(k)} - \mathbf{R} \bullet \mathbf{C}^{(k)} &\le \zeta \\ \mathbf{R} \bullet \mathbf{C}^{(k)} - \rho^{(k)} &\le \zeta \end{aligned} \tag{2.7}$$

Similarly write (2.6) as:

$$\begin{aligned} 2*\mathbf{R}_i - \mathbf{R}_{i-1} - \mathbf{R}_{i+1} &\le \delta_1 \quad \text{(Except for first and last } \mathbf{R}_i\text{)} \\ \mathbf{R}_{i-1} - \mathbf{R}_{i+1} - 2*\mathbf{R}_i &\le \delta_2 \quad \text{(Except for first and last } \mathbf{R}_i\text{)} \end{aligned} \tag{2.8}$$

Here I have introduced separate constraints on the lower and upper limits on the second derivative. Since we expect the sensors to be positive, uni-peaked functions, the absolute value of an acceptable upper limit is more than that for the lower limit. Next it would seem preferable to minimize the left hand side of (2.5), rather than constrain the error interval. In fact, for our purposes, it is better to minimize the sum of squares relative error:

$$\sum_k \left(\frac{\rho^{(k)} - \mathbf{R} \bullet \mathbf{C}^{(k)}}{\rho^{(k)}} \right)^2 \tag{2.9}$$

This amounts to finding the best least squares solution to $\mathbf{MR}=\mathbf{1}$, where the rows of $\mathbf{M}$ are the vectors $\mathbf{C}^{(k)}$ scaled by $\rho^{(k)}$, subject to the above constraints. At this stage, we have a least squares fit problem with linear constraints. For this problem, implementations of standard numerical solution methods are readily available.[2] If it is preferred, implementing the minimization of absolute error is even easier, and minimizing a weighted sum of both is also simple.

---

[2]I used the DBOCLS routine in the SLATEC math library available by anonymous FTP from netlib2.cs.utk.edu.

The experimental set up consisted of a Sony DXC-930 CCD camera with a Canon zoom lens, a Photoresearch 650 spectraradiometer, and a number of lights, surfaces, and filters used to craft a set of **C** with as high a dimension as possible. The camera settings were chosen to be as neutral as possible. Most importantly, the gamma adjustment was turned off. It is necessary to have the camera sensor gains set either for daylight (color balanced to 5400K) or indoor light (color balanced to 3200K), and the choice was made to use the latter (3200K). It was discovered that the camera-digitizer system has an offset of roughly 13 RGB units. In other words, if there is absolutely no light reaching the camera, it records RGB vectors with mean (11.1, 13.2, 12.9) and standard deviations of the order of 2.5. Values well outside the standard error do occur, but it was found that the mean and standard errors are consistent over a period of months. It is critical to subtract such an offset from camera RGB's for practically all color constancy processing (whenever equation (1.1) is assumed to hold). The standard deviation of the offset is taken to be indicative of the sensor error due to noise.

Preliminary measurements verified that the camera is linear within 5%. These were done by increasing the intensity of light reaching the camera by moving a bright light closer to a standard white reflectance seen both by the camera and the spectraradiometer (at essentially the same angle). The camera response was found to increase linearly with the incident light energy, as measured by the spectraradiometer. Preliminary measurements also verified that there are small, but not entirely negligible effects on the chromaticity recorded by the camera under extreme changes in the optics. The magnitude of the camera sensor functions will, of course, change with the aperture, but

there are also slight changes in the chromaticity with both aperture changes and focal length changes. These effects differ across the viewing field, and are mostly confined to the outer half of the field of view. The results provided are for an aperture setting of 2.8, a focal length of 25, and the central 10% of the field of view. The aperture control is too coarse for good reproducibility. Thus the camera sensor curves will predict camera responses for an aperture that is only estimated by 2.8. For our research, this slight uncertainty is not a problem. It *would* be a problem if the aperture changed during the calibration. For this reason, the aperture ring was taped firmly in place during the measurements.

The procedure was to measure alternately the light coming from some combination of lights, surfaces, and filters, with both the camera and the spectraradiometer. In order that the geometry was kept constant, the spectraradiometer was mounted on top of the camera, which itself was on a tripod. Switching between the two sampling devices was achieved by raising and lowering the tripod head. Based on several dry runs, and the examination of many spectra, a collection of light-surface-filter combinations was chosen which provided close to the most variation possible with the equipment at hand. The lights include an incandescent light, a Philips CW fluorescent light, a slide projector, and a black light (a strong source of ultraviolet light close to the visible spectrum). The surfaces consisted of the Macbeth color checker patches and 19 paint chips. The filters were Kodak gelatin filters 29 (red), 58 (green), 47A (light-blue), and 47B (blue). Again due to previous experience, not all possible light-surface-filter combinations were

used, as there is much redundancy in them. In the final run 58 combinations were used.

In order to determine the final camera response, a number of pixels in the area of interest were averaged. The size of the area used is a compromise. The larger the area, the more illumination variation and optical problems there will be. In addition, the pixels will correspond, on average, to points further from the small sampling region of the spectraradiometer. On the other hand, as the area is decreased, more noise and other error is introduced, and it is possible that the sampling region of the spectraradiometer could be missed entirely. For this work, a region which was roughly five times the size of the sampling area of the spectraradiometer was used. This produced image sections of a few hundred pixels.

Computing the sensor functions by the method above is somewhat of an art. We do not know what the functions are, but on the other hand, we assume that we know roughly what they look like. Basically we assume that they start at zero, rise smoothly to a peak, fall smoothly back to zero, and stay there. By imposing these constraints, we hope to better model the actual sensors, and thus gain power in predicting the response to spectra quite different from the test set. If we were only interested in spectra close to the test set, a straight minimization of the error would suffice (the curves produced doing this are very jagged). By adjusting the balance between smoothness and minimum error, a set of smooth curves which predict the camera response over the test set to within 3% RMS relative error were found. The RMS absolute error is 3.5 pixel values. Since the minimum error possible with the non-negativity constraints is about 2%, this is a good fit. The sensor curves are shown in Figure 2.1.

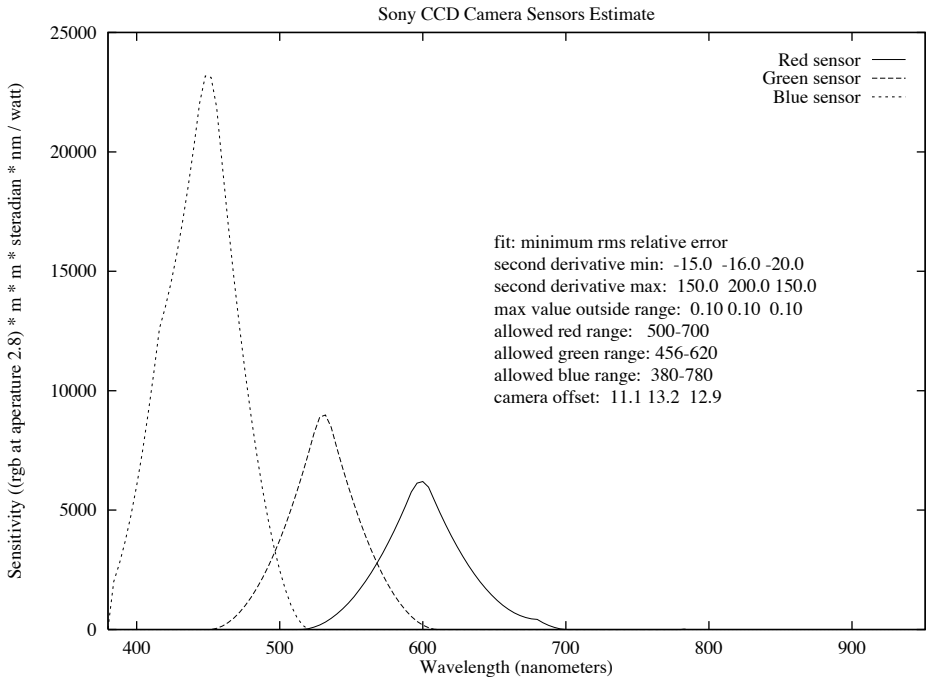

Figure 2.1 Camera response functions as determined by the methods described in this section. The functions are only valid for an aperture setting of 2.8 and camera settings of 3200K, no gain adjustment, and no gamma correction.

2.2 Measurement of Illuminant Spectra

As described in §1.5.2, restricting the set of expected illuminations is a powerful constraint. This leads to the problem of what constitutes an appropriate restriction on the illumination. Finlayson [Fin95] used the published data for the 6 phases of daylight [JMG64] (D48, D55, D65, D75, D100, and D220), the standard CIE illuminants A, B, and C, a 2000K Planckian black body radiator, and uniform white. Thus it was established that a reasonably large set of illuminants is still small enough to be a good constraint.

However, to have confidence that the algorithm will work on real images, one needs to know if the illumination set includes all lights deemed "typical". In addition, it may be possible to make the constraint set smaller. If this was the case, knowing real world lighting would improve the performance of some of the algorithms.[3]

It is possible through the use of filters, or by bouncing light off deeply colored objects, to construct a set of "lights" which is so large as to be useless as a constraint. But the set of lights must include all lights expected in the application domain; otherwise, the constraint will work artificially well when tested on included lights, and may fail when tested on the excluded ones.

Despite the lack of a good definition of "typical" illumination, I set out to measure it. The lighting was measured at various places in and around the SFU campus, at various times of the day, and in a variety of weather conditions. Unusual lighting, such as that beside neon advertising lights, was excluded. However, care was taken to include some reflected light. It seems fair to include lighting which has some component reflected from a concrete building, but not if the building was painted pink. Similarly, the light underneath trees was included. Altogether, roughly 100 different spectra were measured. The chromaticities of the measured spectra are shown in Figure 2.2.

---

[3]In some sense this turned out to be the case. Although including a wide range of illuminants expanded Finlayson's gamut in some directions, no light as red as 2000K was encountered, and hence the measured gamut was more restrictive in the red.

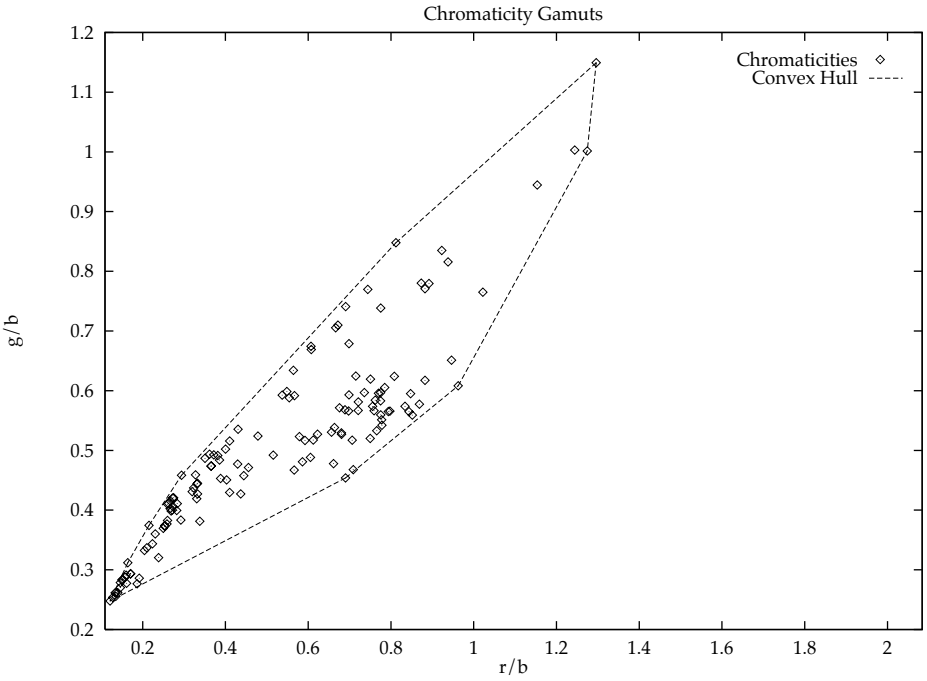

Figure 2.2 Chromaticity gamut of all measured illumination.

2.2 Measurement of Surface Reflectances

In spite of a wealth of published surface reflectance data, surface reflectances were also measured for several reasons. First, one important set of data, the Krinov data set [Kri47], only includes reflectances for wavelengths from 400nm to 650nm. The other data sets are restricted to the range of 400nm to 700nm. Figure 2.1 shows that the camera sensors respond to wavelengths outside these ranges. Second, in the case of the Macbeth color checker which was used for many experiments, it makes sense to use the reflectances of our copy, rather than assume that they are as published. By far the most important impetus for measuring reflectances was that the color of some

objects in our lab did not fall inside the gamut of the published data. An underlying assumption of the gamut mapping algorithms is that the canonical gamuts include responses for all surfaces. If this assumption does not hold, then the algorithms can perform poorly (as was the case). In summary, spectra were measured for essentially the same reason that illumination was measured—to ensure that the properties of the real world were accounted for.

To measure reflectances, the spectra of light reflected from a surface was divided by the spectra of light reflected from a barium oxide standard white surface purchased from Photoresearch. The incident light angle was 45˚, and measurements were taken at an angle of 90˚. Since it was found that it was virtually impossible to illuminate even a relatively small area evenly, it was critical to make sure that the standard reflectance and the test reflectance were in the same place.

The goal was not to produce a complete set of measurements, but simply to validate the use of the published data, and to extend it where necessary. Thus most surfaces were chosen based on how different they were from ones already measured. In addition, a few surfaces that were suspected of causing problems for our color constancy algorithms were measured. The surfaces measured included the Macbeth color checker patches, some paint chips, the covers of some books used in our test scenes, some brightly colored shirts, and a number of pieces of construction paper. In total, 78 surfaces were measured. The chromaticity gamuts of these sets of surfaces as viewed under a Philips CW fluorescent bulb are shown in Figure 2.3.

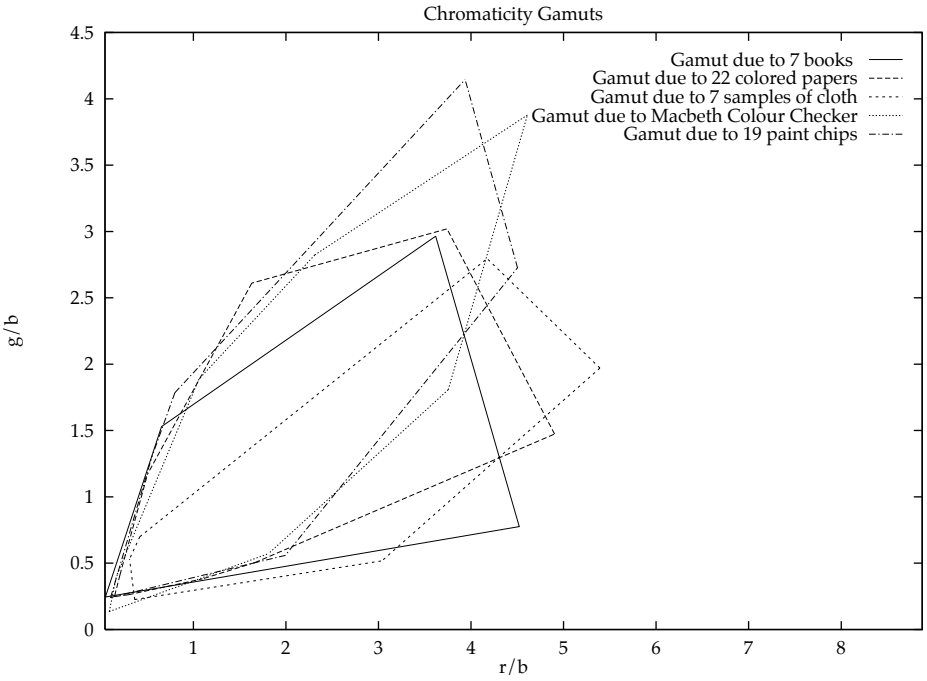

Figure 2.3 Chromaticity gamuts of sets of measured surfaces under a Philips CW fluorescent light as viewed by the Sony CCD camera.

