Unifying Colour Constancy

Graham D. Finlayson and Steven Hordley
School of Information Systems
University of East Anglia
Norwich, NR4 7TJ
UK

Paul M. Hubel
H-P Laboratories
Hewlett-Packard Company
Palo Alto, CA
USA

Abstract

In this paper we consider the problem of colour constancy; how given an image of a scene under an unknown illuminant can we recover an estimate of that light? We develop a general *correlation* framework in which solving for colour constancy is posed as a correlation of the colours in an image with the colours that can occur under each of a set of possible lights. Rather than attempting to recover a single estimate of the illuminant as many previous authors have done, we, in the first instance, recover a correlation measure for each possible illuminant. We then select an estimate of the scene illuminant based on these correlations.

The work presented here follows from previously published [9] work. In this paper we extend that work by showing that the correlation framework is rich enough to allow many existing algorithms to be expressed within it. The grey-world, maximum RGB, gamut mapping, and Maloney-Wandell algorithms, perhaps the algorithms most widely cited in the literature, are presented in this correlation framework. This work together with work published elsewhere [7] shows that almost all published algorithms based on a Mondrian world can be formulated in the framework presented here. Significantly, the correlation framework can be used to add value to existing algorithms. For example, some of the problems associated with the Maloney-Wandell algorithm can be removed.

1. Introduction

An image of a white piece of paper taken with a digital camera under a yellow illuminant will appear yellow. The same piece of paper under a blue or red illuminant will similarly appear blue or red. In contrast the same piece of paper viewed by a human observer under any of the lights, will appear white. The reason for this is that the human visual system is *colour constant.* That is, it is capable of discounting the scene illumination to arrive at a stable percept of object colour.

To ensure that the colours in images accurately match the colours an observer saw in the original scene, it is important that cameras too are colour constant. If the scene illumination is known then it is relatively straightforward [13] to correct an image to account for the illumination. Unfortunately the illumination in a scene is generally unknown, therefore it is necessary to infer it from the image data. To this end many so called colour constancy algorithms have been proposed [15, 16, 3, 12, 17, 10, 6, 11, 2, 5]. The starting point for most algorithms is the Mondrian world model of image formation:

$$\underline{p} = \int_{\omega} E(\lambda)S(\lambda)\mathbf{R}(\lambda) \quad (1)$$

where $\underline{p}$ represents the RGB response of a device with spectral sensitivities $\mathbf{R}(\lambda)$, to a surface reflectance function, $S(\lambda)$, under an illuminant with spectral power distribution (SPD) $E(\lambda)$. Eqn. (1) shows that the colour of the surface is confounded with the illumination so the RGB response of the device to a surface is not a stable descriptor of that surface.

Suppose we define surface colour to be the response of a device to the surface viewed under some reference illuminant and similarly define illuminants by the response of the device to an achromatic surface viewed under that illuminant, then both surfaces and lights can be represented by 3 parameters. Given an image containing N_{surf} surfaces we have $3N_{surf} + 3$ parameters to solve for. However, from Eqn. (1), we have only $3N_{surf}$ equations. That there are fewer knowns than unknowns means that it is difficult to decouple light from surface and solve for colour constancy. In mathematical parlance, colour constancy is an ill posed problem.

In order to make the problem well posed, a variety of additional world constraints have been suggested. Land [15] assumed that every image contains a white patch so that the response of one surface is known which reduces the number of unknowns to the number of knowns. A different approach [3, 12] is to assume that the average surface reflectance in a scene is achromatic, so that the average light leaving the scene is that of the incident illuminant. Other authors [17, 4] have exploited the fact that lights and surfaces can be accurately represented as low-dimensional linear models to develop algebraic schemes to solve for colour constancy.

In this paper, rather than trying to further constrain the problem we consider an alternative point of view. Given the model of image formation in Eqn. (1), and without making any further assumptions about the world, it follows that the solution to colour constancy is in general not unique. That is, a particular set of image data will be consistent with many combinations of light and surfaces. For example, a reddish sensor response is consistent with a white surface viewed under a reddish light, a reddish surface under a white light, and with many other combinations of light and surface. Our approach to solving for colour constancy is, in the first instance, to find all such combinations of light and surfaces which are consistent with the image data and then choose a single combination from amongst these.

