Diagonal Transforms Suffice for Color Constancy

Graham D. Finlayson, Mark S. Drew and Brian V. Funt
Simon Fraser University
Vancouver, B.C.
Canada V5A 1S6
e-mail graham@cs.sfu.ca

Abstract

This paper's main result is to show that under the conditions imposed by the Maloney-Wandell color constancy algorithm, color constancy can in fact be expressed in terms of a simple independent adjustment of the sensor responses—in other words as a von Kries adaptation type of coefficient rule algorithm—so long as the sensor space is first transformed to a new basis. Our overall goal is to present a theoretical analysis connecting many established theories of color constancy. For the case where surface reflectances are 2-dimensional and illuminants are 3-dimensional, we prove that perfect colour constancy can always be solved for by an independent adjustment of sensor responses, which means that the colour constancy transform can be expressed as a diagonal matrix. This result requires a prior transformation of the sensor basis and to support it we show in particular that there exists a transformation of the original sensor basis under which the non-diagonal methods of Maloney-Wandell, Forsyth's MWEXT and Funt and Drew's lightness algorithm all reduce to simpler, diagonal-matrix theories of colour constancy. Our results are strong in the sense that no constraint is placed on the initial sensor spectral sensitivities. In addition to purely theoretical arguments, the paper contains results from simulations of diagonal-matrix-based color constancy in which the spectra of real illuminants and reflectances along with the human cone sensitivity functions are used. The simulations demonstrate that when the cone sensor space is transformed to its new basis in the appropriate manner, a diagonal matrix supports close to optimal colour constancy.

Keywords: Color, color constancy, computer vision, Maloney–Wandell, von Kries adaptation, coefficient rule, Finite–Dimensional Models

1 Introduction

We present a theoretical analysis connecting several well-known color constancy theories—von Kries adaptation[32], Land's retinex[20], the Maloney-Wandell algorithm[23], Funt and Drew's lightness algorithm[10], Forsyth's MWEXT and CRULE[8]—in which we prove that under an appropriate change of basis for the sensor space, for illuminants and reflectances that are well-approximated by low-dimensional finite-dimensional models every one of these methods results in a simple *independent* adjustment of coefficients in this new space.

Color has a long history in machine vision and has proved to be of great use in object recognition [29, 12], image segmentation [16, 1] and a host of other visual tasks [27, 11, 30, 5]. For many of these algorithms, their performance degrades to the extent to which colors change with changing illumination. We need colors to be stable descriptors for object surface properties—the same object seen under a blue sky or under a tungsten lamp should have approximately the same perceived color. The ability to discount the effect of the illuminant is called *color constancy.*

A color camera, like the human eye, has 3 color sensors; hence in a color image each pixel is a 3-vector, one component per sensor channel. A color constancy algorithm maps each color vector $\underline{p}$ to a descriptor vector $\underline{d}$ which is independent of the illuminant. This mapping is usually considered linear—a matrix transform is applied to color vectors. Indeed, under Forsyth's formulation [8] of the color constancy problem, the transform *must* be linear. In this paper we provide a theoretical analysis along with simulation results demonstrating that if the transform is linear, then it need only be diagonal. In other words, a diagonal matrix transform suffices as a vehicle for color constancy. Our results are strong in the sense that they place no constraints on the spectral sensitivities of the visual system.

Competing computational schemes for simulating color constancy apply different structural constraints to the form of the matrix transform. Many authors assume that the transform is a diagonal matrix, and in the model of Maloney and Wandell [23] the transform is a 2×3 projection. Only Forsyth's MWEXT [8] algorithm places no constraints on the form of the transform. In studying color constancy algorithms, therefore, we must ask two questions:

1. Independent of the computational scheme for computing the matrix, how well in principle can a particular matrix form discount the effect of the illuminant?

2. How successful is a given color constancy algorithm in solving for the correct (or best) transform?

Our main focus in this paper is on the first of these questions. In answer to it we show that for cases in which a finite-dimensional linear model of sufficiently low dimension captures the shapes of both the spectral power distribution for illumination and the spectral reflectance function, then the form of the matrix required in order to support perfect color constancy is just diagonal. Since the off-diagonal terms are all zero, this is equivalent to the simple application of a coefficient rule[8, 32].

From the von Kries adaptation model (see [33]) through Land's retinex scheme [20] to Forsyth's recent CRULE theory, the diagonal matrix transform has long been proposed as a viable mechanism for color constancy. Unfortunately many authors [32] have graphically demonstrated that, except for the case of narrow-band sensors, a diagonal matrix often performs poorly in discounting illumination change. Consequently the majority of recent color constancy theories discard the computational simplicity of the diagonal matrix transform for more complex matrix forms which *supposedly* can model illuminant change better.

In contrast to this trend Finlayson et al. [7] have recently proved that diagonal matrix transforms can support perfect color constancy under weak model constraints—the illuminant space linearly spanned by a *2*-dimensional basis and the reflectance space by a *3*-dimensional basis. We term this set of constraints a *2-3* model. That analysis employs a generalization, which we will use here, of the concept of a diagonal matrix transform in which a *sensor transformation* $\mathcal{T}$ is allowed prior to the application of a diagonal matrix:

$$\underline{d} = \mathcal{D}\underline{p} \qquad \text{(1 simple diagonal matrix constancy)}$$

$$\mathcal{T}\underline{d} = \mathcal{D}\mathcal{T}\underline{p} \qquad \text{(2 generalized diagonal matrix constancy)}$$

In the *2-3* case, given a known reference patch in each scene, the correct diagonal matrix transform can be computed to yield perfect color constancy. The elegant color constancy algorithm of Maloney et. al. does not require a reference patch, but it operates under a different set of restrictions. These restrictions, which we will call the *3-2* restrictions, require a 3-dimensional illuminant space and a 2-dimensional reflectance space.