In Figure 2.4 the measured gamuts are combined into one gamut, which is compared to the published data sets. The interesting point is that the measured gamut does extend outside what is available in the published data (in the lower right). Furthermore, extension of the gamut in this direction cannot be explained by the lack of data for wavelengths less than 400nm. A second point of interest is the great extent to which the Munsell chip gamut exceeds the measured gamut in the blue direction. This anomaly is primarily due to a small number of points. In other words, the great majority of the 462

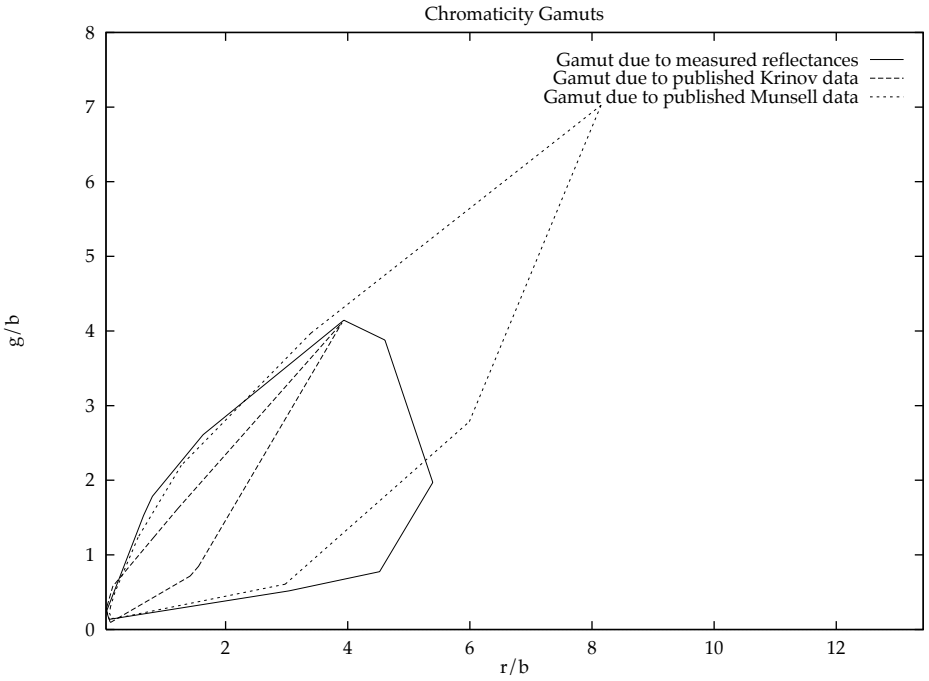

Figure 2.4 Chromaticity gamuts of the measured data together with the gamuts for the published Krinov and Munsell data sets.

surfaces are inside the measured gamut, and a few are far outside it.[4] This can be seen in Figure 2.5 which shows the distribution of the chromaticities in the Munsell chip gamut. Since these surfaces have a very small amount of blue, this could be explained by the cutoff in the measurement range, but the spectral characteristics of the surface would be odd. Other explanations that should be ruled out include a problem with the Munsell measurements, or other standard sources of error. Still, the likeliest explanation is that these

[4]The Munsell chip data used is the set of 462 spectra measured by Nickerson [Nic57]. Spectra 131 and 428 were far outside the measured gamut. Spectra 91, 129, 130, 458 were a little outside, and 6 other spectra were slightly outside.

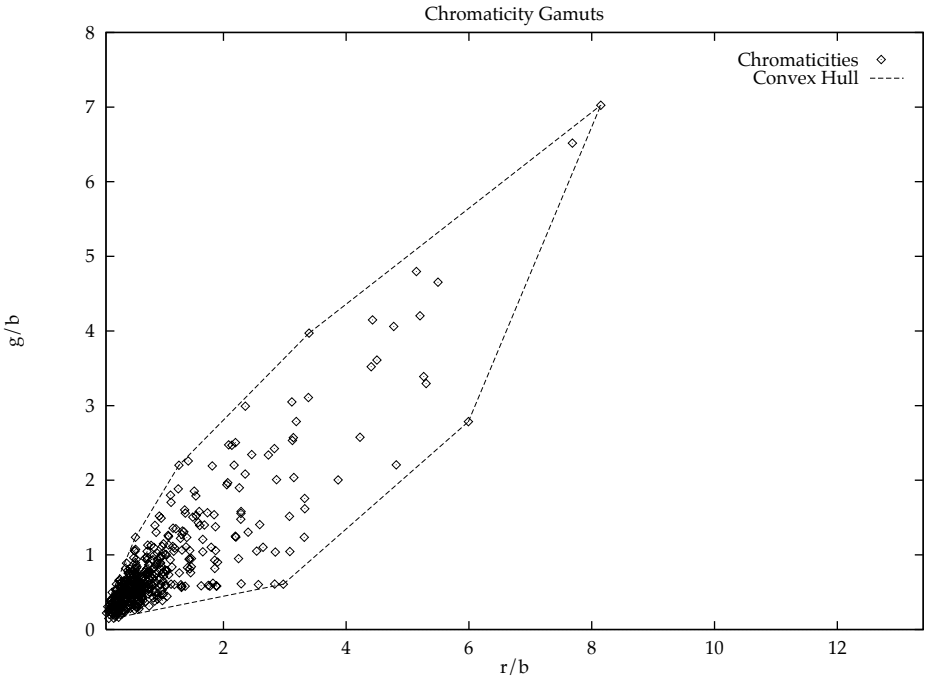

Figure 2.5 Chromaticity gamut of the Munsell data set showing the distribution of chromaticities.

colors do exist in paint chips. Although it technically goes against the philosophy of my approach, the rarity of the offending colors both in the data set, and in the lab, justifies using the measured gamut for convenience. On the other hand, the large extra piece of the gamut due to the Munsell chips does indicate that additional effort spent building (and checking) the reflectance databases is warranted.

2.2.1 Previously Used Canonical Gamuts

Having established the source of the gamuts used in this research, the gamuts used in previous work should be mentioned. In contrast to our approach of generating the gamuts from measured spectra and calibrated

camera sensors, Forsyth determined his gamut by taking pictures under a real, physical, canonical light [For90]. For surfaces he used a collection of 180 colored pieces of paper. Finlayson, on the other hand, generated his gamut from the reflectances of the Macbeth color checker, and then expanded the gamut by 5% to account for the possibility out of gamut chromaticities [Fin95].

Chapter Three

Gamut Mapping Color Constancy

The focus of this part of the research is the implementation of a number of color constancy algorithms in the same framework, in order that meaningful comparisons and observations can be made. Essentially all published algorithms are correct in the sense that if the author's model holds, the expected results will be forthcoming. The problem in going from theoretical to practical situations is that the assumptions, many of which are implicit, often do not hold. But if we compare algorithms (or dare to rank them), then we are saying that we can define a fair input set. Nonetheless, an important first step is to test the algorithms in the same context, and attempt to evaluate the relevant features of the results as they pertain to our necessarily biased input.

Of specific interest in this chapter is the relationship between the three-dimensional approach and the two-dimensional approach, the methods for

choosing a solution from the constraint set, and the usefulness of sensor sharpening.

3.1 Overview

The general idea of the gamut mapping approach is to constrain the set of solutions. We can apply constraints determined from the sensor responses, henceforth referred to as surface or S constraints, and intersect them for a solution set. In three dimensions this is Forsyth's method [For90]. We can additionally intersect this solution set with the illumination constraint, as done by Finlayson in two dimensions. Figure 3.1 illustrates the utilization of multiple constraints in the two-dimensional case. Some specific solutions are also plotted.

3.2 Implementation Details

3.2.1 Primitives

In order to discuss the details of the algorithms, it is helpful to be able to refer to a few primitive operations. The foremost of these is finding the convex hull of a set of points. This was achieved by using the program *qhull* modified so that the hull computation was callable from C programs.[1] Fortuitously, *qhull* provides the facet normals as well as the hull points. Second, it was necessary to compute hull intersections in 2 and 3 dimensions, and for this, a simple approximation was used. First the bounds in all

[1]Qhull is available by anonymous FTP from geom.umn.edu

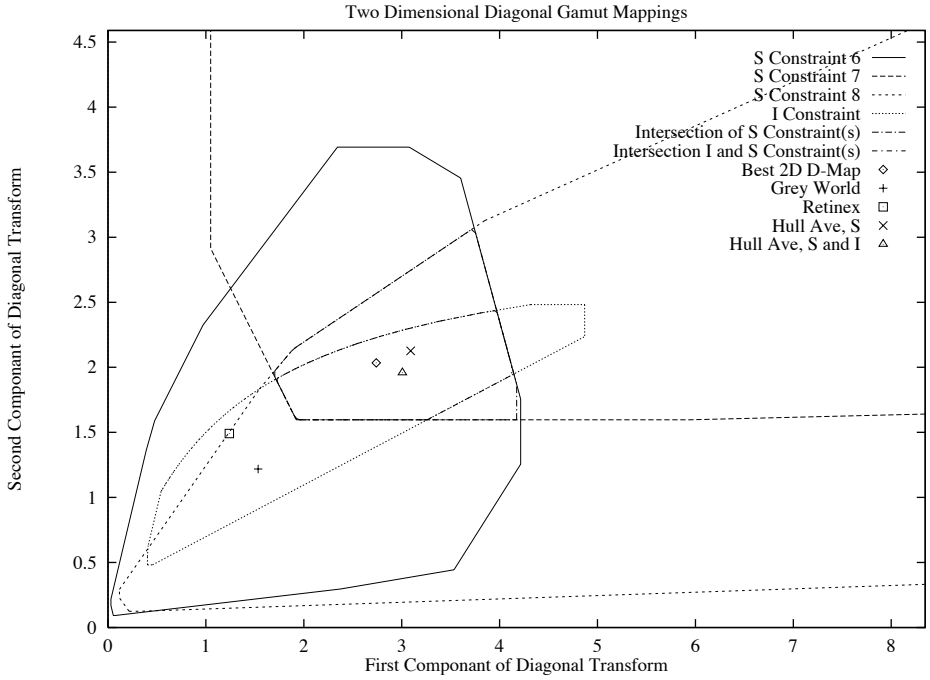

Figure 3.1 Illustration of multiple constraints on the mapping from the unknown illuminant to the canonical.

coordinates of the intersection of the hulls were determined. Then lines inside these bounds, parallel to a specific axis, and passing through a grid of selected resolution, were intersected with all the hulls. From these intersections, the intersection of the line with the intersection of all hulls was determined. Finally the convex hull of these points was computed using *qhull*.

Choosing a solution from the set of possibilities also required implementing a few primitives. One choice is the average of all points inside

the set which is the centroid. This is not simply the average of the hull points. The centroid is defined by:

$$\frac{\int_{\substack{\text{Solution}\\\text{Set}}} \bar{X}\,dV}{\int_{\substack{\text{Solution}\\\text{Stl}}} dV} \tag{3.1}$$

This was calculated by numerical integration over the volume (or area) of the hull. A second choice is the point with the maximum product of its coordinates. This calculation takes some care because the overall maximum does not necessarily lie on a boundary point, nor necessarily (in the three-dimensional case) on a boundary edge. Thus local maxima in the interiors of edges and facets need to be found for comparison with the values at the boundary points.

3.2.2 Gamut Mapping Implementation

The general goal is to constrain the mapping from the unknown illuminant to the canonical illuminant. The first source of constraint is the surfaces. To compute the constraint we first take the convex hull of the input (r,g,b) or (r/b, g/b). Element-wise division of each of the canonical gamut hull points with these points produces the constraint sets. (With respect to the example in §1.4.2, measured hull point "a" is used to compute sets aA, aB, and aC, and similarly for points "b" and "c".) These sets are intersected to give the constraint due to the surfaces.

Calculating the illumination constraint is more involved. In the same coordinates as the surface constraints, the illumination constraint is not convex. Nonetheless, approximating it by its convex hull is reasonable. First,

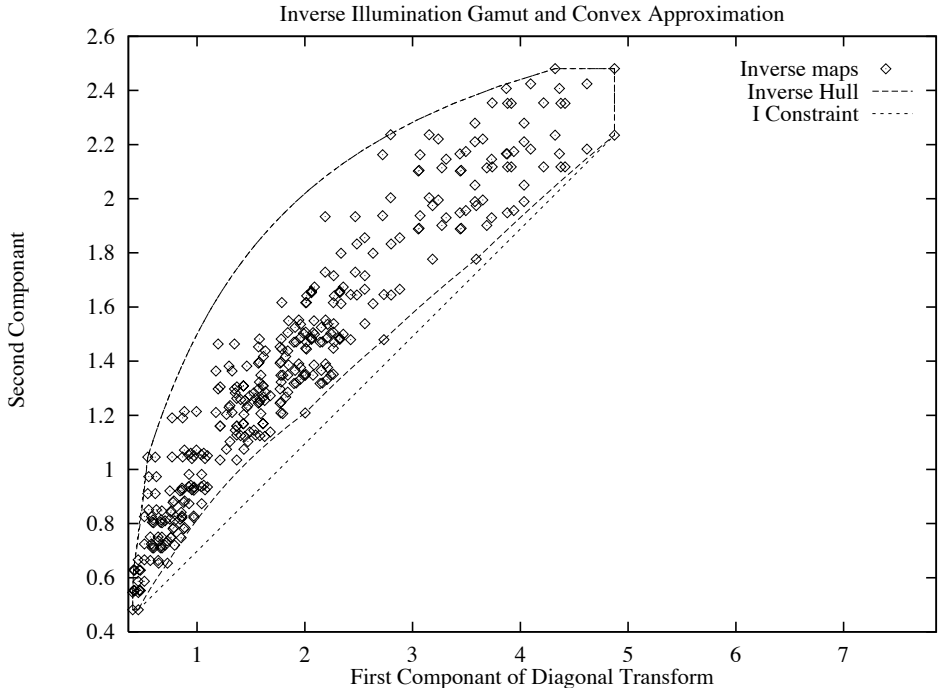

Figure 3.2 The illumination chromaticity gamut showing the data points used to estimate the inverse hull (but not the convex hull directly, as explained in the text). The true constraint is labeled as "Inverse Hull", and the convex approximation as "I constraint". The two differ only by a thin sliver on the bottom edge.

doing this does not exclude any solutions, and second, it is a good approximation. This is clear from Figure 3.2. The illumination constraint is a set of element-wise reciprocals of the convex set representing mappings taking illuminants into the canonical illuminant. This convex set is obtained by element-wise division of the convex hull of the illumination gamut by the canonical illuminant. In order to honor the claim that solutions are not excluded, we cannot simply take the convex hull of the inverses of these mappings. Instead, we compute the convex hull of the mappings, and then invert the result.

Applying the convex approximation in the three-dimensional case takes only a little more effort. It is important to realize that in three dimensions, unlike the surfaces constraints, the illumination constraint is not a polyhedron, but a cone anchored at the origin. This is because it is difficult to characterize the maximum brightness of the expected illuminants; for the most part we assume that they can be arbitrarily bright. Thus the illumination constraint is really a constraint on chromaticity. In order to calculate the appropriate cone, the three-dimensional inverse hull is projected onto the plane x+y+z=1. Then the convex hull of this projection is taken, and the cone defined by the origin and extending through the hull points to infinity is used.

This should give the reader a clear idea of how to compute the constraints on the mappings both with and without the illumination constraint, and in two and three dimensions. I now turn to the problem of selecting a solution from the constraint set. It is unfortunate that a solution has to be chosen at all. It is a great strength of the method that the possible range of the solution is computed. Nonetheless, in order to report results, or demonstrate the algorithm for an audience, a specific solution is necessary.

Originally, a heuristic requiring the volume of the mapped set to be maximal was used [For90]. This translates into maximizing the product of the coordinates of the solution. The heuristic is justified as providing the solution of maximal surface lightness. In the two-dimensional algorithm [Fin95], this precedent was adhered to, but in the two-dimensional case this heuristic loses appeal, as it favors certain chromaticities over others. In deference to the literature, the maximum volume constraint was

implemented and tested. However, if the solutions are evenly distributed in the hull, and the magnitude of the vector distance is used as an error measure, then statistically the best answer will be the hull centroid (see Appendix A). Thus this estimate for the answer was also implemented and tested.

3.2.3 Simple Color Constancy Algorithms

In addition to gamut mapping algorithms, two simple algorithms were implemented such that they conformed to the output model. The first such algorithm is the grey world algorithm. Denote the average RGB for the scene as $(r_A,\ g_A,\ b_A)$, and denote the camera sensor response to grey (a uniform 50% reflectance) under the canonical as $(r_G,\ g_G,\ b_G)$. Then the diagonal transform from unknown to canonical is:

$$\left(\frac{r_G}{r_A},\ \frac{g_G}{g_A},\ \frac{b_G}{b_A}\right) \tag{3.2}$$

The Retinex algorithm in the case of non-varying illumination was also implemented. Here we calculate the maximum in each channel, $(r_M,\ g_M,\ b_M)$ over the entire scene. This is mapped into the camera sensor response to white under the canonical, $(r_W,\ g_W,\ b_W)$. So the diagonal transform from unknown to canonical is:

$$\left(\frac{r_W}{r_M},\ \frac{g_W}{g_M},\ \frac{b_W}{b_M}\right) \tag{3.3}$$

3.3 Sensor Sharpening

As discussed in §1.4, sensor sharpening is a promising method to improve color constancy where diagonal models are used. In sensor sharpening we seek a transform **T** to apply to the data, **U**, so that a diagonal transform **D** found for color constancy yields a better solution when transformed back by $\mathbf{T}^{-1}$. Several methods for producing **T** have been worked out [FDF94a]. It was found that all methods gave similar results, but it should be noted that the test domain was human cone response estimates, not camera sensor functions. In the case of camera sensors, it was recognized that these were already quite sharp. Hence it was expected that the effect would be either negligible, or slightly beneficial. Instead, consistent benefits were difficult to obtain, and the conclusion is that more work is required before sharpening can be applied with confidence in this domain.