The work presented here is a natural extension of previously published work [9, 14]. In this work we posed the colour constancy problem in a *correlation matrix memory* framework. The method works by first building a correlation matrix to correlate image colours with the set of possible scene illuminants. Each column of the matrix is a discrete representation of a colour space (chromaticity space for example) stretched out into a vector. A column of the matrix corresponds to a possible scene illuminant, and its entries tell us which colours can be seen under that illuminant. Given a set of image data we build a corresponding image vector which tells us which colours are present in the image. We then correlate this image vector with each column of the correlation matrix to obtain a measure of the degree of correlation between the image data and each possible scene illuminant. Based on these correlations we choose an estimate of the scene illuminant.

Previous work [9] has shown that this method affords very good colour constancy, and is significantly better than all previously published algorithms. Furthermore, the approach is almost trivial to implement, computationally simple, and robust. The contribution of this paper is to show that the correlation matrix memory is in fact a general framework for solving for colour constancy. We demonstrate that a number of previously published algorithms can be cast in our framework. In this way, the correlation approach unifies colour constancy algorithms which at first glance appear to be very different. Indeed, this work, together with work published elsewhere [7] shows that all algorithms which adopt the model of image formation in Eqn. (1) are either precisely expressible in the correlation framework, or are very closely related to it. Uniting colour constancy algorithms in a common framework, a pleasing result in itself, also helps us understand the relationship between algorithms and their relative strengths and weaknesses. As we shall see it also provides means for improving these algorithms.

2. A framework for colour constancy

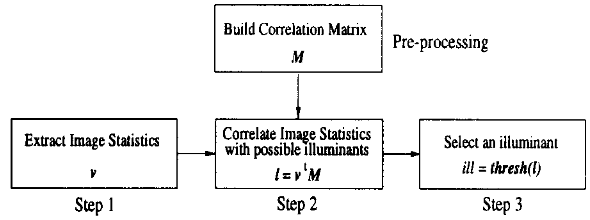

Figure 1: *The general correlation framework for colour constancy.*

Our framework for colour constancy is illustrated in Figure 1. As a first step in solving for colour constancy we must build a correlation matrix memory M which tells us which colours can be observed under which illuminant. Pre-processing aside, there are 3 steps in estimating the scene illuminant for a particular image. The first is to extract some statistics from the image. Typically, this will be a histogram of the colours in the image, but we use the general term image statistics, to allow for other data to be used if and when it is required. Step 2 is to correlate the image statistics with the correlation matrix memory M. Or equivalently, with the set of possible scene illuminants. This results in a set of correlation statistics which tell us how well correlated is the image data with each of the possible illuminants. These correlation statistics are used in the 3^{rd}, illuminant selection, stage of the algorithm, to choose a single illuminant from the set of possible illuminants as an estimate of the scene illuminant. As we shall see, these steps are quite general, and allow many algorithms to be expressed in the framework. In this section we develop the framework by showing how a particular algorithm - Finlayson's [6] *colour in perspective* algorithm - can be formulated within it.

Finlayson's algorithm is founded on the notion of colour gamuts: the set of all image colours possible under a particular illuminant (an idea originally suggested by Forsyth [10]) In Finlayson's algorithm colour gamuts are modelled analytically as closed continuous regions of a 2-d chromaticity space. Illuminants are represented by the mappings (actually 2-d diagonal matrices) which take their colour gamuts to the colour gamut of some reference or canonical light. Given a particular image, solving for colour constancy is then a two stage process. First, the set of mappings taking all image colours to the gamut of the reference light is determined. Each of these mappings represents a possible illuminant. However, Finlayson pointed out that not all of these mappings necessarily correspond to illuminants which occur in the real world. To restrict mappings to those corresponding to plausible illuminants Finlayson defined a gamut of mappings corresponding to all lights that were physically plausible and then intersected this set with the set recovered in the first stage of his algorithm. From the resulting set of mappings a single mapping was chosen as an estimate of the unknown scene illuminant.