The main result of this paper is to show that, in a world in which illuminants and reflectances are governed by Maloney's *3-2* restrictions, color constancy can *always be formulated as a diagonal matrix transform independent of the spectral characteristics of the sensors.* In a world in which these restrictions hold only approximately, a diagonal matrix transform theory of color constancy will still do a good job.

The ramifications of this result for theories of color constancy are widespread. The most immediate implication is that the *3-2* version of Maloney's theory of color constancy is a diagonal-matrix-based theory of color constancy. Finite-dimensional restrictions are also at the foundation of Funt and Drew's [10] color constancy algorithm. Their computational method simplifies, via our analysis, to diagonal matrix operations in the *3-2* case and as such reduces to Blake's version of the Lightness algorithm [2]. Finally, our work plays a unifying role in connecting the theories of Maloney and Forsyth.

Forsyth's work on color constancy consists of two algorithms: MWEXT and the simpler CRULE. In MWEXT, color constancy proceeds by parameterizing all the possible matrices mapping the gamut of image colors into the gamut of descriptors. The more colorful the image, the smaller the set of possible mappings becomes. Unfortunately this algorithm is extraordinarily complex and, as Forsyth suggests, may not be suitable for machine vision. Restricting color constancy transforms to diagonal matrices results in Forsyth's simpler CRULE algorithm. This algorithm can be efficiently implemented and is a suitable candidate for a machine vision implementation of color constancy. Our results prove that Maloney's theory of color constancy is a sub–theory of the computationally tractable CRULE.

Our work also has application in the allied problem of *color balancing*. Video cameras cannot account for changing illumination. Consequently, images taken under different illuminants must be *balanced* before display to a human observer. This balancing usually takes the form of a simple scaling in each color channel—the color video image is transformed by a diagonal matrix. To ensure illumination change is successfully corrected, video cameras are normally equipped with narrow band sensors. The results in this paper indicate that a diagonal matrix transform is a suitable balancing technique independent of the sensor sensitivities used—broad-band sensors are as suitable a choice as narrow-band sensors.

In section 2 we provide the necessary definitions required to develop a mathematical model for color image formation and color constancy. In section 3 we develop techniques for finding the sensor transform $\mathcal{T}$ which affords perfect diagonal matrix color constancy under *3-2* restrictions. It should be noted that this analysis does not place restrictions on the possible form of the initial set of sensors. In section 4 we formally connect our results with other computational theories of color constancy. Finally in section 5 we present simulation results which evaluate the performance of generalized diagonal matrix color constancy. The appendix discusses the role of complex numbers in theories of color constancy.

2 The Model

The light reflected from a surface depends not only on the spectral properties of illumination and surface reflectance, but also on other confounding factors such as specularities and mutual illumination. To simplify our analysis we will, in line with many other authors, develop our theory for the simplified Mondrian world; a Mondrian is a planar surface composed of several, overlapping, matte (Lambertian) patches. We assume that the light striking the Mondrian is of uniform intensity and is spectrally unchanging. In this world the only factor confounding the retrieval of surface descriptors is illumination.

Light reflected from a Mondrian falls onto a planar array of sensors and at each location X in the sensor array there are three different classes of sensors. The value registered by the kth sensor, p_k^X (a scalar), is equal to the integral of its response function multiplied by the incoming color signal. For convenience, we arrange the index X such that each p_k^X corresponds to a unique surface reflectance:

$$p_k^X = \int_\omega C^X(\lambda)\, R_k(\lambda)\, d\lambda \qquad \text{(3 Color appearance)}$$

where λ is wavelength, $R_k(\lambda)$ is the response function of the kth sensor, $C^X(\lambda)$ is the color signal at X and the integral is taken over the visible spectrum ω. The color signal is the product of a single surface reflectance $S(\lambda)$ multiplied by the ambient illumination $E(\lambda)$: $C(\lambda) = E(\lambda)S(\lambda)$. Henceforth we drop the index X.

2.1 Finite-Dimensional Models

Illuminant spectral power distribution functions and surface spectral reflectance functions are well described by finite-dimensional models. A surface reflectance vector $S(\lambda)$ can be approximated as:

$$S(\lambda) \approx \sum_{i=1}^{d_S} S_i(\lambda)\sigma_i \tag{4}$$

where $S_i(\lambda)$ is a basis function and $\underline{\sigma}$ is a d_S-component column vector of weights. Maloney [21] presents evidence which suggests surface reflectances can be well modelled by a set of between 3 and 6 basis vectors. Similarly we can model illuminants with a low–dimension basis set:

$$E(\lambda) \approx \sum_{j=1}^{d_E} E_j(\lambda)\epsilon_j \tag{5}$$

$E_j(\lambda)$ is a basis function and $\underline{\epsilon}$ is a d_E dimensional vector of weights. Judd [17] measured 605 daylight illuminants and showed they are well modelled by a set of 3 basis functions.

Basis functions are generally chosen by performing a principal component analysis of each data set (reflectances and illuminants) in isolation [4, 25, 22]. This type of analysis is weak in the sense that it does not take into account how illuminant, reflectance and sensor interact in forming a color vector (eqn. (3)). Recently Marimont and Wandell [24] developed a method for deriving reflectance and illuminant basis functions which best model *color appearance*—Eqn. (3) is the foundation for their method. They conclude that a 2–dimensional basis set for surface reflectance and a 3–dimensional basis set for illumination is sufficient to model the appearance of the 462 Munsell chips [25] under a wide range of black-body radiator illuminants. This is precisely the *3-2* case.