Two sharpening approaches were investigated. The first is referred to as database sharpening. Here response data is generated for a number of surfaces under a given light and under a second, canonical light. The transform **T** is found by diagonalizing the best least squares transform matrix between the two sets. In this case, the diagonal model, together with sharpening, yields the same least squares fit as the general linear model (see [FDF94a] for details).

A second sharpening method is perfect sharpening. Unlike database sharpening, this method works with the entire space of illuminants. For perfect sharpening, a three-dimensional model of reflectance and a two-dimensional model of illumination is assumed. Two independent illuminants are selected, a canonical and one other. Together these span the

two-dimensional space of illuminants. On this assumption, the lighting matrix for the canonical illuminant, Λ^C, is a linear transform from the lighting matrix for the second illuminant Λ^2. Thus there is a matrix **M** such that:

$$\Lambda^2 = \mathbf{M}\Lambda^C \tag{3.4}$$

M is easy to compute:

$$\mathbf{M} = \Lambda^2(\Lambda^C)^{-1} \tag{3.5}$$

The required sharpening transform **T** is obtained by diagonalizing **M**:

$$\mathbf{M} = \mathbf{T}^{-1}\mathbf{D}\mathbf{T} \tag{3.6}$$

(See [FDF94a] for additional details).

When sensor sharpening was attempted, several problems arose. First, using database sharpening required additional procedures because it is designed to provide good diagonal color constancy only between two illuminants. However, by definition, the color constancy problem must deal with more than two illuminants. Second, when a sharpening transform is applied, it is possible to produce negative input to the color constancy programs. This is a problem for some of the algorithms implemented. The most problems of this type occurred with gamut mapping algorithms in chromaticity space. Problems arise because the inclusion of negative data values can cause some sets to be non-convex. For example, a cornerstone of gamut mapping theory in chromaticity space is the fact that convex sets in RGB space remain convex in chromaticity space. But consider a set with negative as well as positive blue values. Then the set also contains points with zero blue values. Computing chromaticities means that we divide by zero. Previously we could have sets that included the origin, but since it was

not crossed, the singularity could be removed by considering it as a limiting case of points close to the origin. This approach will clearly not work now. Hence other methods of dealing with the problem must be tried.

The severity of the problem depends on which set has the negative values. If it is just a data value, then that value could be ignored. We can certainly pretend that a surface is not in the scene, and this is in fact what is done in the implementation. Since the program was run on numerous combinations of surfaces, it was possible that the input could consist only of responses with negative blue values. Rather than design an appropriate penalty for this, the entire test was designated as having failed.

If the canonical gamut has negative blue values, or if the illumination gamut has negative values for any component, then things are even worse. Truncating negative values at zero would be an option, but this tack lacks supporting theory, and was not tried. In summary, for the gamut mapping approach, it is best to use sensor transforms that tend not to produce negative values. How this should be done is an unanswered question.

I will now return to the problem of dealing with multiple illuminants. One proposal is to average database sharpening results for a number of illuminants.[2] This method has several pitfalls. One is that if a large number of these matrices are diagonalized, then complex eigenvalues can and do occur.[3] More troublesome is that averaging the matrices is a somewhat ill-

---

[2]This suggestion is due to G. Finlayson

[3]It is not known what to do with complex eigenvalues. Apparently these never occurred in the original experiments. One assumes that if there not too many, then they could safely be ignored.

defined procedure. This is because the order of the eigenvectors must be arranged before averaging, since the ordering of the change of basis coordinates, as delivered by diagonalization, is arbitrary. A heuristic is required for this rearrangement. In addition, some of eigenvectors need to be negated via another heuristic. Nonetheless, the sharpening transform produced by averaging the database results for the test lights was tried with two different canonical illuminants. Using the test lights, of which there are only a small number, means that the above problems can be dealt with manually, but the experiment is biased.

In response to the above difficulties, another method of database sharpening is proposed. For standard database sharpening we form a matrix **U** of the responses for all surfaces under one light, and a matrix **C** of the responses under the canonical light, and diagonalize the best map between them. Now if all lights are equally likely, then intuitively we want to have the best map from **all** surfaces under all lights to all surfaces under the canonical light. This means **U** should be the responses of the first surface under all N lights, followed by the responses of the second surface under all N lights, and so on. **C** becomes the response of the first surface under the canonical light N times, followed by the response of the second light N times, and so on. Following Finlayson et. al. [FDF94a], we then solve $\mathbf{C} = \mathbf{GU}$ in the least squares sense by:

$$\mathbf{G} = \mathbf{C}\mathbf{U}^{+} \tag{3.7}$$

where + denotes the Moore-Penrose inverse which is defined by:

$$\mathbf{A}^{+} = \mathbf{A}^{\mathbf{T}}[\mathbf{A}\mathbf{A}^{\mathbf{T}}]^{-1} \tag{3.8}$$

The matrix **G** is then diagonalized to obtain the transform **T**:

$$\mathbf{G} = \mathbf{T}^{-1}\mathbf{D}\mathbf{T} \tag{3.9}$$

In contrast to the original formulation of database sharpening, perfect sharpening is put forth as a method that overcomes the multi-illuminant problem. However, there are a few points to consider before treating it as such. The sharpening transforms are calculated from the lighting matrices of two linearly independent illuminants. Under the assumption that illuminants are two-dimensional, it does not matter which ones they are. However, this is only a crude approximation. Since illuminants are not two-dimensional, the choice of which ones to use requires some attention. One choice is to be as faithful as possible to the two-dimensional model. Here the first two least squares fitting vectors as computed by SVD are used. An alternative is to include the canonical illuminant as motivated by the above development. Two possible choices for the second vector are the first SVD basis vector, and the first SVD basis vector of the space reduced in dimension by the canonical. This latter method simply recognizes that since the canonical is already available for fitting, any component in this direction in the basis is redundant. Thus a slightly better fit will be possible if the canonical is excluded.

3.4 Results

3.4.1 Input Data

The experiments in this chapter used three different data sets. The first will be referred to as carefully controlled image data. Results using this data

are only reported for the sensor sharpening experiments. For this data, a picture of the Macbeth color checker was taken under five test lights. These were a pure incandescent studio light, the light from a slide projector, a Sylvan warm white fluorescent light, a Philips cool white fluorescent light, and the Philips cool white with a blue filter. This last combination was chosen to simulate blue sky, and it was verified with the spectraradiometer that it was close to deep blue sky both in camera chromaticities and CIE XYZ. Effort was made to illuminate the color checker as evenly as possible, although with the equipment available it was not possible to do this particularly well. Once the pictures were taken, the pixels of the inner 2/3 of the patches were averaged to produce the data. The patches in the middle two rows of the color checker were used for a total of 12 surfaces.

The second data set was generated from the camera sensors and the measured spectra described in detail in §2.2 and §2.3. The same test lights used for the carefully controlled image data were used. Results using this data set are reported for the sensor sharpening experiments in §3.3.3, and the simulation tests of color constancy algorithms in §3.3.4.

The third data set was a sequence of four scenes taken under three lights. Care was taken not to bump the camera between the lighting changes, as proper registration between the images was required for performance measurement. This is explained more fully in the next section, and the image data is described further in §3.4.5.

The canonical light was chosen to be the Philips cool white fluorescent light, which is the whitest of the lights. For some experiments, the studio light was also tried. This light is the closest to the ideal light to use when the

camera is setup for indoor lighting, as was the case for all experiments. Thus it is the appropriate canonical light for color correction.

3.4.2 Format of Results

Color constancy performance is expressed in terms of a direct comparison to the desired answer, which is the camera sensor response under the canonical light. In the case of the simulations, it is important to note that the measure used is based on mapping the responses for all the surfaces in the data set into those for the canonical illuminant, regardless of the number of surfaces used to estimate the mapping. This reflects the philosophy that there is a right answer to be found for the particular illuminant which applies to all surfaces, and we wish to know how well we estimate it. In the case of image data, all surfaces are used, both for computation and for performance measurement, so there is no ambiguity. The best linear and diagonal fits for both 2 and 3 dimensions are provided to help characterize the "right answer". When chromaticity results are reported (as is usual), the three-dimensional best fit results are the best fits in three dimensions, projected into chromaticity space.

The difference between the estimate and the target is the RMS of the chromaticity vector difference magnitudes taken over all surfaces. In all cases where more than one test contributed to the result, the average was reported. For most results, an error estimate is provided. This is the estimated error of the mean, which is the standard deviation divided by the square root of the

sample size.[4] This should not be confused with the standard deviation itself. The error of the mean is an estimate of the variation of the tabulated averages, if the entire experiment was done repeatedly, rather than only once.

In order to tabulate results compactly, the algorithms are referred to by the following abbreviations:

Best 2D linear transform	BT2
Best 2D diagonal transform	BD2
Grey world algorithm	GW
Retinex	RET
Two-dimensional gamut mapping, surface constraints only, with solution chosen by the maximum volume heuristic	SMV2
Two-dimensional gamut mapping, surface constraints only, with solution chosen by the hull average	SHA2
Two-dimensional gamut mapping, surface and illumination constraints, with solution chosen by the maximum volume heuristic	SIMV2
Two-dimensional gamut mapping, surface and illumination constraints, with solution chosen by hull average	SIHA2

The three-dimensional algorithms will be similarly abbreviated, except that the "2" is replaced with a "3".

3.4.3 Sensor Sharpening Results

A number of sharpening methods were tried (see §3.3 for technical details). First, database sharpening was applied to the carefully controlled

---

[4]The validity of the estimate depends on the assumption that the distribution of the results is approximated by a Gaussian distribution. This assumption was not confirmed.

image data, which spans the five test illuminants. The transforms for all five to the canonical were averaged, minding the pitfalls mentioned in §3.3. This method is favorably biased as the test illuminants were used exclusively. Second, the alternative method of database sharpening proposed above was used. All three variants of perfect sharpening described in §3.3 were also tried, using the measured data to obtain basis vectors. It should be noted that the data for the illuminants was normalized for all the sharpening calculations. The abbreviations used for the tabulation of results are as follows:

Database sharpening on small data set	DBS
New database sharpening on all data	DBA
Perfect with 1st and 2nd principal component	P12
Perfect with canonical and 1st principal component	PC1
Perfect with canonical and 1st P.C of reduced data.	PCR1
Best Diagonal Transform	BDT
Best Transform	BT

The first test of sharpening is how well the data generated from all measured surfaces and illuminants can be mapped to the desired canonical result (in the least squares sense). The RMS error results are tabulated in Table 3.1. Note that the result for DBA is (provably) the same as that for BT3. For this experiment, a quantity proportional to RGB (not chromaticity) was used. The magnitude of the numbers is a consequence of the units of the sensor curves because the illuminants were normalized. Hence only the relative sizes of the numbers are significant.

Sharpening method	Result with P-CW canonical	Result with studio canonical
DBS	1735	948
DBA	1587	815
P12	1731	893
PC1	1722	875
PCR1	1723	876
BD3	1736	879
BT3	1587	815

Table 3.1 Results of sensor sharpening tests for mapping all responses that can be generated from the measured data into the appropriate response for the canonical light.

The most important result above is that the camera sensors are already quite sharp. The best diagonal transforms have only about 9% more error than the best linear map. The second observation is that in general, other than the DBA method which is tuned for this particular test, sharpening does not significantly improve the answer, since the results are all comparable to BD3. Thus, if the distribution of illuminants and reflectances in the generated sets are representative of the data over which the color constancy experiments will be run, then the other sharpening methods will not improve the results significantly. For the typical image, this assumption will only be approximately true. In the case of the illuminants, some attempt was made to obtain a representative set. In the case of the reflectances, the opposite is true; these were chosen to expand the gamut. Much more work is required to make a definitive statement about which sharpening method is best for the ill-defined average collection of scenes.

The second test of sharpening is the degree to which it improves the performance of color constancy experiments. I tested all the sharpening methods on both the carefully controlled image data and the generated data, with both choices of canonical lights. In the case of the carefully controlled image data, tests with 4 and 8 surfaces out of 12 were done. In the case of the generated data, 4, 8, 12, and 16 surfaces out of 78 were done. For each combination of data set and number of surfaces, 100 randomly generated surfaces were used in conjunction with each of the 5 test lights, giving 500 data points altogether. The actual set of surfaces was the same for all tests for a given number of surfaces of a given data set. Since the implications of all the results are consistent, only data for selected algorithms in the case of 8 surfaces with the measured data and studio light as a canonical are provided. This particular result was chosen for inclusion because it had a relatively small number of missing values due to the problems with negative sensor input mentioned in §3.3. Such missing values were more prevalent when the cool white fluorescent was used as a canonical.

Results are provided for RGB (Table 3.2) and chromaticity (Table 3.3). The main thrust of these results is that none of the sharpening methods tested yield clear improvements for all the diagonal model algorithms. The results show no major effect, with a given sharpening method typically improving one algorithm, but hindering another. This variation persists in the data not included for illustration. Thus, in the case of the current collection of surfaces, the current estimate of typical illumination, and our camera sensors, more work is required to identify the parameters affecting the utility of sharpening.

Algorithm	None	DBS	DBA	P12	PC1	PCR1
BT3	5.1 (0.2)	5.1 (0.2)	5.1 (0.2)	5.1 (0.2)	5.1 (0.2)	5.1 (0.2)
BD3	7.0 (0.2)	8.7 (0.3)	7.7 (0.3)	8.6 (0.3)	6.3 (0.2)	6.3 (0.2)
GW	95 (1)	94 (1)	110 (2)	98 (1)	95 (1)	95 (1)
RET	78 (1)	98 (1)	64 (1)	77 (1)	75 (1)	76 (1)
SIHA3	31.3 (0.5)	28 (0.3)	24 (0.5)	30 (0.5)	31 (0.3)	31 (0.3)

Table 3.2 RGB results of sharpening experiments for 100 random groups of 8 surfaces for each of the five test lights using the carefully controlled image data. The canonical light is the studio light. The values shown are the average magnitude of the RGB vector difference between the estimate and the desired answer, averaged over all 500 results. The value in parenthesis is the error estimate of this average. The sample standard deviations are roughly 20 times this estimate.

It is possible that more extensive testing would reveal some advantage to using one of these transforms for a particular algorithm. However, the results for the transform best tuned to this data (DBS) are not encouraging in this regard. Furthermore, it seems that attempts to use transforms which were more global lead to worse results for some of the particular cases, especially with respect to producing data that some of the algorithms cannot deal with. For these reasons, sharpening was not used for subsequent experiments.

Algorithm	None	DBS	DBA	P12	PC1	PCR1
BT2	0.26 (0.01)	0.26 (0.01)	0.26 (0.01)	0.26 (0.01)	0.26 (0.01)	0.26 (0.01)
BD2	0.31 (0.01)	0.36 (0.01)	2.57 (0.01)	0.38 (0.01)	0.30 (0.01)	0.30 (0.01)
GW	1.10 (0.02)	1.12 (0.02)	6.2 (1)	1.2 (0.02)	1.1 (0.02)	1.1 (0.02)
RET	1.19 (0.01)	1.04 (0.01)	6.1 (1)	1.24 (0.02)	1.15 (0.01)	1.15 (0.01)
SIHA2	0.53 (0.02)	0.54 (0.02)	*	0.64 (0.02)	0.50 (0.02)	0.49 (0.01)

Table 3.3 Chromaticity results of sharpening experiments for 100 random groups of 8 surfaces for each of five test lights taken from the carefully controlled image data The canonical light is the studio light. The measurement labeled * could not work at all due to negative RGB values in the transformed data . The values shown are the average magnitude of the chromaticity vector difference between the estimate and the desired answer, averaged over all 500 results. The value in parenthesis is the error estimate of this average. The sample standard deviations are roughly 20 times this estimate.

3.4.4 Color Constancy Simulation Results

I will now present two sets of performance results for the various algorithms. In the first, 500 sets of a variety of numbers of surfaces were randomly selected,[5] and the predicted responses to these surfaces under the simulated blue sky light were generated. These responses were presented to each algorithm. The average magnitude of the error as a function of the number of surfaces is shown in Tables 3.4 and 3.5.

[5]Except for the case of one surface, where only 78 possibilities were available. In this case, all 78 where used.

The second set excludes the three-dimensional algorithms (which take much longer to run), but is more comprehensive. Here 1000 sets of a variety of numbers of surfaces were run against each of the 5 test lights, for a total of 5000 tests for each number of surfaces.[6] The results are shown in Table 3.6.

With the current implementation, the run time in the case of chromaticities is of the order of a second on a Sun Sparc 10 workstation, and of the order of a minute in the case of RGB's. This remains true even if thousands of surfaces is used, as is the case of image data. This may seem a little anomalous, but the computationally intense part of the process works on the convex hull of the input set.