2.1. Building the correlation matrix

To implement Finlayson's algorithm in our framework we first build a correlation matrix M_{Fin}. This entails characterising the set of possible scene illuminants, and determining the range of image colours that can be observed under each of these lights. We characterise the i^{th} possible illuminant by its chromaticity co-ordinate $\underline{c}_i$ in some 2-d chromaticity space. In Finlayson's algorithm this space is $(r/b, g/b)$. The key difference between the framework we present here and Finlayson's formulation is that we use a discrete representation of this chromaticity space, so that the set of possible illuminants between which we wish to distinguish is finite. For convenience we represent these illuminants in an $N_{ill} \times 2$ matrix C_{ill} whose i^{th} row is the chromaticity co-ordinate of the i^{th} illuminant. We further assume that possible colours are specified in the same discretised chromaticity space and that there are a finite number of such colours. Since we are characterising illuminants and surfaces in terms of device sensor responses the discrete representation is wholly valid because a digital camera can only return a finite and discrete set of sensor responses. For example the responses of a camera giving 8-bit data are RGB triplets where each of R, G, and B is an integer between 0 and 255.

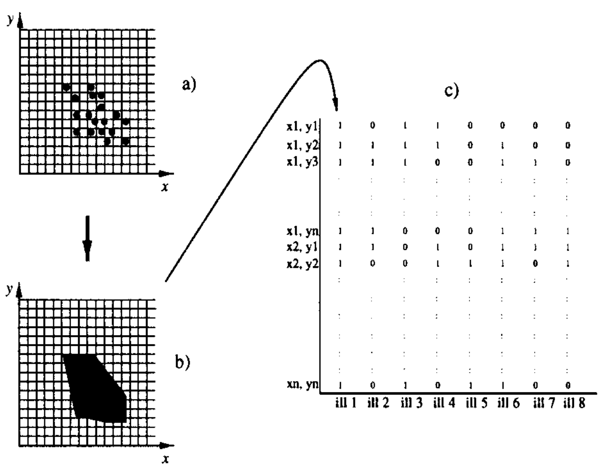

Figure 2: Three steps in building a correlation matrix. (a) The sensor responses of a set of reference surfaces under an illuminant are determined. (b) The gamut of colours possible under this light is found. (c) The process is repeated for all possible illuminants to form the correlation matrix M.

For each possible scene illuminant we want to determine which colours can be observed under it. To do this we take a set of reference surfaces, representative of the range of naturally occurring surface reflectances and use Eqn. (1) to determine the corresponding sensor responses. From these sensor responses we calculate the chromaticities (plotted in Fig. 2a). The set of possible chromaticities under the illuminant is taken to be the convex hull of these points (Fig. 2b). From this we form a vector, each element of which corresponds to a binary representation of the discrete chromaticity space. Entries in the vector are set to one if the corresponding chromaticity is possible under the illuminant, and to zero otherwise. This process is repeated for all possible illuminants and the vector corresponding to the i^{th} illuminant forms the i^{th} column of the correlation matrix M_{Fin} (Fig. 2c). To restrict illuminants to only those which occur in the real world we can either remove a column corresponding to an implausible illuminant from M_{Fin} or alternatively set all the entries of the column to zero. Representing our knowledge about which image colours are possible under which illuminant is the heart of our correlation framework. As we will see later, implementing different algorithms mostly amounts to changing the entries of the correlation matrix M_{Fin}.

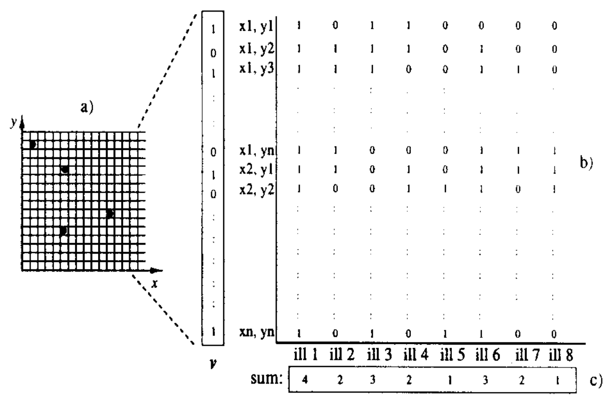

Figure 3: Solving for colour constancy. The computation proceeds in 3 stages. (a) Histogram the chromaticities in the image. (b) Determine the degree of correlation between the image data and each possible scene illuminant. (c) Use these correlations to select an estimate of the unknown illuminant.