2.2 Lighting and Surface Matrices

Given finite-dimensional approximations to surface reflectance, the color appearance eqn. (3) can be rewritten as a matrix transform. A *lighting matrix* $\Lambda(\underline{\epsilon})$ maps reflectances, defined by the $\underline{\sigma}$ vector, onto a corresponding color vector:

$$\underline{p} = \Lambda(\underline{\epsilon})\underline{\sigma} \tag{6}$$

where $\Lambda(\underline{\epsilon})_{ij} = \int_\omega R_i(\lambda)E(\lambda)S_j(\lambda)d\lambda$. The lighting matrix is dependent on the illuminant weighting vector $\underline{\epsilon}$, with $E(\lambda)$ given by eqn. (5). The roles of illumination and reflectance are symmetric; we can rewrite the color appearance matrix of eqn. (6) as a *surface matrix* transforming an epsilon vector:

$$\underline{p} \;=\; \Omega(\underline{\sigma})\underline{\epsilon} \tag{7}$$

where $\Omega(\underline{\sigma})_{ij} = \int_\omega R_i(\lambda)E_j(\lambda)S(\lambda)d\lambda$, with $S(\lambda)$ defined in eqn. (4). This symmetry is a key part of the analysis presented in section 3.

2.3 The Color Constancy Problem

The aim of any color constancy algorithm is to transform the color vector $\underline{p}$ to its corresponding illuminant independent descriptor $\underline{d}$.

$$\underline{d} \;=\; \mathcal{Q}\underline{p} \tag{8}$$

where $\mathcal{Q}$ is a linear transform. However, there is no consistent definition for a descriptor. For example Maloney [23] uses the surface weight vector $\underline{\sigma}$ for the descriptor (eqn. (9)); in contrast Forsyth defines a descriptor to be the appearance of a surface seen under a canonical illuminant, defined by the weight vector $\underline{c}$ (eqn. (10)).

$$\underline{d}^M \;=\; [\Lambda(\underline{\epsilon})]^{-1}\Lambda(\underline{\epsilon})\underline{\sigma} \qquad \text{(9 Maloney's descriptor)}$$

$$\underline{d}^F \;=\; \Lambda(\underline{c})[\Lambda(\underline{\epsilon})]^{-1}\Lambda(\underline{\epsilon})\underline{\sigma} \qquad \text{(10 Forsyth's descriptor)}$$

Because each color constancy algorithm applies a linear transform to color vectors, different descriptor definitions differ only by a fixed linear transform, for example $\underline{d}^F = \Lambda(\underline{c})\underline{d}^M$. Therefore, demonstrating the adequacy of a diagonal matrix for one descriptor form demonstrates its adequacy for color constancy in general. In the analysis of section 3 we use Forsyth's descriptor form.

2.4 Illuminant Invariance

Color constancy seeks illuminant invariant color descriptors. A closely related problem is to find illuminant invariant *relationships* between color vectors instead. One candidate relationship, which we will call *diagonal invariance*, is the diagonal matrix mapping between the color vectors of the two surfaces:

$$\mathcal{D}^{ij}\underline{p}^{i,x} \;=\; \underline{p}^{j,x} \;\; . \tag{11}$$

Here i and j index two different surface reflectances, x refers to a particular (single) illuminant indexed by x, and $\mathcal{D}^{ij}$ is the diagonal invariant matrix. [1] [2]

[1]Note that $\mathcal{D}^{ij}$ means the entire 3×3 diagonal matrix relating $\underline{p}^{i,x}$ and $\underline{p}^{j,x}$, not the ij component of a matrix $\mathcal{D}$.

[2]Diagonal invariance is sometimes referred to as *ratio invariance*, because the diagonal elements of $\mathcal{D}^{ij}$ equal the ratios of the components of $\underline{p}^{j,x}$ over $\underline{p}^{i,x}$.

In this diagonal model, for *all* illuminants x, $\mathcal{D}^{ij}$ maps the color vector for surface i to the color vector for surface j. Diagonal invariance plays a key role in the lightness computations of Horn [14] and Blake [2], the image segmentation work of Hurlbert [15] and in the object recognition work of Funt and Finlayson [12]. Brill [3] develops a more general theory of illuminant invariance, where the relationship between surfaces can be a general linear transform.

3 Diagonal Matrix Transform and the *3-2* Case

Finlayson et al. [7] proved that assuming illumination is 2-dimensional and reflectance 3-dimensional (the *2-3* case), there exists a transformed sensor basis in which a diagonal matrix supports perfect color constancy. In this section we prove the equivalent result for the *3-2* case.

Theorem 1 *If illumination is 3-dimensional and surface reflectance 2-dimensional then there exists a sensor transform* $\mathcal{T}$ *for which a diagonal matrix supports perfect color constancy.*

We prove Theorem 1 in two stages. First we demonstrate a symmetry between diagonal invariance and diagonal matrix color constancy. Then we prove the existence of a sensor transform which supports diagonal invariance.

Lemma 1 *A diagonal matrix supports perfect color constancy if and only if there is diagonal invariance.*

Proof. When a diagonal matrix supports perfect color constancy, illumination change is exactly modelled by a diagonal matrix.

$$\underline{p}^{i,c} = \mathcal{D}^{ec}\underline{p}^{i,e} \; ; \; \underline{p}^{j,c} = \mathcal{D}^{ec}\underline{p}^{j,e} \tag{12}$$

where i, j index surface reflectance and the diagonal matrix $\mathcal{D}^{ec}$ maps the appearance of surfaces under an arbitrary illuminant e to their appearance with respect to the canonical illuminant c. Clearly we can map $\underline{p}^{i,e}$ to $\underline{p}^{j,e}$ by applying a diagonal matrix.