The results verify that all algorithms work better as the number of surfaces increase. The strength of the illumination constraint is also verified, which, as expected, makes the biggest difference when only a small number of surfaces are available. Using Table 3.3 for three-dimensional results we see that the illumination constraint improves the answer by factors of roughly 5, 3.6, and 2.5, for 1, 2, and 4 surfaces respectively (comparing hull average algorithms). The differences in Table 3.5 for two dimensions is even more pronounced, but the reader should be cautioned that in the case of a small number of surfaces, the chromaticity results are overly influenced by a few instances where the blue value is close to zero. This means that the uncertainty in the chromaticity becomes very large. Thus the large errors are due to a small number of bad data points. These problems could be deal with,

---

[6]Except for the case of one surface, where only 78 possibilities were available. Here all 78 combinations were used instead of 1000 randomly selected ones.

but it is not worth the effort, since they do not occur with the illumination constraint.

The gamut mapping approach generally does better than the grey world and Retinex algorithms provided that the illumination constraint is used. The exceptions are the two-dimensional results for all test lights in the case of a small number of surfaces. This result is somewhat counter-intuitive, and suggests that adding surfaces constrains the solutions more effectively than it incrementally makes the world grey. Overall, the grey world and Retinex algorithms do well given their simplicity. As a rule of thumb, the more sophisticated approaches are only roughly twice as effective. Of course, the statistics do not deal with the common criticism of these algorithms; namely, that it is easy to design a scene for which they give very poor results.

As the number of surfaces increases, the illumination constraint becomes less important. It is interesting to note that for a very large number of surfaces, the addition of the illumination constraint degrades the answer a little. The explanation is that in the case of the simulated blue sky, which is near the edge of the illumination gamut, the best fit is slightly outside the illumination constraint. The error is measured explicitly by mapping the results of surfaces under one light to that for the canonical, and a large number of surface constraints will push the answer towards a best diagonal fit of the surface data.

As expected, in the two-dimensional case, the hull average performed better than the maximum volume heuristic. We had hoped to model an evenly distributed data set, and a counter-result would be indication that this goal was not achieved. Except where the error was small, and thus largely due

to factors that cannot be resolved by these algorithms (such as a failure of the diagonal model), using the hull average typically improved answers by a factor of two.

The difference is distinctly less in the three-dimensional case. In RGB space, the maximum product point will be somewhere towards the middle of the possibilities once it is projected. However, it is clear that the maximum volume heuristic would be suspect if we were trying to recover full RGB values, given a least squares error measurement. Here it is expected that the average would do better.

One very interesting observation on this topic is that when a large number of surfaces were included in the three-dimensional case, the maximum volume heuristic gave a better answer. As the mapping volume increases, the measured gamut is fit into the unknown gamut with less slack. When the volume is maximal there is, of course, no slack with one or more responses. In the case of a large number of surfaces, the fit must be good. What is interesting is the possibility that this may give a better answer than the average in some cases. However, it does not work as well in the case of a small number of surfaces, and this is where it is most important to do well.

Along similar lines it is interesting that the three-dimensional method, together with the maximum volume heuristic, gave the best results out of the algorithms tested when the number of surfaces was large. To understand how this can be the case (given that the results are being projected), it is useful to note that the three-dimensional surface constraints are in fact more powerful than the two-dimensional ones. Once the three-dimensional sets are projected, some information is lost. Another way to understand this is to consider that the projection of the intersection of two

three-dimensional sets is **not** necessarily the intersection of the projections. The former will always be a subset, but it may be a proper subset. Of course, for the sets that occur in this sort of application, the difference tends to be small. The situation can be visualized by realizing that the two-dimensional surface constraints become cones in three dimensions, but the three-dimensional constraints are not cones; sometimes the boundary opposite the origin plays a role.

The above does not address the issue that the two-dimensional algorithm is automatically robust with respect to shading effects, which do not occur in the generated data. In the three-dimensional case, it is possible that solving for three parameters with variation in illumination intensity as a confound may degrade the chromaticity recovery significantly. Another possibility is that the variation in illumination intensity would simply become uncertainties in the estimate of the magnitudes of the reflectances, and that the chromaticity results would be comparable to the two-dimensional case. This latter possibility is consistent with the results for image data presented shortly.

Thus additional work is required to determine if using three dimensions is warranted in the interesting case of a small number of surfaces. As implied above, the experiments done so far have not been designed to properly test this. One expects that if there is a gain, it will not be enough to offset the large jump in processing cost introduced by adding the third dimension. Nonetheless, it has been verified that if the three-dimensional model is desired, as would be the case if one wishes to recover lightness, then chromaticities will be estimated at least as well as they are using the two-dimensional algorithm, provided that there is no variation in illumination.

	1	2	4	6	8
BF2	0.063	0.063	0.063	0.063	0.063
BF3	0.063	0.063	0.063	0.063	0.063
BD2	0.103	0.103	0.103	0.103	0.103
BD3	0.102	0.102	0.102	0.102	0.102
GW	1.54 (0.2)	1.03 (0.04)	0.65 (0.02)	0.53 (0.01)	0.48 (0.01)
RET	1.54 (0.2)	1.14 (0.02)	0.68 (0.02)	0.55 (0.02)	0.48 (0.01)
SMV2	54 (12)	16 (2)	4.3 (0.3)	2.9 (0.1)	1.3 (0.1)
SHA2	30 (6)	9 (1)	2.0 (0.1)	1.2 (0.1)	0.55 (0.04)
SIMV2	0.838 (0.01)	0.81 (0.01)	0.78 (0.01)	0.75 (0.01)	0.60 (0.01)
SIHA2	0.334 (0.06)	0.325 (0.003)	0.320 (0.003)	0.305 (0.003)	0.260 (0.05)
SMV3	2.7 (0.4)	1.61 (0.1)	0.67 (0.02)	0.46 (0.02)	0.29 (0.01)
SHA3	2.8 (0.4)	1.68 (0.1)	0.71 (0.01)	0.50 (0.02)	0.32 (0.01)
SIMV3	0.80 (0.04)	0.67 (0.02)	0.446 (0.02)	0.36 (0.01)	0.26 (0.01)
SIHA3	0.56 (0.04)	0.46 (0.02)	0.380 (0.01)	0.34 (0.01)	0.28 (0.05)

Table 3.4 Results of color constancy experiments for 500 sets of 1, 2, 4, 6, and 8 surfaces as viewed under simulated blue sky light. The canonical illuminant was a Philips CW fluorescent light. The values shown are the average magnitude of the chromaticity vector difference between the estimate and the desired answer, averaged over all 500 results. The value in parenthesis is the error estimate of this average. The sample standard deviations are roughly 20 times this estimate (except for the first column where it is 8 times the estimate). Since there is only one unknown illuminant, the errors for the best fits are zero.

	12	16	20	24	32
BF2	0.063	0.063	0.063	0.063	0.063
BF3	0.063	0.063	0.063	0.063	0.063
BD2	0.103	0.103	0.103	0.103	0.103
BD3	0.102	0.102	0.102	0.102	0.102
GW	0.44 (0.01)	0.43 (0.01)	0.41 (0.01)	0.40 (0.01)	0.39 (0.01)
RET	0.40 (0.01)	0.36 (0.01)	0.33 (0.01)	0.30 (0.01)	0.256 (0.01)
SMV2	0.78 (0.02)	0.6 (0.03)	0.45 (0.01)	0.40 (0.01)	0.332 (0.01)
SHA2	0.31 (0.01)	0.281 (0.005)	0.234 (0.005)	0.228 (0.004)	0.226 (0.003)
SIMV2	0.51 (0.01)	0.42 (0.01)	0.37 (0.01)	0.34 (0.01)	0.291 (0.005)
SIHA2	0.26 (0.05)	0.266 (0.005)	0.256 (0.005)	0.253 (0.005)	0.246 (0.003)
SMV3	0.25 (0.01)	0.219 (0.005)	0.199 (0.005)	0.181 (0.005)	0.149 (0.003)
SHA3	0.24 (0.01)	0.218 (0.005)	0.199 (0.004)	0.186 (0.003)	0.168 (0.002)
SIMV3	0.24 (0.01)	0.212 (0.005)	0.194 (0.005)	0.177 (0.005)	0.146 (0.003)
SIHA3	0.32 (0.005)	0.310 (0.005)	0.297 (0.004)	0.288 (0.003)	0.271 (0.003)

Table 3.5 Results of color constancy experiments for 500 sets of 12, 16, 20, 24 , and 32 surfaces as viewed under simulated blue sky light. The canonical illuminant was a Philips CW fluorescent light. The values shown are the average magnitude of the chromaticity vector difference between the estimate and the desired answer, averaged over all 500 results. The value in parenthesis is the error estimate of this average. The sample standard deviations are roughly 20 times this estimate. Since there is only one unknown illuminant, the errors for the best fits are zero.

	1	2	4	8	16
BF2	0.073 (0.001)	0.073 (0.001)	0.073 (0.001)	0.073 (0.001)	0.073 (0.001)
BF3	0.076 (0.001)	0.076 (0.001)	0.076 (0.001)	0.076 (0.001)	0.076 (0.001)
BD2	0.116 (0.002)	0.116 (0.002)	0.116 (0.002)	0.116 (0.002)	0.116 (0.002)
BD3	0.132 (0.002)	0.132 (0.002)	0.132 (0.002)	0.132 (0.002)	0.132 (0.002)
GW	1.62 (0.03)	1.01 (0.01)	0.69 (0.01)	0.513 (0.004)	0.428 (0.003)
RET	1.62 (0.03)	1.10 (0.01)	0.72 (0.01)	0.478 (0.004)	0.354 (0.003)
SMV2	21.9 (0.5)	8.4 (0.1)	3.43 (0.1)	1.37 (0.02)	0.62 (0.01)
SHA2	12.4 (0.3)	4.4 (0.1)	1.65 (0.05)	0.585 (0.01)	0.285 (0.003)
SIMV2	5.28 (0.05)	3.87 (0.02)	2.28 (0.05)	1.12 (0.02)	0.581 (0.005)
SIHA2	2.275 (0.2)	1.677 (0.02)	0.99 (0.02)	0.480 (0.01)	0.271 (0.003)

Table 3.6 Results of two-dimensional color constancy experiments for 1000 sets of 1, 2, 4, 8, and 16 surfaces for each of the five test lights. The canonical illuminant was a Philips CW fluorescent light. The values shown are the average magnitude of the chromaticity vector difference between the estimate and the desired answer, averaged over all 5000 results. The value in parenthesis is the error estimate of this average. The sample standard deviations are roughly 70 times this estimate (except for the first column where it is 20 times the estimate).

3.4.5 Image Data Results

The algorithms dealt with in this chapter were also tested on image data. The images included a Macbeth color checker, a "Mondrian" made up of colored construction paper, an off-white wall with some colored pieces of construction paper, and a shelf of books. The shelf of books is considered a reasonable attempt at producing a standard scene which has all the complications inherent in real scenes such as complex boundaries, saturated pixels, shadows, shape, and much non-Lambertion reflection.

All the algorithms in this chapter can work with the assumption of one surface per pixel. Some pixels have contributions from more than one physical surface, but the resulting signal will be a linear combination of the contributions, and thus will always be a valid input. This situation remains true if we average the scene to reduce resource use. Of course, as the resolution is reduced, either physically or by averaging, some information will be lost.

The illuminants used included some of the test illuminants of the previous sections as well as three others; a bright halogen bulb which was a little redder than a standard incandescent light, the studio light bounced off the ceiling, and the overhead light in the lab which is quite red. Not all lights were tried on all pictures. Again the Philips CW fluorescent light was used as the canonical light.

Performance measurements are provided on the basis of a picture of the scene taken under the canonical light. Care was taken to avoid bumping the camera while switching the lights, as the pixel correspondence between images was used to compare the estimated result to the actual result. The use

of disparate lights, rather than filters, means that the lighting geometry was not uniform between test and canonical lights. The studio light is roughly a point source, but the glass around the filament distorts the intensity significantly. The fluorescent light is quite spread out, and had to be placed such that it illuminated the image unevenly in intensity. Every light produced a different geometry. In the case of the bookshelf this led to different parts of the image being in shadow with different lights. These differences were embraced as part of the challenge confronting the algorithms, and were thus accepted without any attempt to control them.

The numerical results obtained are tabulated at the end of this section. Before turning to them, three specific results will be presented visually. These consist of the Macbeth color checker under the vision lab overhead light (quite red), a color Mondrian under a regular incandescent light (red, but less so), and the book scene under simulated blue sky.

The results for the Macbeth color checker taken under the overhead light is shown in Figure 3.3. The input is in the upper left corner. The reddening effect on the non-saturated grey on the bottom row is clear. The bottom right corner is the same scene under the fluorescent canonical light. The upper right corner shows the result of running the grey world algorithm, and the lower left corner shows the result of running the algorithm for gamut mapping in chromaticity space with surface and illumination constraints and using the hull average to choose a solution. Both algorithms work well, which is to be expected due to the nature of the input.

Figure 3.3 Results of color constancy processing on the Macbeth color checker viewed under the vision lab overhead light (quite red). The input is in the upper left corner, the result of the grey world algorithm is in the upper right corner, the result of the two-dimensional gamut mapping algorithm with both illumination and surface constraints is in the lower left corner. The same object under the canonical light (Philips CW fluorescent) is shown in the lower right corner.

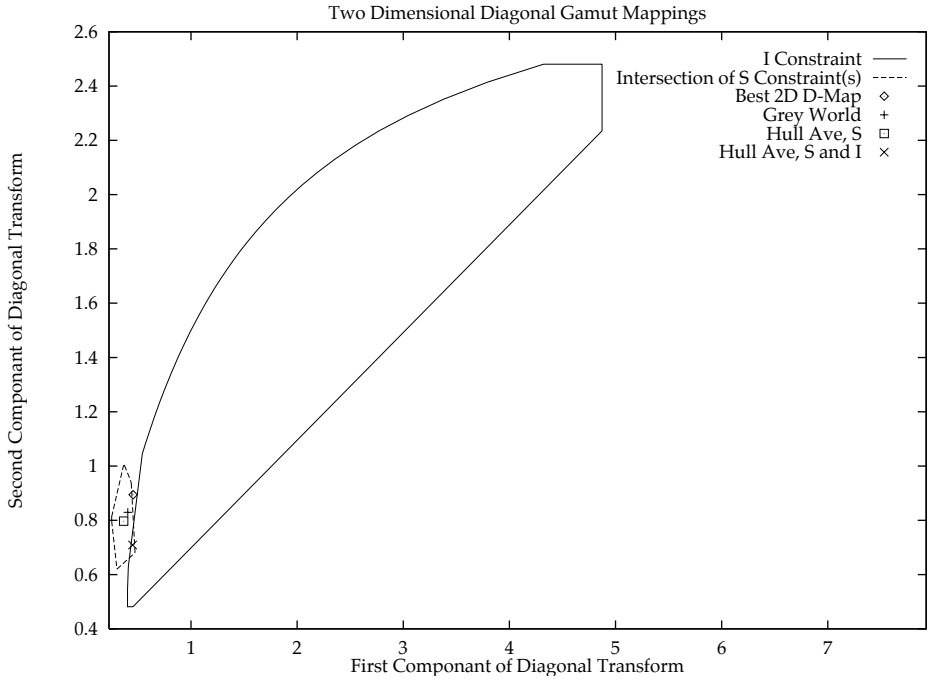

Figure 3.4 Illumination gamut mapping constraints for Macbeth color checker under the vision lab overhead light (reddish) with Philips CW fluorescent used as the canonical light.

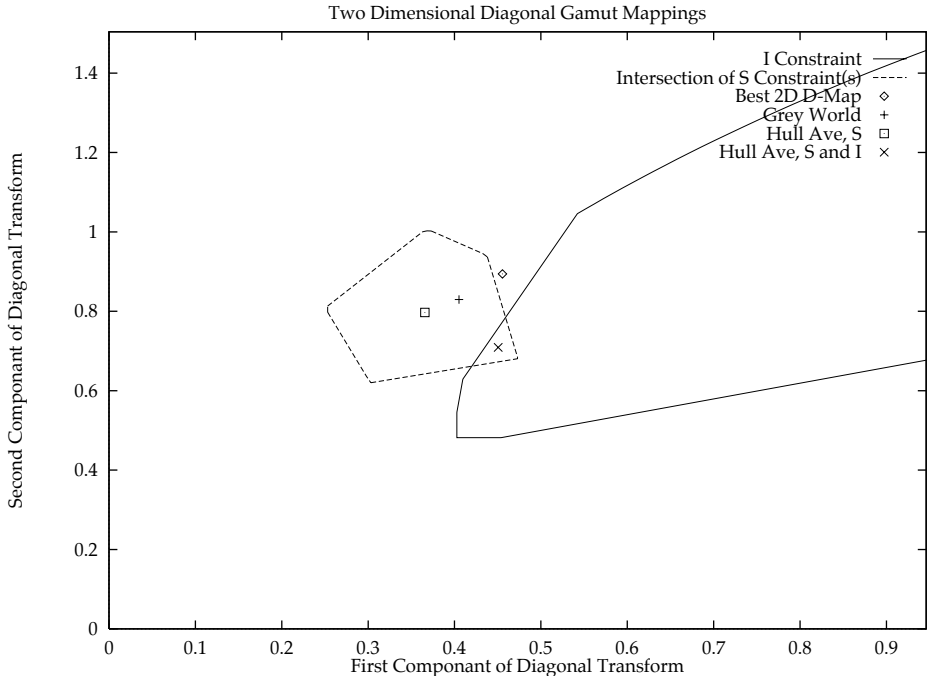

Figure 3.5 Figure 3.4 enlarged to show intersection region. (The relative sizes of the units on the two axes has also changed.) Clearly the large numbers of colors in the Macbeth color checker allow good color constancy.