2.2. Extracting image statistics

Given a particular image Finlayson's algorithm first determines the set of chromaticities present in the image. We represent these in an $N_{surf} \times 2$ matrix C_{im} whose i^{th} row corresponds to the i^{th} surface in the image. From these image chromaticities (plotted in Figure 3a) we determine an image vector $\underline{v}$ (Fig 3b). Each element of this vector represents a bin of the discrete chromaticity space, and an element is set to one if its corresponding chromaticity is present in the image, and to zero otherwise. We can define a binary histogramming operation $binhist2d()$ to represent this process, so that:

$$\underline{v} = binhist2d(C_{im}) \tag{2}$$

2.3. Correlating the image data

Figure 3c illustrates how the image data is correlated with the correlation matrix. Highlighted rows of the matrix in this figure correspond to entries of $\underline{v}$ which are one (that is the corresponding chromaticity is present in the image). Summing the highlighted entries of each column in the matrix tells us how many image chromaticities are consistent with each illuminant. Mathematically we can write this process as:

$$\underline{l} = \underline{v}^t M \quad = \quad binhist2d(C_{im})^t M \tag{3}$$

so that the i^{th} entry of $\underline{l}$ contains the number of image chromaticities consistent with illuminant i. That is it expresses the degree of correlation between the image data and each of the possible scene illuminants.

2.4. Illuminant selection

From this correlation data we can determine the set of feasible scene illuminants recovered by Finlayson's algorithm. For an illuminant to be feasible it must be consistent with all image colours. For example, the image in Figure 3 contains 4 chromaticities ($\sum_i v_i = 4$), so the set of plausible illuminants I_{plaus} are those whose corresponding entries in $\underline{l}$ are 4. However, factors such as image noise often mean that not all image chromaticities are consistent with any choice of illuminant. So we find the illuminant that is consistent with most image colours:

$$I_{plaus} = diag(thresh(\underline{l}))C_{ill} \tag{4}$$

where $thresh()$ is defined such that:

$$thresh(l_i) = \begin{cases} 1, & \text{if } l_i = max(\underline{l}) \\ 0, & \text{otherwise} \end{cases} \tag{5a}$$

$$thresh(\underline{l}) = [thresh(l_1), \ldots, thresh(l_{N_{ill}})] \tag{5b}$$

and the operation $diag(\underline{a})$ returns a diagonal matrix whose diagonal entries are the elements of $\underline{a}$.
The final stage in Finlayson's algorithm is to select an illuminant from this plausible set as an estimate of the unknown illuminant. We could do this by, for example, averaging all the illuminants in I_{plaus}. In our framework this is written:

$$\hat{\underline{c}} = thresh2(\underline{l})C_{ill} \tag{6}$$

where $thresh2()$ is $thresh()$ modified such that:

$$thresh2(l_i) = \begin{cases} 1/N_{plaus}, & \text{if } l_i = max(\underline{l}) \\ thresh2(l_i) = 0, & \text{otherwise} \end{cases} \tag{7}$$

and N_{plaus} is the number of plausible illuminants $N_{plaus} = \sum_i thresh(l_i)$.

In our framework Finlayson's algorithm (with mean selection) can be summarised as:

$$\hat{\underline{c}} = thresh2(binhist(C_{im})^t M_{Fin})C_{ill} \tag{8}$$

Eqn. (8) is remarkable because it is so simple. The original colour in perspective algorithm, based on continuous representations of colours, was necessarily based on a series of complex geometric calculations.

In the next section we show that Eqn. (8) is a general framework for colour constancy and that by simply changing the correlation matrix M_{Fin} and modifying the histogramming and thresholding operations a number of other algorithms can be posed in the same form.