$$\underline{p}^{i,e} = \mathcal{D}^{ij}\underline{p}^{j,e} \tag{13}$$

Applying the color constancy transform $\mathcal{D}^{ec}$ to both sides of equation (13) we see that:

$$\mathcal{D}^{ec}\underline{p}^{i,e} = \mathcal{D}^{ec}\mathcal{D}^{ij}\underline{p}^{j,e} \tag{14}$$

Because transformation by diagonal matrices is commutative we can rewrite equation (14) as

$$\mathcal{D}^{ec}\underline{p}^{i,e} = \mathcal{D}^{ij}\mathcal{D}^{ec}\underline{p}^{j,e} \tag{15}$$

Substituting equations (12) into equation (15) we see that

$$\underline{p}^{i,c} = \mathcal{D}^{ij}\underline{p}^{j,c} \tag{16}$$

Equation (16) is a statement of diagonal invariance. The above argument is clearly symmetric—given diagonal invariance, diagonal matrix color constancy must follow. For the proof, we need only change the meaning of the superscripts in equations (12)-(16) so the first indexes the illuminant and the second reflectance ($\mathcal{D}^{ec}$ becomes a diagonal invariant and $\mathcal{D}^{ij}$ a color constancy transform).

Lemma 2 *Given 3-2 restrictions, there exists a transformation of the sensor response functions for which, independent of the illuminant, color vectors are diagonally invariant.*

Proof. Under the *3-2* restrictions the appearance of a reflectance $\underline{\sigma}$ under an illuminant $\underline{\epsilon}$ can be written in terms of two surface matrices. To see this, first note that matrix $\Omega(\underline{\sigma})$ in eqn. (7) can be decomposed into two parts. If the 2–vector $\underline{\sigma}$ has components $(\sigma_1, \sigma_2)^T$, then defining two special Ω matrices associated with the two basis directions in σ–space,

$$\Omega(1) \leftrightarrow (1,0)^T, \;\; \Omega(2) \leftrightarrow (0,1)^T ,$$

we have

$$\Omega(\underline{\sigma}) = \sigma_1\Omega(1) + \sigma_2\Omega(2) .$$

Therefore eqn. (7) becomes

$$\underline{p} = \sigma_1\Omega(1)\underline{\epsilon} + \sigma_2\Omega(2)\underline{\epsilon} \tag{17}$$

Let us define a *canonical* surface reflectance, $\underline{s}$, and examine its relationship to the color appearance of other surfaces. Without loss of generality we choose the first surface basis function as the canonical surface. The appearance of the second surface basis function is an illuminant-independent, linear transform of the canonical surface appearance:

$$\Omega(2)\underline{\epsilon} = \mathcal{M}\Omega(1)\underline{\epsilon} \tag{18}$$

$$\mathcal{M} = \Omega(2)[\Omega(1)]^{-1} \tag{19}$$

Now we can rewrite eqn. (17), the general appearance of arbitrary surfaces, as a fixed transform from the canonical surface appearance.

$$\underline{p} = [\sigma_1\mathcal{I} + \sigma_2\mathcal{M}]\Omega(1)\underline{\epsilon} \tag{20}$$

where $\mathcal{I}$ is the identity matrix. Therefore we have shown that the appearance of the canonical surface can be mapped to the appearance of any other surface reflectance by applying a linear combination of the identity matrix $\mathcal{I}$ and the matrix $\mathcal{M}$. We define a *generalized diagonal transform* as a basis transformation followed by a diagonal matrix transform. That there exists a generalized diagonal transform mapping canonical surface appearance follows from the eigenvector decomposition of $\mathcal{M}$:

$$\mathcal{M} = \mathcal{T}^{-1}\mathcal{D}\mathcal{T} \tag{21}$$

We can also express the identity matrix $\mathcal{I}$ in terms of the eigenvectors of $\mathcal{M}$:

$$\mathcal{I} = \mathcal{T}^{-1}\mathcal{I}\mathcal{T} \tag{22}$$

Consequently we can rewrite eqn. (17) as a generalized diagonal matrix transform.

$$\mathcal{T}\underline{p} = [\sigma_1\mathcal{I} + \sigma_2\mathcal{D}]\mathcal{T}\Omega(1)\underline{\epsilon} \tag{23}$$

Equation (23) states that diagonal invariance holds between the *canonical* surface and all other surfaces given the sensor transformation $\mathcal{T}$. In fact eqn. (23) implies that diagonal invariance holds between *any* two surfaces. Let i and j index two arbitrary surfaces with σ 2–vectors $\underline{\sigma}^i$ and $\underline{\sigma}^j$. From eqn. (23), under any illuminant, we can write $\mathcal{T}\underline{p}^i$ and $\mathcal{T}\underline{p}^j$ as *fixed* diagonal transforms of $\mathcal{T}\underline{p}^s$ (the canonical surface appearance):

$$\mathcal{T}\underline{p}^i = [\sigma_1^i\mathcal{I} + \sigma_2^i\mathcal{D}]\mathcal{T}\underline{p}^s \tag{24}$$

$$\mathcal{T}\underline{p}^j = [\sigma_1^j\mathcal{I} + \sigma_2^j\mathcal{D}]\mathcal{T}\underline{p}^s \tag{25}$$

Clearly we can write $\mathcal{T}\underline{p}^i$ as a diagonal matrix times $\mathcal{T}\underline{p}^j$:

$$\mathcal{T}\underline{p}^i = \mathcal{D}^{ij}\mathcal{T}\underline{p}^j \tag{26}$$

where

$$\mathcal{D}^{ij} = [\sigma_1^i\mathcal{I} + \sigma_2^i\mathcal{D}][\sigma_1^j\mathcal{I} + \sigma_2^j\mathcal{D}]^{-1} \tag{27}$$

This completes the proof of Lemma (2). In the *3-2* case there exists a sensor transformation $\mathcal{T}$ with respect to which there is diagonal invariance and this invariance implies that a diagonal matrix is sufficient to support perfect color constancy (Lemma (1)). Therefore, this also completes the proof of Theorem (1). ∎

The crucial step in the above derivation is the eigenvector decomposition of the transform matrix $\mathcal{M}$. To relate this analysis to traditional theories of diagonal matrix color constancy we would like the eigenvalues of $\mathcal{M}$ to be real-valued. However, whether or not they are depends on the form of the surface matrices (and hence the initial sensor spectral sensitivities).