The next set of results is for a Mondrian made up of colored paper. The unknown light is a regular incandescent light, and the canonical light is the Philips CW fluorescent. The results are shown in Figure 3.6. Again, the images in the composite are the input at upper left, the grey world result at upper right, two dimensional gamut mapping with both surface and illumination constraints at lower left, and the scene under the canonical illuminant at lower right. For this input, the grey world algorithm fares less well. This is due to the inclusion of more yellow and red areas and less blue and green ones, and thus the grey world result is distinctly too blue. The distribution of the colors was not designed for this effect, but nonetheless, this result nicely demonstrates the weakness of the grey world algorithm. It is also clear that the gamut mapping result is a little too red. However, Figure 3.7, which plots the constraints on the mappings, shows that the best solution is well within the constraints, demonstrating a benefit of the approach. The correct solution is characterized, even it the required value cannot be well estimated.

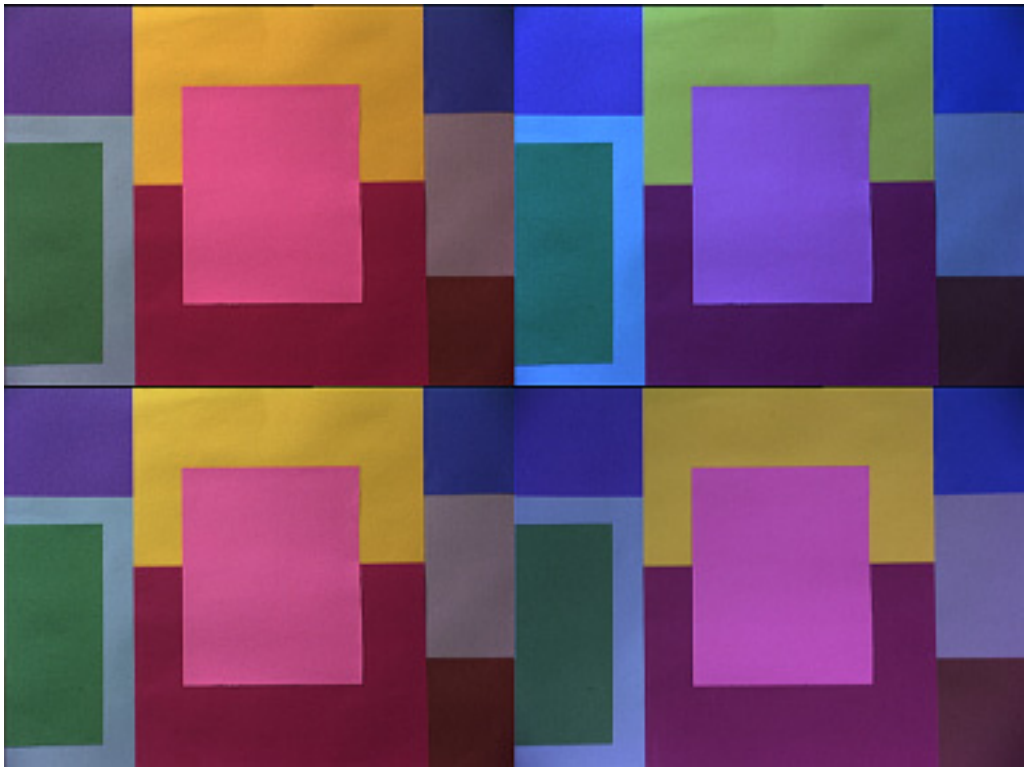

Figure 3.6 Results of color constancy processing on a scene made from colored construction paper. The upper left corner is the scene under a regular incandescent light. The upper right corner is the result of applying the grey world algorithm. The lower left corner it the result of applying the two dimensional gamut mapping method with both surface and illumination constraints applied. The desired answer is the scene under the canonical light which is shown in the bottom right corner. This image is a little blue because the camera is set for a redder light than the canonical (Phillips CW fluorescent).

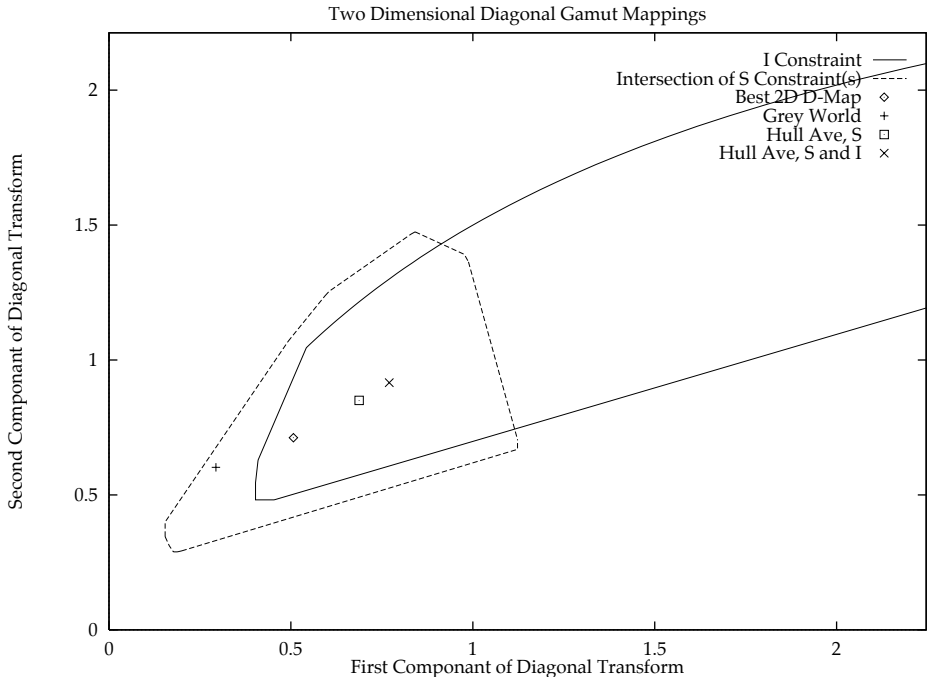

Figure 3.7 Constraints on the illumination mappings for the Mondrian shown in Figure 3.6 viewed under incandescent light. The canonical illuminant is the Phillips CW fluorescent light.

The final result shown in detail is the book scene as viewed under simulated blue sky. The corresponding composite of input, results, and output is shown in Figure 3.8.

Figure 3.8 Results of color constancy processing the book scene. The upper left corner is the scene under a simulated blue sky. The upper right corner is the result of applying the grey world algorithm. The lower left corner is the result of applying the two dimensional gamut mapping method with both surface and illumination constraints applied. The desired answer is the scene under the canonical light which is shown in the bottom right corner.

The results sampled in the previous pages are now presented numerically. The results are for all algorithms on the three scenes already shown, as well as a fourth scene of colored paper on an off-white wall. In all cases the canonical light is the Philips CW fluorescent. Table 3.7 shows the results when the unknown illuminant is an incandescent light, Table 3.8 shows the results when the unknown is the Philips CW fluorescent,[1] and Table 3.9 shows the result for simulated blue sky. A problem with the Retinex algorithm should be mentioned at this point. If the image contains saturated pixels, then the Retinex algorithm will not work optimally because the maximum possible lightness due to the scene makeup is not available to the algorithm. Thus the results for the Retinex algorithm should be considered with this in mind, since most of the images have saturated pixels in at least one channel. This problem does not occur in simulations, nor are there any saturated data in the carefully controlled data set.

Certainly this small set of pictures is not representative of all images under all lights. However, chromaticity gamut mapping with both illumination and surface constraints using the hull average to select a solution (SIHA2) gives better results than the grey world approach in 11 cases out of 12. Forsyth's CRULE together with the illumination constraint (SIMV3) also gives a better answer than the grey world 11 times out of 12 (not the same 11). Thus we tentatively conclude that these algorithms are good candidates for use with image data.

---

[1]It is important to test the algorithms even when the correct answer is to do nothing.

	Macbeth color checker	Mondrian	Wall with colored paper	Books
BF2	0.114	0.064	0.139	0.114
BF3	0.211	0.075	0.178	0.211
BD2	0.132	0.091	0.236	0.132
BD3	0.169	0.098	0.325	0.169
GW	0.317	0.413	0.330	0.317
RET	0.291	0.350	0.265	0.291
SMV2	0.231	1.109	0.334	0.231
SHA2	0.269	0.370	0.317	0.269
SIMV2	0.233	1.108	0.329	0.233
SIHA2	0.156	0.531	0.241	0.156
SMV3	0.287	0.215	0.287	0.287
SHA3	0.299	0.152	0.297	0.299
SIMV3	0.294	0.142	0.264	0.294
SIHA3	0.148	0.332	0.309	0.148

Table 3.7 Results of color constancy processing on the four real scenes under incandescent light. All algorithms discussed in detail in this thesis are shown. The canonical light is the Philips CW fluorescent light. Since there is only one scene of each class, no error estimate can be made.

	Macbeth color checker	Mondrian	Wall with colored paper	Books
BF2	0.000	0.000	0.000	0.000
BF3	0.000	0.000	0.000	0.000
BD2	0.000	0.000	0.000	0.000
BD3	0.000	0.000	0.000	0.000
GW	0.280	0.395	0.351	0.143
RET	0.436	0.342	0.280	0.449
SMV2	0.090	1.006	0.479	0.216
SHA2	0.229	0.297	0.126	0.177
SIMV2	0.091	1.006	0.469	0.216
SIHA2	0.236	0.283	0.091	0.142
SMV3	0.191	0.226	0.263	0.140
SHA3	0.243	0.172	0.210	0.095
SIMV3	0.191	0.222	0.260	0.140
SIHA3	0.159	0.095	0.113	0.047

Table 3.8 Results of color constancy processing on the four real scenes under the canonical light (Philips CW fluorescent). All algorithms discussed in detail in this thesis are shown. Since there is only one scene of each class, no error estimate can be made. The error in all forms of the best fit is zero because the identity map works perfectly.

	Macbeth color checker	Mondrian	Wall with colored paper	Books
BF2	0.134	0.053	0.043	0.108
BF3	0.114	0.043	0.041	0.095
BD2	0.190	0.054	0.054	0.109
BD3	0.145	0.075	0.051	0.110
GW	0.730	0.553	0.387	0.388
RET	0.614	0.478	0.297	0.698
SMV2	0.180	1.245	0.895	0.389
SHA2	0.502	0.342	0.117	0.219
SIMV2	0.151	0.969	0.638	0.232
SIHA2	0.537	0.135	0.150	0.264
SMV3	0.357	0.260	0.164	0.414
SHA3	0.451	0.238	0.169	0.309
SIMV3	0.357	0.213	0.170	0.420
SIHA3	0.514	0.429	0.344	0.407

Table 3.9 Results of color constancy processing on the four real scenes under incandescent light. All algorithms discussed in detail in this thesis are shown. The canonical light is the Philips CW fluorescent light. Since there is only one scene of each class, no error estimate can be made.

Chapter Four

Color Constancy with Varying Illumination

In this chapter we deal with the case where the illumination varies across the scene. As found by Finlayson et. al. [FFB95], if varying illumination can be identified, then instead of being detrimental to color constancy processing, it can be useful. However, in that work the illumination variation is identified manually, and the varying illumination for only one surface is used. These issues are dealt with here. First the theory is modified to use all scene illumination variations in conjunction with the illumination and surface constraints. Then a method for identifying the illumination variation based on an appropriate segmentation is introduced. Finally the comprehensive algorithm is tested on simple, easy to segment, real images.

4.1 The Varying Illumination Algorithm

As implied above, the algorithm used for varying illumination is based on the one provided by Finlayson et. al. [FFB95]. This algorithm is outlined in §1.6.2. In the following text a different approach is introduced which retains the essence of the original idea. The motivation for the new approach was to enable the integration of the varying illumination constraints with the surface constraints, as well as provide for the use of more than one such constraint. In addition, the ability to use an arbitrary illumination gamut was sought. As a further benefit, the modified approach is valid for both the two-dimensional and three-dimensional cases.

To begin, assume that we have an illumination-variation map for the image. This is defined by the diagonal transform required to map the illumination at the center of the image to every other point. The idea is to solve for the illumination in the center using this map. (The strategy used for determining such a map for a real image is the topic of §4.3). The approach is intuitively motivated by the following argument. Suppose that the left side of the image is illuminated by a blue light. This means that the entry in the illumination-variation map for a pixel on the left side transforms illuminants so that they are more blue. Now the illumination at the center of the image cannot be so blue that making it even more blue produces an illumination that falls outside the gamut. Thus the illumination at the center is constrained by the jump towards blue. All entries in the map contribute this sort of constraint. This will now be made a little more formal.

First we verify the intuitive claim that the constraint due to a given illumination-variation map, **D**, is the illumination gamut scaled by $\mathbf{D}^{-1}$. Consider the non-inverted illumination gamut, $\mathcal{I}$ which is a convex set:

$$\mathcal{I} = \left\{ \mathbf{X} \;\middle|\; \mathbf{X} = \sum_i \lambda_i \mathbf{X}_i \text{ where } \sum_i \lambda_i = 1 \right\} \text{ for hull points } \left\{ \mathbf{X}_i \right\} \tag{4.1}$$

We have the constraint that we can map the illumination by the diagonal map **D** and still be in this set:

$$\mathbf{XD} \in \mathcal{I} \tag{4.2}$$

This means that:

$$\mathbf{XD} = \sum_i \lambda_i \mathbf{X}_i \quad \text{for some } \lambda_i \text{ with } \sum_i \lambda_i = 1,\ \lambda_i \geq 0 \tag{4.3}$$

And

$$\mathbf{X} = \sum_i \lambda_i \left(\mathbf{X}_i\ \mathbf{D}^{-1} \right) \text{ for some } \lambda_i \text{ with } \sum_i \lambda_i = 1,\ \lambda_i \geq 0 \tag{4.4}$$

So we define a new constraint set $\mathcal{V}$ as:

$$\mathcal{V} = \left\{ \mathbf{X} \;\middle|\; \mathbf{X} = \sum_i \lambda_i \left(\mathbf{X}_i\ \mathbf{D}^{-1} \right) \text{where } \sum_i \lambda_i = 1,\ \lambda_i \geq 0 \right\} \tag{4.5}$$

It is clear that for all $\mathbf{X} \in \mathcal{I}$, $\mathbf{XD} \in \mathcal{V}$. Furthermore, the argument is reversible. That is, if $\mathbf{Y} \in \mathcal{V}$, $\exists \mathbf{X} \ni \mathbf{Y} = \mathbf{XD}$ for some $\mathbf{X} \in \mathcal{I}$. It should be noted that the above also shows that we can identify the convex constraint set with the mapped hull points $\mathbf{X}_i\ \mathbf{D}^{-1}$.

Next, consider the set of constraints determined from the illumination-variation map. We will verify that the convex hull of these constraints are just as powerful the entire set. The motivation for using the hull is that it saves a significant amount of processing time. We are free to

use the hull regardless, but it is comforting to know that doing so does not weaken the algorithm. To demonstrate this we need to show that given two diagonal transforms $\mathbf{D}_1$ and $\mathbf{D}_2$, the corresponding constraint sets $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ used together include all constraints in the set $\boldsymbol{\nu}$ due to $\alpha\mathbf{D}_1 + \beta\mathbf{D}_2$ ($\alpha + \beta = 1,\ \alpha,\beta \geq 0$). Now the constraints in $\boldsymbol{\mathcal{I}}$ can be expressed as a series of inequalities in matrix form (again post-multiplication is used for mappings):

$$\mathbf{XM} < \mathbf{b} \tag{4.6}$$

Our sets $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ are constructed by assuming that if $\mathbf{D}_1$ and $\mathbf{D}_2$ are applied, that the result is in $\boldsymbol{\mathcal{I}}$:

$$\begin{aligned} \mathbf{XD}_1\mathbf{M} &< \mathbf{b} \\ \mathbf{XD}_2\mathbf{M} &< \mathbf{b} \end{aligned} \tag{4.7}$$

This means that:

$$\begin{aligned} \alpha\mathbf{XD}_1\mathbf{M} &< \alpha\mathbf{b} \\ \beta\mathbf{XD}_2\mathbf{M} &< \beta\mathbf{b} \end{aligned} \tag{4.8}$$

Adding these two equations, and insisting that $\alpha + \beta = 1$ gives:

$$\mathbf{X}\left(\alpha\mathbf{D}_1 + \beta\mathbf{D}_2\right)\mathbf{M} < \mathbf{b} \tag{4.9}$$

This is the condition imposed by using the mapping $\alpha\mathbf{D}_1 + \beta\mathbf{D}_2$, and so we are done.

Although not necessary for the algorithm to work, it proved convenient to be able to apply the constraints in the inverse space. This is acceptable because, as discussed in §3.2.2, using the convex hull of the inverse is a reasonable approximation. If this were not the case, then it would be necessary do the computations in the non-inverse space, and deal with the

inverse by other means.[1] To further justify working in the inverse space, consider mapping a point in the inverse hull by **D**. Let $\mathbf{X} \in \mathcal{I}$. Then a point in the inverse hull, $1/\mathbf{X}$, mapped by **D** will be simply:

$$\left(1/\mathbf{X}\right)\mathbf{D} = 1/\left(\mathbf{X}\mathbf{D}^{-1}\right) \qquad \text{(element-wise division)} \tag{4.10}$$

This means that our inverse space mapped point ($1/\mathbf{X}$) is the inverse of a point in the mapped set of the non-inverse set (denoted by $\mathcal{V}$ above). Thus if the convex approximation is good, then we can choose to work in the inverse space.

It is hoped that the basic idea has not been lost in the details. The additional constraint is very simple. It says that we have to be able to scale the illuminant by a certain amount, and *still* satisfy the illumination gamut constraint. This constraint is realized by simply scaling the illumination gamut by the inverse. As a simple example, consider the one-dimensional line segment [0,1]. If we have a condition on these points that if they are scaled by a factor of two, the result must still be in that segment, then the set of points in our constrained set must be $[0, \frac{1}{2}]$. In other words, the set was scaled by the inverse of the scale factor.

4.2 Simulation Results

The algorithm described in the previous section was tested first by simulation. One thousand surfaces were randomly generated and used in

---

[1] In the two dimensional case one could use the method in [Fin95]. Alternately, the non-convex set could be approximated.

conjunction with combinations of the five test lights as described below. The Philips CW fluorescent was used as the canonical. Results are reported for 1, 2, 4, 8, and 16 surfaces used in conjunction with up to four simulated varying illumination sources. The input to the algorithm was the response of each of the surfaces as viewed under the unknown light (as in chapter 3), as well as the response as viewed under the extra lights. The extra lights were test lights other than the unknown. For each of the five unknown lights, all possible combinations of extra lights were used. It should be noted that the roles of a light as the unknown, and as an extra for varying illumination are not reversible. Both maximum volume and hull average results were recorded. As expected for a comprehensive test, the hull average fared better by roughly a factor of two, and so only the hull average results are reported. For each surface the number of tests done and abbreviations used for tabulation are as follows:

Single unknown light (regular case)		5 tests
One extra light	SIV1HA	20 tests
Two extra lights	SIV2HA	30 tests
Three extra lights	SIV3HA	20 tests
Four extra lights	SIV4HA	5 tests

The abbreviations introduced in §3.3.2 will also be used to cover the algorithms where no varying illumination was used. The results for these tests are shown in Table 4.1.

The results are exactly what was hoped for. As either the number of surfaces, or the number of extra lights increases, the answer consistently improves. Thus it was verified that varying illumination is a powerful constraint, and furthermore, it can be effectively integrated with the other constraints.

	1	2	4	8	16
BF2	0.073 (0.001)	0.073 (0.001)	0.073 (0.001)	0.073 (0.001)	0.073 (0.001)
BF3	0.076 (0.001)	0.076 (0.001)	0.076 (0.001)	0.076 (0.001)	0.076 (0.001)
BD2	0.116 (0.002)	0.116 (0.002)	0.116 (0.002)	0.116 (0.002)	0.116 (0.002)
BD3	0.132 (0.002)	0.132 (0.002)	0.132 (0.002)	0.132 (0.002)	0.132 (0.002)
GW	1.62 (0.03)	1.01 (0.01)	0.69 (0.01)	0.513 (0.004)	0.428 (0.003)
RET	1.62 (0.03)	1.10 (0.01)	0.72 (0.01)	0.478 (0.004)	0.354 (0.003)
SMV2	21.9 (0.5)	8.4 (0.1)	3.43 (0.1)	1.37 (0.02)	0.62 (0.01)
SHA2	12.4 (0.3)	4.4 (0.1)	1.65 (0.05)	0.585 (0.01)	0.285 (0.003)
SIMV2	5.28 (0.05)	3.87 (0.02)	2.28 (0.05)	1.12 (0.02)	0.581 (0.005)
SIHA2	2.275 (0.2)	1.677 (0.02)	0.99 (0.02)	0.480 (0.01)	0.271 (0.003)
SIV1HA2	1.65 (0.02)	1.26 (0.01)	0.79 (0.01)	0.420 (0.01)	0.256 (0.002)
SIV2HA2	1.154 (0.01)	0.896 (0.01)	0.620 (0.008)	0.351 (0.004)	0.242 (0.002)
SIV3HA2	0.800 (0.01)	0.656 (0.008)	0.488 (0.006)	0.311 (0.004)	0.231 (0.001)
SIV4HA2	0.384 (0.002)	0.359 (0.002)	0.317 (0.002)	0.274 (0.002)	0.228 (0.002)

Table 4.1 Results of color constancy experiments for 1000 sets of 1, 2, 4, 8, and 16 surfaces under all combinations of test lights and extra lights for varying illumination. The canonical illuminant was a Philips CW fluorescent light. The values shown are the average magnitude of the chromaticity vector difference between the estimate and the desired answer, averaged over all results. The value in parenthesis is the error estimate of this average. The sample standard deviations vary from roughly 20 to 170 times this estimate (depending on the number of extra lights). The first part of this table is identical to Table 3.6.

4.3 Finding the Illumination Map

The previous sections have shown that varying illumination can be nicely integrated with the existing gamut mapping algorithms, provided that a varying-illumination map can be found. In this section a method to determine this map will be proposed. It assumes that an image can be segmented into regions representing different surfaces. The segmentation problem for arbitrary scenes is unsolved in general, and allowing varying illumination makes the problem even more difficult. Consider the simple K-means algorithm which could be used to bin sections based on color. If the illumination varies evenly, but significantly from blue to red across a white wall, then the blue pixels will be binned as blue, and the red ones as red. The distinction will be made on the basis of an arbitrary threshold. Thus this approach will not work.

The segmentation method used for this research is based on an assumption similar to that used by the Retinex algorithm to identify lightness changes. Regions are grown based on small, local changes, which are assumed to be due to illumination variation or noise. It was found to be more robust to threshold on both RGB and chromaticity differences than either one used alone. This method will run into problems in complex images, but it worked well in the case of simple Mondrian images, and served the desired purpose of allowing us to demonstrate that if the appropriate segmentation can be found, then the varying-illumination map can be robustly determined. The method used for this will now be explained.

Suppose that we have an image segmented on the basis of scene surfaces. The goal is to determine the diagonal transform which maps the illumination at the center to that at every other point. We reduce this to the problem of finding the illumination at the center of each region relative to that at the image center. Since the center of a region, as defined by the center of mass, need not be inside the region, the implementation used the point in the region closest to the center. The illumination at a point relative to that of the region center is simply the ratio of the response to the response of the center point. This follows directly from the assumption that the pixels are from the same surface, given that we accept a diagonal model for illumination change. Thus the map at an arbitrary point is simply the map at the center, adjusted by this relative jump.

To determine the maps at the center points we make the assumption that illumination does not change significantly at the region boundaries. Thus every jump across a boundary gives a condition on the relative maps of the centers of the two adjacent regions. More specifically, consider two regions A and B, with centers C_A and C_B, and boundary points B_A and B_B close to each other. Denote responses by R subscripted by the point label, and denote the diagonal map relative to the grand central point as D, also subscripted by the point label. Each channel or chromaticity component is dealt with independently, so the quantities in the equations are scalars. The assumption that the illumination does not change significantly at the boundary is simply:

$$D_{B_A} = D_{B_B} \tag{4.11}$$

Since we are assuming a diagonal model of illumination change, and C_A is

on the same surface as B_A, and similarly for the center and boundary of surface B, we have:

$$D_{B_A} = D_{C_A}\left(R_{B_A} \big/ R_{C_A}\right)$$
$$D_{B_B} = D_{C_B}\left(R_{B_B} \big/ R_{C_B}\right) \tag{4.12}$$

Combining (4.11) and (4.12) yields:

$$D_{C_A}\left(R_{B_A} \big/ R_{C_A}\right) = D_{C_B}\left(R_{B_B} \big/ R_{C_B}\right) \tag{4,13}$$

Taking logarithms of both sides of (4.13) gives:

$$\ln(D_{C_A}) + \ln(R_{B_A}) - \ln(R_{C_A}) = \ln(D_{C_B}) + \ln(R_{B_B}) - \ln(R_{C_B})$$

And moving terms reveals:

$$\ln(D_{C_A}) - \ln(D_{C_B}) = \ln(R_{B_B}) - \ln(R_{B_A}) + \ln(R_{C_A}) - \ln(R_{C_B}) \tag{4.14}$$

This final equation is at the heart of the method. Here we have conditions on the map component for two of the regions. Other boundary points produce additional equations. In order to have a robust method, one would like long boundaries to have more weight in the process than short ones, since the latter may due to a small region consisting entirely of noise. But this is exactly what we will get if we enter one equation for each boundary pair, and solve the resulting system of equations in the least squares sense. Furthermore, some boundary pairs can be identified as being of high quality, and these are weighted even more by scaling the equation by a number greater than one (typically five). In addition, some boundary pairs should contribute less, and

their equations are scaled by a number less than unity. These lesser quality pairs arise dealing with the pitfall described next.

In order to have a solution to the set of equations, it must be ensured that all segments connect to each other through the boundary pairs. This can be accomplished by simply assigning a region to every point, and using each break in both the horizontal and vertical directions to produce a boundary pair. This is usually not an option because normally some parts of the image should not be used. For example, the area may be too dark. Therefore connectedness was enforced in the following manner. Boundary pairs were assigned at each horizontal and vertical change of region. If one of the regions was to be ignored, a good region was sought in the same direction, taking as many pixels as required. The resulting equation was weighted inversely to the distance taken to find a good region. Thus such a boundary would contribute little to the solution, but connectivity was not a problem.[2]

As mentioned above, there is some variation in boundary pair quality. It was found to be better to use pixels one unit towards the insides of the respective regions, if these were available. This way the pixels would tend to have contributions that were solely due to a single surface, as opposed to the possibility that they sampled over more than one surface. These boundary pairs were weighted by a factor of five compared to ones where it was necessary to use pixels exactly on the boundary.

---

[2]Using this method, it is still possible to have an unconnected image. However, these are well defined and are not likely to occur in practice. Thus this possibility was ignored. For the image to still be unconnected, it would need a cross of unassigned pixels extending from side to side, and top to bottom, dividing it into four sections.

4.4 Putting it all Together

The methods of the preceding sections were tested on three real images. In all cases the scenes were illuminated by an incandescent desk light on the left, and simulated blue sky on the right. Thus the test covered a common real world situation—an office with a window. All scenes were chosen to produce images which were easy to segment, as the segmentation method used cannot handle complex images. Figure 4.1 shows the first scene illuminated as described above. Figure 4.2 shows the segmentation and all points that contribute to the boundary pairs (in red). The boundary areas are quite wide for two reasons. First 20% resolution was used for computation. Second, as described in §4.3, whenever possible boundary pairs were taken one unit inside the regions. Figure 4.3 shows the image with the illumination chromaticity variation removed, and Figure 4.4 shows the illumination map deduced. The input for color constancy processing is this image (for the surface constraints), and the illumination map (for the varying illumination constraints). Figure 4.5 shows the result of color constancy processing. The input is reproduced in the upper left corner. Unlike the images in §3.4, the upper right corner is the result of applying surface and illumination constraints. The lower left corner is the result of applying these constraints and the varying illumination constraints, and the lower right is the same scene under the canonical illuminant (Phillips CW fluorescent). Due to a fortuitous selection of colors, the varying illumination constraints are almost superfluous in this case, as both algorithms do well. It is important to note that even if the varying illumination constraints are not needed, the processing to remove the variation is still required. The result of blindly

applying the non-varying illumination algorithms to the image with varying illumination is shown in Figure 4.6. Figure 4.7 shows the mapping constraints for this image. It is clear that the varying illumination constraints are very strong here, and alone would be sufficient to give a good answer.

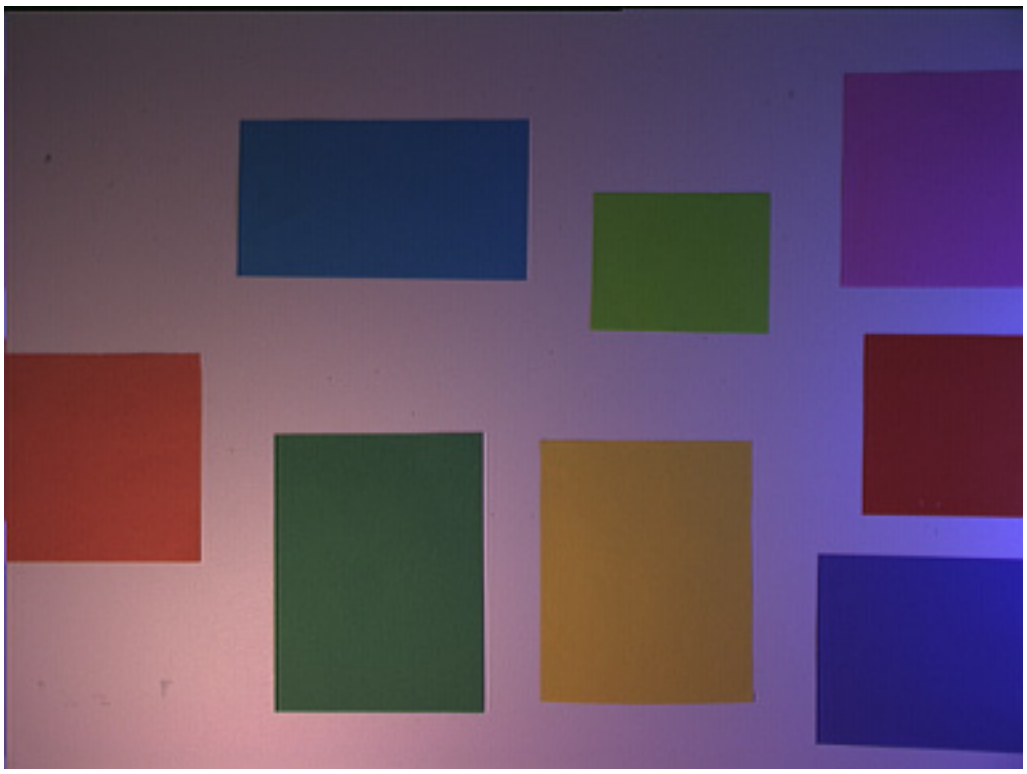

Figure 4.1 Image of a wall with colored paper illuminated on the left by incandescent light, and on the right by simulated daylight.

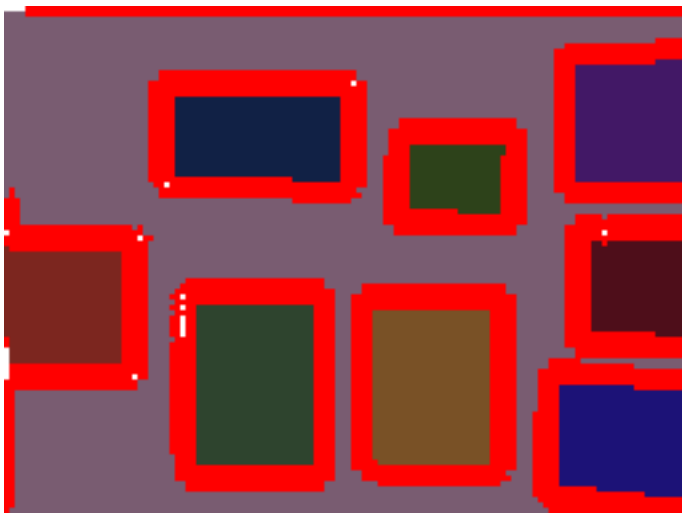

Figure 4.2 The results of segmenting the image shown in Figure 4.1, with all points contributing to the equations used to solve for the varying illumination map shown in red.