On first consideration complex eigenvalues appear problematic—e.g., transforming the sensors by a complex matrix of eigenvectors does not have a plausible physical interpretation. The problem lies in the fact that the new sensors would be partly imaginary; however, we show in the Appendix that complex eigenvalues fit seamlessly into our generalized theory of diagonal matrix color constancy.

4 Implications for Other Theories of Color Constancy

Under the *3-2* conditions the lighting matrices $\Lambda(\underline{\epsilon})$ are 3×2 surjective maps—color vectors are linear combinations of the two column vectors of $\Lambda(\underline{\epsilon})$—and surfaces seen under a single illuminant span a plane in the 3-dimensional receptor space. Maloney and Wandell[23] exploit this plane constraint in their algorithm for color constancy. Maloney[21] proves that each illuminant corresponds to a unique plane of response vectors. This uniqueness condition is sufficient to solve for the illuminant weight vector $\underline{\epsilon}$ and hence the pseudo-inverse $[\Lambda(\underline{\epsilon})]^{-1}$. Consequently the surface weight vector (or Maloney descriptor) can be recovered via equation (9).

We present an alternative *simpler* color constancy algorithm for the *3-2* world. Our algorithm solves for the diagonal matrix mapping the gamut of observed responses into the gamut of canonical responses.

4.1 Simpler Color Constancy

In the *3-2* world the response vectors for surface reflectances under the canonical illuminant lie on the 'canonical plane' $\mathcal{P}^c$. The span of the canonical plane is defined by the column vectors, $\underline{v}_1$ and $\underline{v}_2$, of the 3×2 spanning matrix V, and are calculated prior to the color constancy computation.

Under each other illuminant, response vectors for surfaces lie on the observed (or image) plane $\mathcal{P}^o$. Assuming that there are at least 2 linearly independent surfaces in our image we can solve for the spanning matrix W—the columns of W, $\underline{w}_1$ and $\underline{w}_2$ are simply the response vectors of any two distinct surfaces.

Theorem 2 *The diagonal transform mapping $\mathcal{P}^o$ onto $\mathcal{P}^c$ is unique.*

Lemma 3 *The only diagonal matrix mapping a plane onto itself is the identity matrix $\mathcal{I}$.*

Proof of Lemma 3. If S is a 3×2 matrix defining the span of a plane, $\underline{n}$ denotes the plane normal, and $\mathcal{D}$ is a diagonal matrix then $[\mathcal{D}\underline{n}]^t S = 0$. This is true only when $\mathcal{D} = \mathcal{I}$. ∎

Proof of Theorem 2. Let us assume that there are two diagonal matrices, $\mathcal{D}_1$ and $\mathcal{D}_2$, which differ by more than a simple scaling, mapping $\mathcal{P}^o$ onto $\mathcal{P}^c$:

$$\mathcal{D}_1 W = V A_1 \tag{28}$$

$$\mathcal{D}_2 W = V A_2 \tag{29}$$

where A_1 and A_2 are 2×2 matrices transforming the span V. Solving for W in eqn. (28) and substituting into eqn. (29) we see that

$$\mathcal{D}_2[\mathcal{D}_1]^{-1} V A_1 = V A_2 \tag{30}$$

By Lemma 3, only the identity matrix maps a plane onto itself (eqn. (30)). Hence $\mathcal{D}_1 = k\mathcal{D}_2$ (where k is a scalar), which contradicts our initial assumption; thus Theorem 2 follows. ∎

The first spanning vector of W, $\underline{w}_1$, can be mapped onto $\mathcal{P}^c$ by applying a linear combination of two diagonal matrices

$$[\alpha\mathcal{D}^{11} + \beta\mathcal{D}^{12}]\underline{w}_1 = \alpha\underline{v}_1 + \beta\underline{v}_2 \tag{31}$$

Similarly $\underline{w}_2$ can be mapped onto $\mathcal{P}^c$ by applying linear combinations of the diagonal matrices $\mathcal{D}^{21}$ and $\mathcal{D}^{22}$. Because the diagonal matrix mapping $\mathcal{P}^o$ to $\mathcal{P}^c$ is unique, the set of diagonal matrices defined by $\mathcal{D}^{11}$ and $\mathcal{D}^{12}$ must intersect those defined by $\mathcal{D}^{21}$ and $\mathcal{D}^{22}$ in a unique diagonal matrix.

We can use this property to develop a simple algorithm for color constancy. The algorithm requires two distinct colors in the image. It proceeds in 3 stages:

1. Find the set D_1 of diagonal matrices mapping the first image color to the set of all canonical colors.
2. Find the set D_2 of diagonal matrices mapping a second image color to the set of all canonical colors.
3. The unique diagonal matrix mapping all image colors to their canonical appearance is equal to $D_1 \bigcap D_2$.