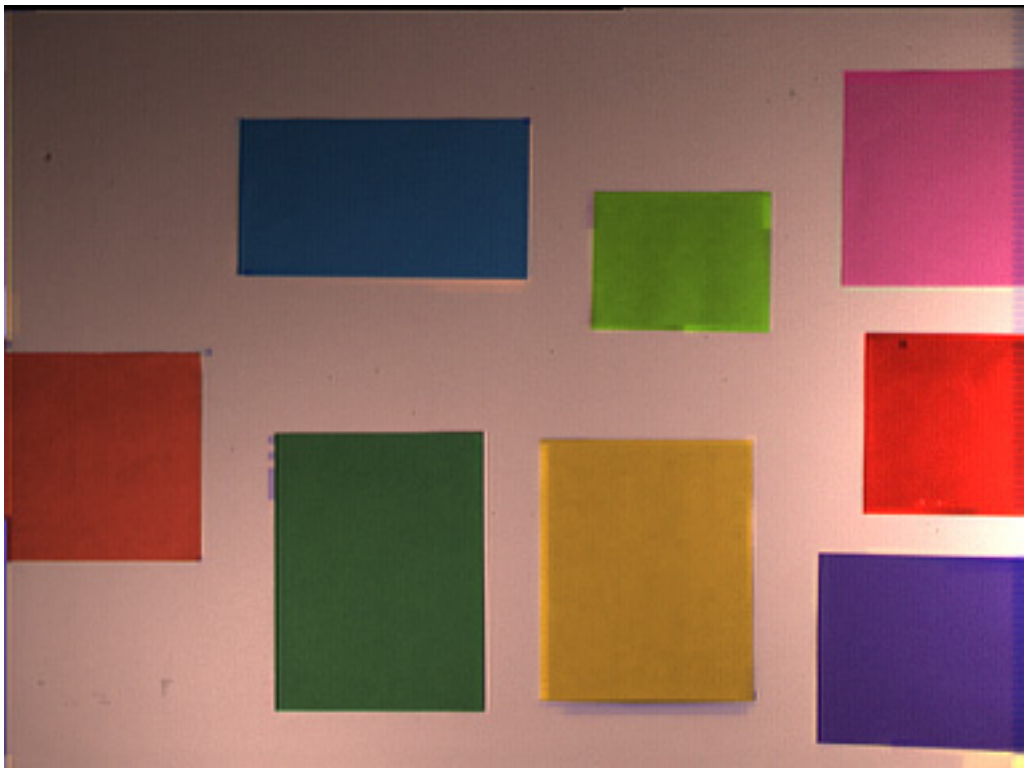

Figure 4.3 The results of removing the illumination chromaticity variation from the image shown in Figure 4.1. The image is adjusted so that the estimated chromaticity of the light is the same as that for the center of the original image.

Figure 4.4 The illumination chromaticity variation map for the image shown in Figure 4.1.

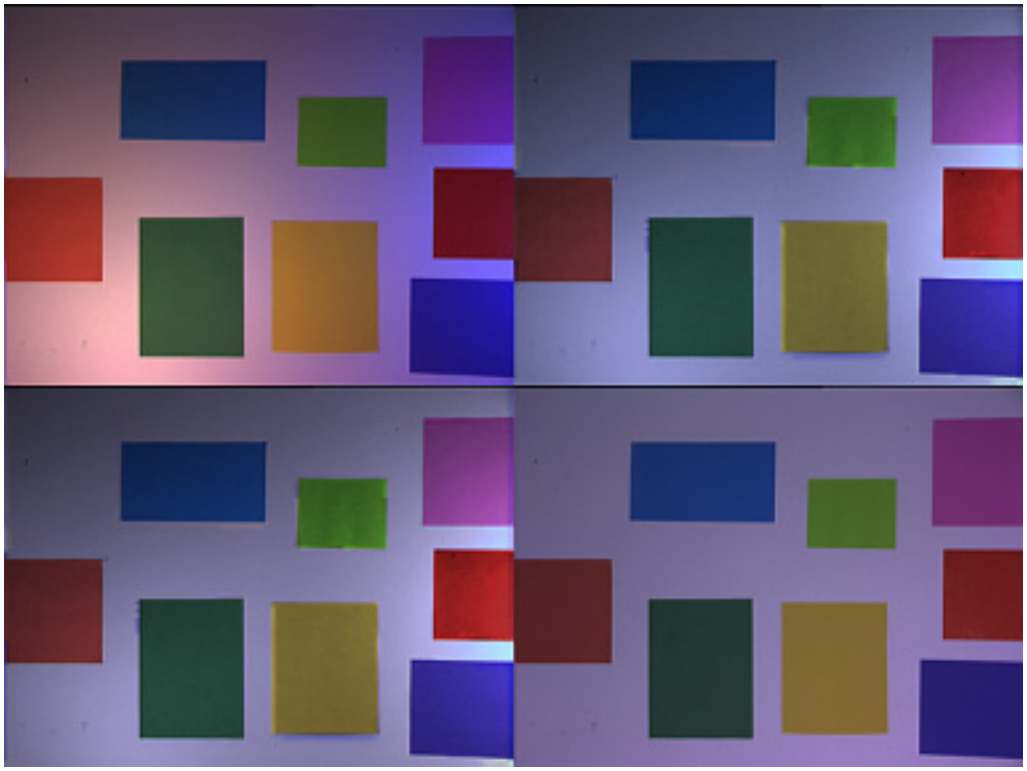

Figure 4.5 The result of color constancy processing for the image shown in Figure 4.1, and reproduced in the upper left. The upper right is the result of applying surface and illumination constraints to the image with illumination color variation removed (see Figure 4.3). The lower left is the result of including the varying illumination constraints. It is very much the same due to the fortuitous selection of paper colors. The desired colors are shown in the lower right, which is the same scene taken under the Philips CW fluorescent light used as the canonical. It should be pointed out that illumination intensity could be dealt with by the methods used for chromaticities, but for these experiments only chromaticity was corrected for. Hence the algorithm results show variation in intensity which is not present in the target image.

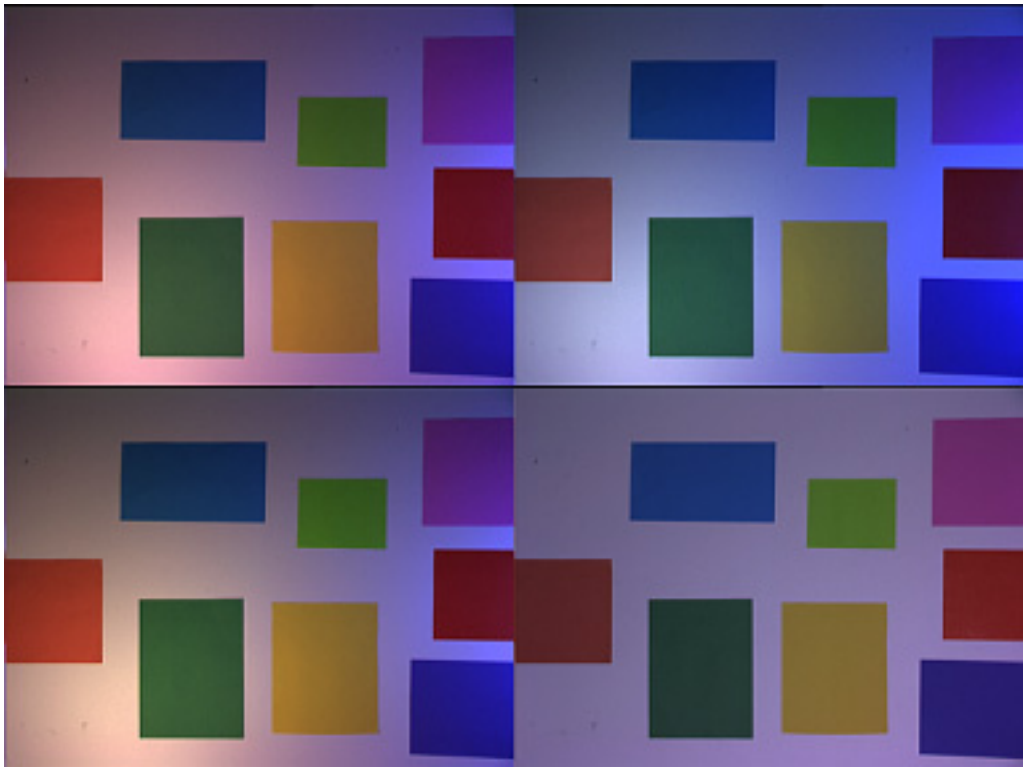

Figure 4.6 The result of applying the grey world algorithm and the chromaticity gamut mapping algorithm to the image shown in Figure 4.1, without any preprocessing to deal with the illumination variation. The upper left is the input, the upper right is the grey world result, the lower left is the result using the chromaticity gamut mapping algorithm, and the lower right shows the image of the scene under the canonical light. This result emphasizes that blindly applying algorithms which are not designed for varying illumination does not give good results. The chromaticity variation is maintained and thus any adjustment will, at best, be correct for only part of the scene. Furthermore, the adjustment itself is calculated from incorrect assumptions and can no longer be justified as being close to the best single map between the input and result (which is not even the desired answer in this case).

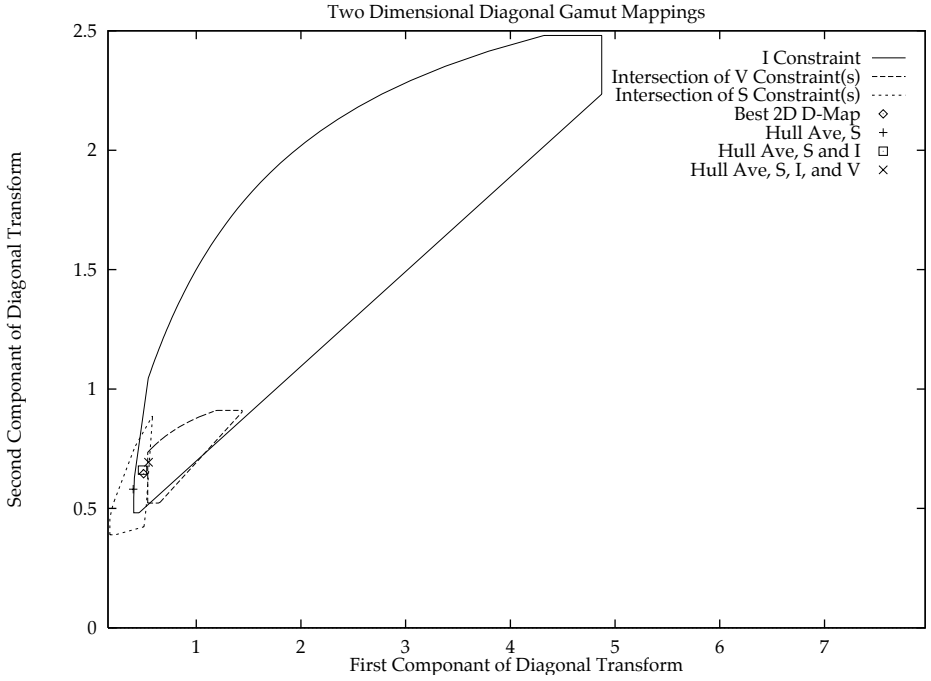

Figure 4.7 The constraints on the mappings to the canonical illuminant determined from the image with the illumination color removed, and including the varying illumination constraints. Note that the best fit (Best 2D D-Map) is for mapping the image with the illumination chromaticity removed to the image taken with the canonical light.

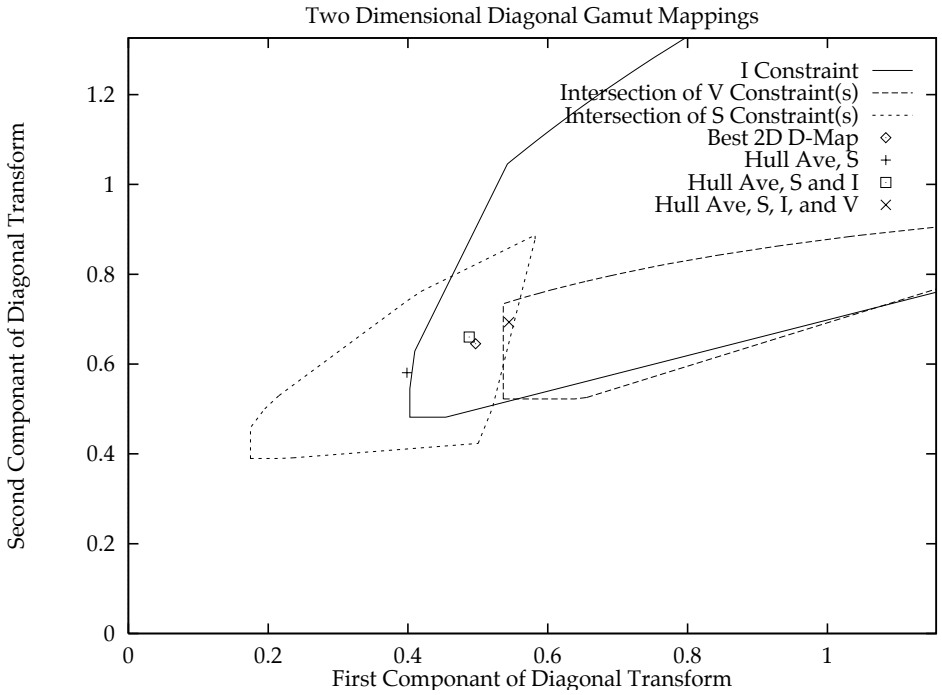

Figure 4.8 Figure 4.7 magnified to show the intersections in more detail.

The image analyzed in the preceding text had enough fortuitous colors that once the variation in the illumination chromaticity was removed, excellent color constancy was possible even without using the illumination variation. However, illumination variation can constrain the desired solution even when there are few colors in the scene. Figure 4.9 shows the results obtained by applying the algorithm to an image of a single green card taken under similar circumstances to the previous image. Figure 4.10 shows the constraints obtained, and it is clear that the varying illumination alone provides a significant restrictions on the possible mappings.

The varying illumination was also tested on the Mondrian image illuminated under similar conditions to the previous two images. The results are shown in Figure 4.11. Again, the results are good, and indicate that this is a very promising method.

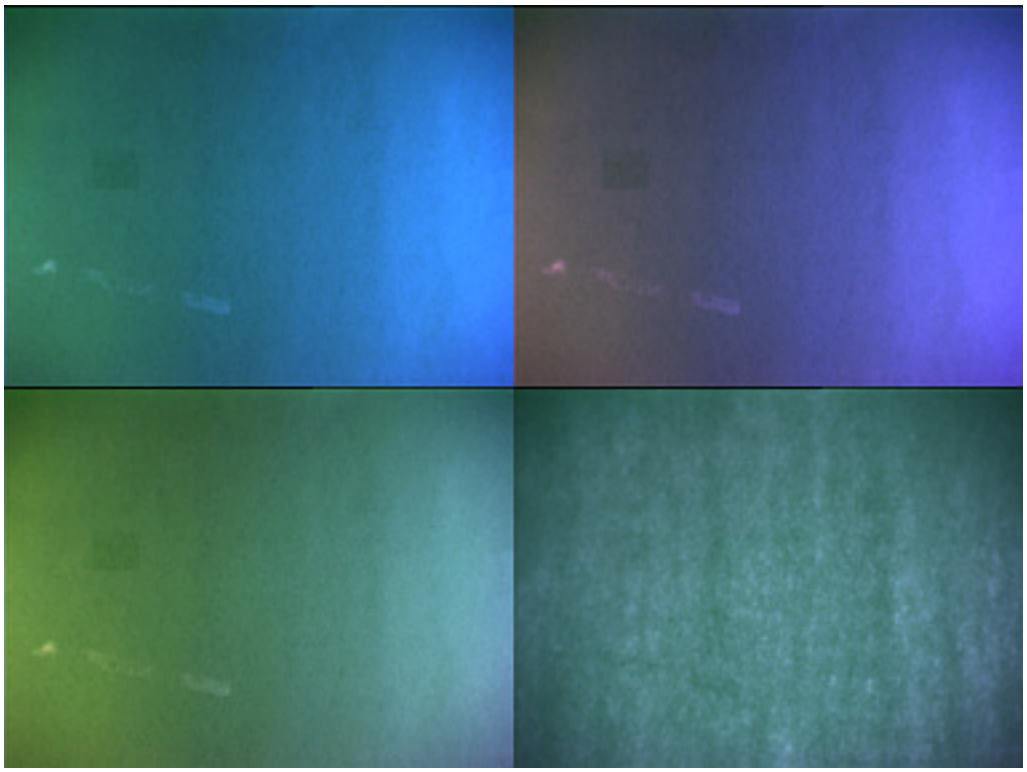

Figure 4.9 The results of the comprehensive algorithm applied to a single green card illuminated on the left by a regular incandescent light, and on the right by simulated blue sky. The upper left is the input. The upper right is the result of blindly using the grey world algorithm. The lower left is the result of using the varying illumination algorithm described in this chapter. The lower right is the image taken under the canonical illuminant (Phillips CW fluorescent).