This algorithm is closely related to Forsyth's CRULE[8]. One difference, however is that through our analysis we can solve for the unique diagonal matrix by examining the color appearance of only two surfaces. In contrast CRULE would examine all observed response vectors. This is a simple incarnation of Forsyth's CRULE algorithm. Consequently Maloney's computational method is equivalent to CRULE under a sensor transformation. However, CRULE is less restrictive than Maloney's algorithm and can achieve color constancy even when the *3-2* conditions are relaxed. This is not true for Maloney's algorithm. Therefore, for trichromatic color constancy Maloney's theory is a sub-theory of Forsyth's CRULE. Previously Forsyth had cast *3-2* color constancy as a subtheory of his more complex MWEXT theory.

4.2 Other Theories

The color constancy problem is made more difficult if the illuminant intensity varies across the image. Horn[14] presented an algorithm for removing intensity gradients from images of a Mondrian world. Unfortunately his approach imposed strong constraints on the form of the Mondrian boundary. Later Blake[2] extended this algorithm to allow less restrictive boundary constraints. Key to their algorithms is diagonal invariance, and hence diagonal matrix color constancy. Therefore lightness recovery was thought to be applicable only for visual systems with narrow band sensors.

Funt and Drew[10] presented a non-diagonal lightness algorithm for illuminants and reflectances that are well-approximated by finite-dimensional models. Their method is independent of the sensor spectral sensitivities. However through our analysis of diagonal invariance and diagonal color constancy in section 3 holds for arbitrary spectral sensitivity functions under an appropriate sensor transformation. Our analysis therefore circumvents the need for a non-diagonal lightness theory—Funt and Drew's algorithm reduces to Blake's algorithm under a sensor transformation in the case of *3-2* world conditions.

Land's Retinex theory[19] and its precursor, von Kries adaptation[32], assume that color constancy is achieved if each image contains a known reference patch. By assuming diagonal invariance between the appearance of arbitrary surfaces with the appearance of the reference patch, implies that a diagonal matrix supports color constancy. Our analysis demonstrates that diagonal invariance holds for all sensor sets given only weak constraints.

Color balancing is equivalent to reference patch color constancy. The colors in a video image are adjusted such that a known reference patch appears to have its correct color. Our analysis demonstrates that a simple diagonal matrix balancing is sufficient for color correction—even when the sensors are broad-band. In the next section we evaluate reference patch color constancy (and hence color balancing) before and after a sensor transformation.

5 Experimental Results

Real illuminants are not 3-dimensional and real surfaces are not 2-dimensional—the *3-2* conditions only provide an approximation of color appearance—and hence a diagonal matrix can achieve only approximate color constancy. Here we perform simulations, using measured surface reflectances and measured illuminants, comparing the performance of diagonal matrix and generalized diagonal matrix color constancy.

The color appearance of surfaces viewed under different illuminants are generated using eqn. (3). The human cone responses measured by Vos and Walraven[31] are used as our sensors, the 462 Munsell Spectra[25] for surfaces and the 5 Judd Daylight phases[17](D48, D55, D65, D75 and D100) and CIE A[33] for illuminants. All spectra are sampled at 10nm (nanometer) intervals from 400 to 650nm. Consequently the integral of eqn. (3) is approximated as a summation.

The sensor transformation $\mathcal{T}$ was calculated via the technique outlined in section 3. Singular value decompositions of the Munsell and illuminant spectra were performed to derive the required surface and illuminant basis functions. Figure 1 displays the cone functions before and after the sensor transformation $\mathcal{T}$. Notice that the transformed sensors appear more narrowband—this is consistent with the pragmatic observation that narrow-band sensors afford better diagonal matrix color constancy. A similar narrowing has been observed in various psychophysical experiments[9, 13, 28, 18, 26, 6] involving the human visual system.

There are many algorithms for diagonal matrix color constancy; each differs in its strategy for determining the diagonal matrix. Here we present simulation results for the simplest diagonal matrix algorithm—the *white patch normalization*. The starting point for that algorithm is diagonal invariance. A color vector $\underline{p}_i$ is assumed to be diagonally invariant to the appearance of a white patch $\underline{p}_w$.

$$\underline{p}^i = \mathcal{D}^{iw}\underline{p}^w \tag{32}$$

Hence it is the diagonal matrix $\mathcal{D}^{iw}$ which is independent of the illuminant, and consequently can be used as a descriptor. Usually $\mathcal{D}^{iw}$ is written in vector (or descriptor) form $\underline{d}^{iw}$ where $d_k^{iw} = \frac{p_k^i}{p_k^w}$. By the symmetry between diagonal matrix color constancy and diagonal invariance we can rewrite eqn. (32) as a color constancy transform.

$$\underline{d}^{iw} = [diag(\underline{p}^w)]^{-1}\underline{p}^i \tag{33}$$

where the function *diag* converts the vector $\underline{p}^w$ to a diagonal matrix (diagonal elements correspond to the rows of $\underline{p}^w$). Arbitrarily we chose the white patch descriptor vectors calculated for D55 as the *canonical descriptor vectors*—these provide a reference for determining color constancy performance. Under each of the other 5 illuminants we calculate white patch descriptors. The Euclidean distance between these descriptors and their canonical counterparts, normalized with respect to the canonical descriptor's length, provides a measurement of constancy performance. The percent normalized fitted distance (NFD) metric is defined as:

$$\text{NFD} = 100 * \frac{\| \underline{d}^{iw,e} - \underline{d}^{iw,c} \|}{\| \underline{d}^{iw,c} \|} \tag{34}$$

where $\underline{d}^{iw,c}$ denotes a canonical descriptor and $\underline{d}^{iw,e}$ a descriptor for some other illuminant e. For each illuminant we calculated the following 3 cumulative NFD histograms:

1. the NFD error of white patch normalized responses for the cone functions.

2. the NFD error of generalized white patch normalized responses. (Generalized in the sense of $\underline{d}^{iw,e} = \mathcal{T}^{-1}[diag(\mathcal{T}\underline{p}^w)]^{-1}\mathcal{T}\underline{p}^{i,e}$ [3]

3. the optimal color constancy performance for a general linear transform.

We define optimal color constancy performance to be a least-squares fit relating the responses of all surfaces under an illuminant e to their appearance under the canonical illuminant c. This optimal case, serves as a control for evaluating the color constancy performance afforded by a diagonal matrix.

Figure 2 displays these 3 cumulative histograms for the test illuminants CIE A, D48, D65, D75 and D100 (dashed lines for simple white patch normalization, dotted lines for generalized white patch normalization and solid lines for the optimal constancy performance). In all cases generalized diagonal matrix color constancy constancy outperforms, by a large margin, simple diagonal matrix

[3] The descriptors for all 3 cumulative histograms are with respect to the same sensor basis.

constancy. Generalized diagonal matrix constancy also compares favorably with optimal color constancy. Only for the extremes in test illuminants, CIE A and to a lesser extent D100, is there a significant performance difference.

6 Conclusion

A diagonal matrix is the simplest possible vehicle for color constancy. Indeed, it is its inherent simplicity which has motivated research into more complex matrix forms—if a diagonal matrix can give good color constancy a non-diagonal matrix, which has 9 instead of 3 parameters, must be able to support better color constancy, or so the reasoning goes. The analysis presented in this paper concludes that this is in fact **not** the case. Under weak world constraints a diagonal matrix, in conjunction with an appropriate transformation of the sensor basis, has been shown to suffice for the support of perfect color constancy. This result is strong in the sense that no constraints are placed on the spectral sensitivities of the sensors.

Our simulation studies investigated whether the optimal sensors as expressed in the new sensor basis derived for the *3-2* world would continue to support good color constancy when the *3-2* restrictions were relaxed. For many real reflectances imaged under real illuminants, a diagonal matrix continued to give close to optimal color constancy.

Our analysis establishes a relationship among several theories of color constancy. For a world where illumination is 3-dimensional and surface reflectance 2-dimensional, the Maloney-Wandell[23] algorithm, Forsyth's MWEXT[8] and the lightness theory of Funt and Drew[10] all reduce to diagonal matrix color constancy, and diagonal transforms are already at the heart of Forsyth's CRULE and von Kries adaptation. These non-diagonal algorithms are therefore more complex than necessary and can all be simplified by a fixed transformation of the sensor basis.

Appendix: Complex Eigenvalues

Complex eigenvalues may arise in the eigenvector decomposition of the transform matrix $\mathcal{M}$, but as we will show, they do not present a serious problem.

In traditional theories of diagonal matrix color constancy it is clear that each diagonal constancy transform can be expressed as the sum of three basis transforms. Indeed it is this condition which makes diagonal matrix color constancy so appealing. For example, suppose we observe the color vector $\underline{p}$ and this corresponds to the descriptor $\underline{d}$. This information is sufficient to solve for the constancy transform:

$$\underline{d} = \mathcal{D}\underline{p}\ ,\ \mathcal{D}_{kk} \ = \ \frac{d_k}{p_k} \tag{35}$$

This same *uniqueness condition* is clearly true in generalized diagonal matrix color constancy if the sensor transformation $\mathcal{T}$ is real–valued. In fact, the uniqueness condition also holds in the general case where the elements of $\mathcal{T}$ can have complex terms.

Theorem 3 *Under any sensor transformation $\mathcal{T}$ (where $\mathcal{T}$ can have complex elements) there are exactly 3 linearly independent diagonal matrices consistent with generalized diagonal matrix color constancy. Consequently the mapping between a color vector and its descriptor is unique.*

Proof. Our original statement of diagonal matrix color constancy, eqn. (2), can be written in the following mathematically equivalent form.

$$\underline{d} \;=\; \mathcal{T}^{-1}\mathcal{D}\mathcal{T}\underline{p} \tag{36}$$

Both $\underline{d}$ and $\underline{p}$ are real-valued vectors and hence $\mathcal{T}^{-1}\mathcal{D}\mathcal{T}$ must be a real-valued matrix. Theorem 2 follows if we can demonstrate that there exist only 3 linearly independent, real-valued matrices with the same eigenvectors—the columns of $\mathcal{T}^{-1}$.

A diagonal matrix $\mathcal{D}$ has 6 variable components, 3 reals and 3 imaginary numbers. Consequently there are in general 6 linearly independent matrices sharing the same eigenvectors. The matrices $\mathcal{T}^{-1}\mathcal{I}\mathcal{T}$, $\mathcal{T}^{-1}\mathcal{D}\mathcal{T}$ and $\mathcal{T}^{-1}\mathcal{D}^{-1}\mathcal{T}$ are all linearly independent, real-valued matrices. Similarly $\mathcal{T}^{-1}\mathcal{I}j\mathcal{T}$, $\mathcal{T}^{-1}\mathcal{D}j\mathcal{T}$ and $\mathcal{T}^{-1}\mathcal{D}^{-1}j\mathcal{T}$ are all linearly independent, purely imaginary matrices (j is the square root of -1). The sum of imaginary numbers is always imaginary and conversely the sum of real numbers is always real; hence these 6 matrices span the set of all matrices with eigenvectors $\mathcal{T}^{-1}$. Including complex numbers in the field over which we form a span, this means that only 3 matrices form a basis for the span of all real valued matrices with eigenvectors $\mathcal{T}^{-1}$. This completes the proof for Theorem 2. ∎

Theorem 2 states that generalized diagonal matrix constancy holds equally well even when the sensor transformation is complex. For any sensor transformation the diagonal color constancy transform can be expressed as the sum of three diagonal basis matrices $\mathcal{D}$, $\mathcal{D}^{-1}$ and $\mathcal{I}$. The mapping $\mathcal{D}^{ij}$, in equation (13), taking $\underline{p}^{i,e}$ to $\underline{p}^{j,e}$ is still unique and is independent of the illuminant.