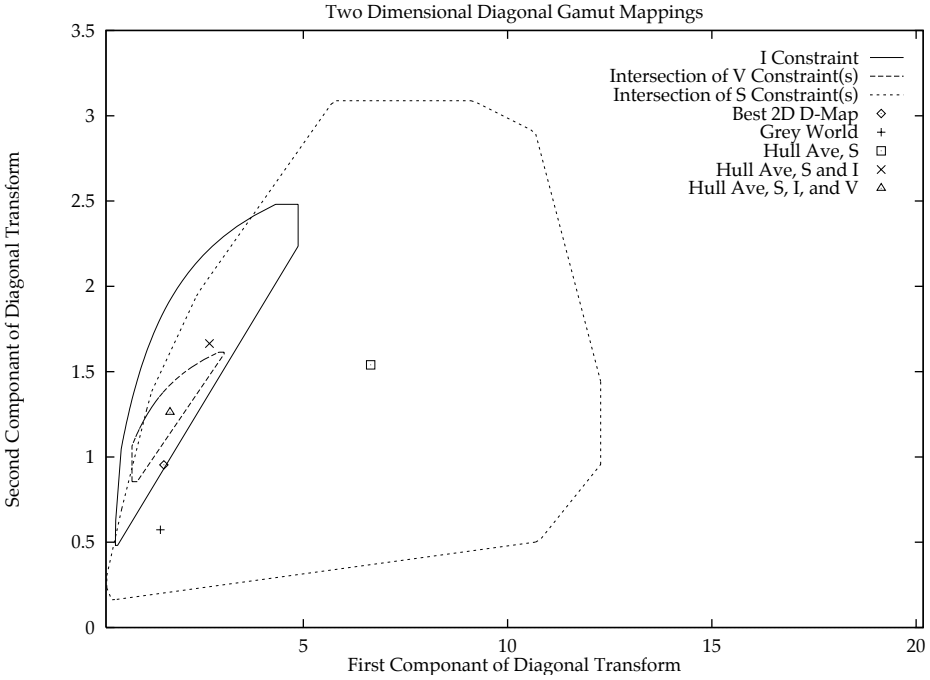

Figure 4.10 The constraints on the illumination mappings for the image shown in Figure 4.9. As one would expect, the single surface constraint does not help much. However, the varying illumination yields a reasonable solution. It should be noted that the best fit shown on the plot is the best fit between the image with the illumination chromaticity variation removed and the image under the canonical illuminant.

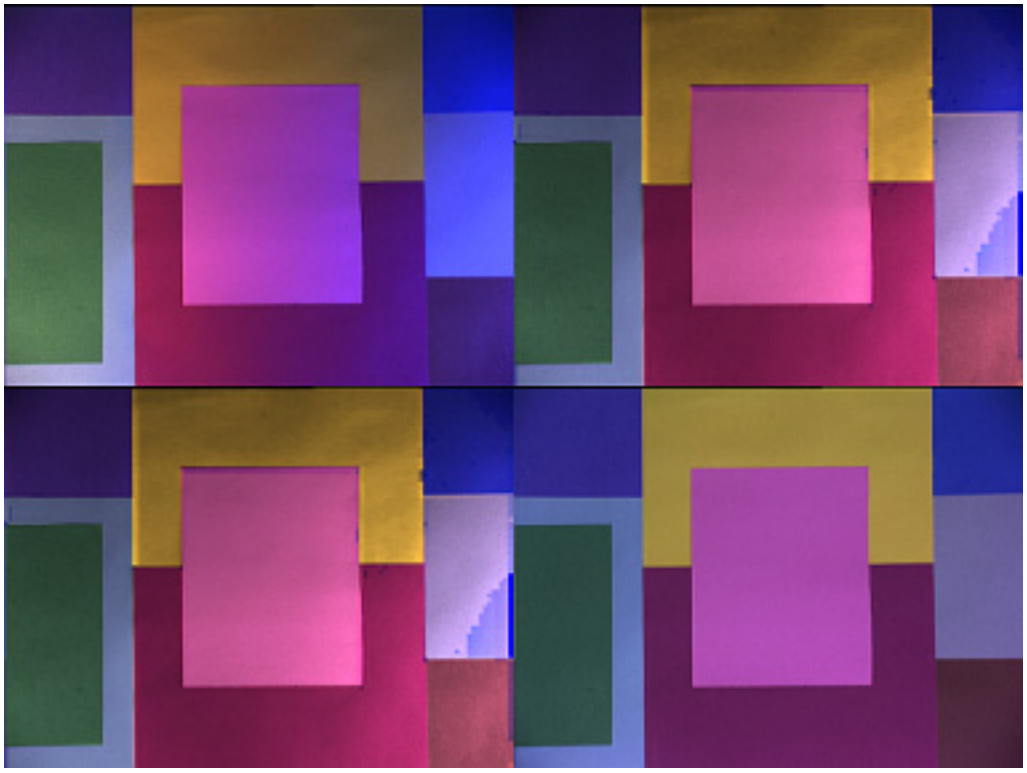

Figure 4.11 The results of the comprehensive algorithm applied to the Mondrian illuminated on the left by a regular incandescent light, and on the right by simulated blue sky. The upper left is the input. The upper right is the result of using the illumination and surface constraints on the image with the illumination chromaticity removed. The lower left is the result of using the illumination variation constraints as well. Again, in the case of sufficient color, it does not make much difference. The obviously incorrect region on the middle right of these two images is due to insufficient information to solve for the variation in illumination due saturation in the original image. This problem could be dealt with. The lower right is the image taken under the canonical illuminant (Phillips CW fluorescent).

Conclusion

The goal of this work was to investigate the application of color constancy algorithms to image data. This proved to be possible in the case of chromatically uniform illumination under arbitrary conditions. In addition, it was possible to extend the results of a recent algorithm for color constancy under varying illumination so that it worked very well in the case of simple images.

The journey began with a series of measurements to explore the nature of the lights, surfaces, and camera sensors which contribute to the images that need to be analyzed. It was found that published data was not sufficient to cover the surfaces in our laboratory. Specifically, success with the book scene presented in chapter 3 required adding measured data to the canonical gamut. An additional contribution was a simple implementation of a camera calibration technique based on that of Vhrel and Trussel [VT92].

Then some current ideas in color constancy were investigated. First, sensor sharpening was considered as a method for improving the performance of the algorithms. Here it proved difficult to identify a

sharpening transform for a wide range of illuminants which unambiguously improved the algorithms. The search was made difficult by the complicating factor that the negative sensor values that may be produced by sharpening can lead to problems with the gamut mapping algorithms. It seems that insisting on the best solution for all algorithms with all illuminants (as is the case with overall optimization proposed in §3.2) leads to poor performance for some of the algorithms under the test illuminants. Thus it is suggested that sensor sharpening must be further investigated with respect to specific algorithms in the context of a wider illumination set.

Fortunately the problems found with sharpening can be ignored because the camera sensors are already quite sharp (as verified in Table 3.1). Working with unmodified sensors, a number of color constancy algorithms were tested on generated data. It was found that the gamut mapping approach performed better than more naive methods, but it was necessary to include the illumination constraint proposed by Finlayson [Fin95]. Furthermore, using the hull average as suggested in this work, as opposed to the maximum-volume-mapping heuristic used previously, increased the number of cases where the gamut mapping algorithms performed better. Although the results were not completely unanimous, given its relative efficacy, combined with the esthetics of constraining rather than guessing a solution, the gamut mapping approach is the current method of choice. This is heavily supported by the results of the experiments on image data. In these experiments, the grey world algorithm was preferred only once out of 12 combinations of unknown illuminant and image scene.

In the final leg of the journey, a different challenge for color constancy research was confronted, namely dealing with scenes with varying illumination. First a very promising algorithm [FFB95] was modified so that it could be used as part of a comprehensive color constancy algorithm. Then experiments on generated data verified that all three classes of constraints, specifically those due to surfaces, illumination, and varying illumination, worked together to give better color constancy. At this point, we were still without a method for identifying the varying illumination, and thus a robust method for doing this was proposed. This method requires segmenting the image, which is a difficult problem especially if the illumination may vary. Nonetheless, simple images were successfully segmented by region growing using small jumps in both chromaticity and RGB as the condition for inclusion into a region. This segmentation method allowed the comprehensive color constancy algorithm to be tested on image data, and the results were excellent.

Thus to a reasonable extent, the original goal has been achieved. It is worth pointing out that no color constancy processing on real image data with a non-negligible amount of varying illumination has been reported in the literature. Along similar lines, even for the common assumption of uniform illumination, results in the literature for image data are very sparse. Certainly there are no published results for color constancy processing on images as general as our book scenes. Hence the work in this thesis significantly extends the quality and quantity of practical color constancy results.

Appendix A

Selecting Solutions by Centroid

In this work the preferred method to select solutions was to use the centroid of the solution set. Intuitively, this is a good choice, but it can also be justified formally without too much effort. To do this, we make two simple assumptions. First, we assume that all candidate solutions are equally likely. Second, we accept the definition of the error to be the vector magnitude difference between the estimate and the actual value.

The assumption that all candidate solutions are equally likely means that the probability density function, P(**X**), is a constant over the solution set:

$$P(\mathbf{X})=C \tag{A.1}$$

Although the value of C is not required, it is easily specified. Since we must have:

$$\int_{\substack{\text{Solution}\\\text{Set}}} P(\mathbf{X})\, dv = 1 \tag{A.2}$$

C is given by:

$$C = \frac{1}{\int_{\substack{\text{Solution}\\\text{Set}}} dv}$$

Now suppose that $\mathbf{X}^*$ is a proposed solution. Given that all solutions are equally likely, the expected value of the error squared is given by:

$$E^2 = \int_{\substack{\text{Solution}\\\text{Set}}} C \left\| \mathbf{X}^* - \mathbf{X} \right\|^2 dv \tag{A.3}$$

Working with the components we have:

$$E^2 = \int_{\substack{\text{Solution}\\\text{Set}}} C \sum_i \left(x_i - x_i^* \right)^2 dv \tag{A.4}$$

Now consider varying $\mathbf{X}^*$ to minimize the error. To minimize the error, we can minimize the error squared. To find the best value of $\mathbf{X}^*$ we set the partial derivative with respect to each component to zero. Using "*I*" to designate a specific choice of component:

$$\frac{\partial}{\partial x_I^*}\left(E^2\right) = 0 \tag{A.5}$$

$$\frac{\partial}{\partial x_I^*} \int_{\substack{\text{Solution}\\\text{Set}}} C \sum_i \left(x_i - x_i^* \right)^2 dv = 0 \tag{A.6}$$

$$\int_{\substack{\text{Solution}\\\text{Set}}} \left(x_I - x_I^* \right) dv = 0 \qquad \text{(Ignoring a factor of –2C)} \tag{A.7}$$

$$\int_{\substack{\text{Solution}\\\text{Set}}} x_I \, dv = \int_{\substack{\text{Solution}\\\text{Set}}} x_I^* \, dv \tag{A.8}$$

$$x_I^* \int_{\substack{\text{Solution}\\\text{Set}}} dv = \int_{\substack{\text{Solution}\\\text{Set}}} x_I \; dv \tag{A.9}$$

Finally, for each component x_I^* of the minimum error solution we get:

$$x_I^* = \frac{\int_{\substack{\text{Solution}\\\text{Set}}} x_I \; dv}{\int_{\substack{\text{Solution}\\\text{Set}}} dv} \tag{A.10}$$

This final equation defines the components of the centroid, and so we are done.

Bibliography

[BF94] D.H. Brainard and W.T. Freeman, "Bayesian method for recovering surface and illuminant properties from photosensor responses", *Human Vision, Visual Processing, and Digital Display V*, pp. 364-376, (1994)

[Buc80] G. Buchsbaum, "A spatial processor model for object colour perception", *Journal of the Franklin Institute*, 310, pp. 1-26, (1980)

[BW81] M.H. Brill and G. West, "Contributions to the Theory of Invariance of Color Under the Condition of Varying Illumination", *Journal of Mathematical Biology*, 11, pp, 337-350, 1981

[Coh64] J. Cohen, "Dependency of The Spectral Reflectance Curves of The Munsell Color Chips", *Psychon. Sci.*, 1, pp. 369-370, (1964)

[Dix78] E.R. Dixon, "Spectral distribution of Australian daylight", *Journal of the Optical Society of America*, 68, pp. 437-450, (1978)

[FB95] W. Freeman and David Brainard, “Baysian Desion Theory, the Maximum Local Mass Estimate, and Color Constancy”, in *Proceedings: Fifth International Conference on Computer Vision*, pp 210-217, (IEEE Computer Society Press, 1995)

[FDB] B. V. Funt, M. S. Drew, M. Brockington, “Recovering Shading from Color Images”, in *Proceedings: Second European Conference on Computer Vision*, G.Sandini, ed., 1992.

[FDF94a] G.D. Finlayson and M.S. Drew and B.V. Funt, "Spectral Sharpening: Sensor Transformations for Improved Color Constancy", *Journal of the Optical Society of America A*, 11, 5, pp. 1553-1563, (1994)

[FDF94b] G.D. Finlayson and M.S. Drew and B.V. Funt, "Color Constancy: Generalized Diagonal Transforms Suffice", *Journal of the Optical Society of America A*, 11, pp. 3011-3020, (1994)

[FF94] G.D. Finlayson and B.V. Funt, "Color Constancy with Shadows", *Perception*, 23, (Special issue on the 17th European Conference on Visual Perception, Eindhoven), pp. 89-90, (1994)

[FFB95] G. D. Finlayson, B. V. Funt, and K. Barnard, “Color Constancy Under Varying Illumination”, in *Proceedings: Fifth International Conference on Computer Vision*, pp 720-725, 1995.

[Fin94] G.D. Finlayson, "Color Constancy and a Changing Illumination", In *Human Vision, Visual Processing, and Digital Display V*, pp. 353-363, (1994)

[Fin95] G. D. Finlayson, “Color Constancy in Diagonal Chromaticity Space”, in *Proceedings: Fifth International Conference on Computer Vision*, pp 218-223, (IEEE Computer Society Press, 1995).

[For90] D. Forsyth, "A novel algorithm for color constancy", *Int. J. Computer. Vision*, 5, pp. 5-36, (1990)

[GJT86] R. Gershon and A.D. Jepson and J.K. Tsotsos, "Ambient illumination and the determination of material changes", *Journal of the Optical Society of America A*, 3, pp. 1700-1707, (1986)

[GW87] R. C. Gonzalez and P. Wintz, *Digital Image Processing*, Second Edition, (Addison-Wesley 1987)

[Hor74] B.K.P. Horn, "Determining lightness from an image, *Computer Vision, Graphics, and Image Processing*, 3, pp. 277-299, (1974)

[JMW64] D.B. Judd and D.L. MacAdam and G. Wyszecki, "Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature", *Journal of the Optical Society of America*, 54, pp. 1031-1040, (August 1964)

[Kri47] E.L. Krinov, "Spectral Reflectance Properties of Natural Formations", Technical Translation TT-439, National Research Council of Canada, (1947)

[Lan77] E.H. Land, "The Retinex theory of Color Vision", *Scientific American,* 108-129, (1977)

[Lan86] E.H. Land, "Recent advances in Retinex theory", *Vision Res.*, 26, pp. 7-21, (1986)

[LM71] E. H. Land and J. J. McCann, “Lightness and Retinex Theory”, *Journal of the Optical Society of America*, 61, 1, pp. 1-11, (1971).

[MMT76] John J. McCann, Suzanne P. McKee, and Thomas H. Taylor, “Quantitative Studies in Retinex Theory”, *Vision Research*, 16, pp. 445–458, (1976)

[MW86] L.T. Maloney and B.A. Wandell, "Color constancy: a method for recovering surface spectral reflectance", *Journal of the Optical Society of America A,* 3, pp. 29-33, (1986)

[MW92] D.H. Marimont and B.A. Wandell, "Linear models of surface and illuminant spectra", *Journal of the Optical Society of America A,* 9, 11, pp. 1905-1913, (1992)

[Nic57] D. Nickerson, *Spectrophotometric Data for a Collection of Munsell Samples*, U.S. Department of Agriculture. Wahsington D.C., (1957)

[PHJ89] J. P. S. Parkkinen and J. Hallikanen and T. Jaaskelainen, "Characteristic spectra of Munsell Colors", *Journal of the Optical Society of America A*, 6, pp. 318-322, (1989)

[Sha85] S.A. Shafer, "Using color to separate reflection components", *Color Res. Appl.*, 10, pp. 210-218, (1985)

[ST93] G. Sharma and H.J. Trussel, "Characterization of Scanner Sensitivity", In *IS&T and SID's Color Imaging Conference: Transforms & Transportability of Color*, pp. 103-107, (1993)

[TO90] M. Tsukada and Y. Ohta, "An Approach to Color Constancy Using Multiple Images", in *Proceedings Third International Conference on Computer Vision*, (IEEE Computer Society, 1990)

[Tom94] S. Tominaga, "Realization of Color Constancy Using the Dichromatic Reflection Model", in *The second IS&T and SID's Color Imaging Conference*, pp. 37-40, (1994)

[VT93] M.J. Vrhel and H.J. Trussel, "Physical Device Illumination Correction", in *Device-Independent Color Imaging and Imaging Systems Integration*, Vol. 1909, pp. 84-91, (1993),

[Wan87] B.A. Wandell, "The synthesis and analysis of color images", *IEEE Trans. Patt. Anal. and Mach. Intell.*, 9, pp. 2-13, (1987)

[WB82] G. West and M.H. Brill, "Necessary and sufficient conditions for von Kries chromatic adaption to give colour constancy", *J. Math. Biol.*, 15, pp. 249-258, (1982)

[WS82] G. Wyszecki and W.S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, 2nd edition, (Wiley, New York, 1982)

[ZI93] M. D'Zmura and G. Iverson, "Color constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces", *Journal of the Optical Society of America A*, 10, pp. 2148-2165, (1993)