Figure Captions

Figure 1. Result of sensor transformation $\mathcal{T}$. Solid lines: Vos–Walraven cone fundamentals; dashed lines: transformed sensors.

Figure 2. Cumulative histograms showing improved performance of generalized diagonal color constancy. Dashed lines: simple diagonal color constancy; dotted lines: generalized diagonal color constancy; solid lines: optimal (non-diagonal) color constancy.

References

[1] R. Bajcsy, S.W. Lee, and A. Leonardis. Color image segmentation and with detection of highlights and local illumination induced by inter-reflections. In *International Conference on Pattern Recognition, Atlantic City*, volume 1, pages 785–790. IEEE, 1990.

[2] A. Blake. Boundary conditions for lightness computation in Mondrian world. *Computer Vision, Graphics, and Image Processing*, 32:314–327, 1985.

[3] M.H. Brill. A device performing illuminant–invariant assessment of chromatic relations. *J. Theor. Biol.*, 71:473–478, 1978.

[4] J. Cohen. Dependency of the spectral reflectance curves of the Munsell color chips. *Psychon. Sci.*, 1:369–370, 1964.

[5] M.S. Drew and B.V. Funt. Variational approach to mutual illumination in color images, 1991. Submitted for publication.

[6] G.D. Finlayson, M.S. Drew, and B.V. Funt. Spectral sharpening: an optimal sensor transformation for colour constancy, 1991. In preparation.

[7] G.D. Finlayson, M.S. Drew, and B.V. Funt. Enhancing von kries adaptation via sensor transformations. SPIE, 1993.

[8] D. Forsyth. A novel algorithm for color constancy. *Int. J. Comput. Vision*, 5:5–36, 1990.

[9] D.H. Foster and R.S. Snelgar. Initial analysis of opponent-colour interactions revealed in sharpened field sensitivities. In J.D. Mollon and L.T. Sharpe, editors, *Colour Vision : Physiology and Psychophysics*, pages 303–312. Academic Press, 1983.

[10] B.V. Funt and M.S. Drew. Color constancy computation in near–Mondrian scenes using a finite dimensional linear model. In *Computer Vision and Pattern Recognition Proceedings*, pages 544–549. IEEE Computer Society, June 1988.

[11] B.V. Funt and M.S. Drew. Color space analysis of mutual illumination. *IEEE Trans. Patt. Anal. and Mach. Intell.*, 1991. In Press.

[12] B.V. Funt and G.D. Finlayson. Color constant color indexing. Technical Report CSS/LCCR TR 91-09, Simon Fraser University School of Computing Science, 1991.

[13] W. Jaeger H. Krastel and S. Braun. An increment-threshold evaluation of mechanisms underlying colour constancy. In J.D. Mollon and L.T. Sharpe, editors, *Colour Vision : Physiology and Psychophysics*, pages 545–552. Academic Press, 1983.

[14] B. K. P. Horn. Determining lightness from an image. *Computer Vision, Graphics, and Image Processing*, 3:277–299, 1974.

[15] A. Hurlbert. Formal connections between lightness algorithms. *J. Opt. Soc. Am. A*, 3:1684–1692, 1986.

[16] A.C. Hurlbert. *The Computation of Color (PhD Thesis)*. MIT Artificial Intelligence Laboratory, 1989.

[17] D.B. Judd, D.L. MacAdam, and G. Wyszecki. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 54:1031–1040, August 1964.

[18] M. Kalloniatis and R.S. Harwerth. Spectral sensitivity and adaptation characteristics of cone mechanisms under white-light adaptation. *J. Opt. Soc. Am. A*, 7:1912–1928, 1990.

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

[20] E.H. Land and J.J. McCann. Lightness and retinex theory. *J. Opt. Soc. Amer.*, 61:1–11, 1971.

[21] L.T. Maloney. *Computational Approaches to Color Constancy.* PhD thesis, Stanford University, 1985. Applied Psychology Lab.

[22] L.T. Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *J. Opt. Soc. Am. A*, 3:1673–1683, 1986.

[23] L.T. Maloney and B.A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *J. Opt. Soc. Am. A*, 3:29–33, 1986.

[24] D.H. Marimont, B.A. Wandell, and A.B. Poirson, 1992. Tech. Report.

[25] Newhall, Nickerson, and D.B. Judd. Final report of the osa subcommittee on the spacing of the munsell colors. *J. Opt. Soc. Am. A*, 33:385–418, 1943.

[26] A.B. Poirson and B.A. Wandell. Task-dependent color discrimination. *J. Opt. Soc. Am. A*, 7:776–782, 1990.

[27] S.A. Shafer. Using color to separate reflection components. *Color Res. Appl.*, 10:210–218, 1985.

[28] H.G. Sperling and R.S. Harwerth. Red-green cone interactions in the increment-threshold spectral sensitivity of primates. *Science*, 172:180–184, 1971.

[29] M.J. Swain and D.H. Ballard. Indexing via color histograms. In *Proceedings: International Conference on Computer Vision, Osaka , Dec.4-7/90*, pages 390–393. IEEE, 1990.

[30] F. Tong and B.V. Funt. Specularity removal for shape from shading. In *Proceedings: Vision Interface 1988*, pages 98–103, Edmonton, Alberta, Canada, 1988.

[31] J.J VOS and P.L. WALRAVEN. On the derevation of the foveal receptor primaries. *Vision Research*, 11:799–818, 1971.

[32] G. West and M.H. Brill. Necessary and sufficient conditions for von kries chromatic adaption to give colour constancy. *J. Math. Biol.*, 15:249–258, 1982.

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