3. Other algorithms in the framework

3.1. Grey-World

We begin with the so called *grey-world* algorithm. This algorithm has been proposed in a variety of forms by a number of different authors [3, 12, 16] and is based on the assumption that the spatial average of surface reflectances in a scene is achromatic. Since the light reflected from an achromatic surface is changed equally at all wavelengths it follows that the spatial average of the light leaving the scene will be the colour of the incident illumination. If we wish to recover an estimate of the scene illuminant in the form of the sensor response of a device to the illuminant, then the grey-world algorithm is trivial to implement; we simply take the average of all sensor responses in the image:

$$\underline{p}^E = mean(RGB_{im}). \tag{9}$$

where RGB_{im} is an $N_{pix} \times 3$ matrix whose i^{th} row is the RGB of the i^{th} pixel in the image. Equivalently we can define a matrix RGB_{ill} the rows of which contain all possible RGB sensor responses and re-write Eqn. (9) as:

$$\underline{p}^E = hist(RGB_{im})^t \mathcal{I} RGB_{ill} \tag{10}$$

where the operation $hist()$ returns a vector $\underline{h}$ such that:

$$\begin{gathered} \underline{h} = vec(H) \\ H(R,G,B) = \textstyle\sum_{j=1}^{N_{pix}} f(R,G,B), \\ f(R,G,B) = \begin{cases} 1 & \text{if } RGB_{im}(j) = (R,G,B) \\ 0 & \text{otherwise} \end{cases} \end{gathered} \tag{11}$$

In Eqn. (11) $H(x, y, z)$ returns the count of the number of pixels in the image such that: $R = x$, $G = y$ and $B = z$. The operator $vec()$ takes the 3-dimensional array H and stretches it out as a long vector of numbers.

Returning to Eqn. (10) we see that the mean RGB can be calculated by multiplying the colour histogram of an image with the possible illuminants RGB_{ill}. However, we have added the matrix $\mathcal{I}$ into this equation (which we needn't have since $\mathcal{I}RGB_{ill} = RGB_{ill}$). But, the addition is useful as $\mathcal{I}$ plays exactly the same role as the correlation matrix M_{Fin}: it represents our knowledge about the interaction between image colours and surfaces. The columns and rows of $\mathcal{I}$ represent possible illuminants and possible image colours respectively. Hence, $\mathcal{I}$ tells us that given an RGB sensor response in an image, the only illuminant consistent with it is the illuminant characterised by the same sensor response. Correspondingly the vector

$$\underline{l} = hist(RGB_{im})^t \mathcal{I} \qquad (12)$$

whose elements contain the number of image colours consistent with each illuminant, can be interpreted as a measure of the likelihood that each illuminant (each distinct RGB present in the image) is the scene illuminant. Based on these likelihoods an estimate of the scene illuminant is calculated by taking the weighted average of all illuminants. In this framework it is clear why grey-world is an inadequate solution to the colour constancy problem; the matrix $\mathcal{I}$ does not accurately encode our knowledge about the correlation between lights and image colours. We have seen in Section 2 how this information can be more accurately represented in the correlation matrix framework and correspondingly, the colour in perspective algorithm affords significantly better colour constancy [8].

3.2. Modified Grey-World

Gershon *et al* [12] noted another limitation of the grey-world algorithm. They pointed out that the spatial average computed in Eqn. (9) is biased towards surfaces of large spatial extent. To alleviate this problem they modified the algorithm by segmenting the image into patches of uniform colour prior to estimating the illuminant. The sensor response from each segmented surface is then counted only once in the spatial average, so that surfaces of different size are given equal weight in the average. It is trivial to add this feature in our framework; we simply need to use the histogramming operation $binhist3d()$, ($binhist2d()$ defined in Section 2, modified to work on RGBs rather than chromaticities) rather than $hist()$:

$$\underline{p}^E = binhist3d(RGB_{im})^t \mathcal{I} RGB_{ill} \qquad (13)$$

3.3. Gamut Mapping Algorithms

In Section 2 of this paper we formulated Finlayson's colour in perspective algorithm in the correlation matrix framework. This algorithm is a refinement of work by Forsyth [10] who was the first to use the notion of colour gamuts in solving for colour constancy. Rather than modelling colour gamuts and illuminants in a 2-d chromaticity space as did Finlayson, Forsyth's algorithm was formulated in 3-d camera RGB space. With this change, a 3-D gamut mapping solution to colour constancy can be written in our framework as:

$$\hat{\underline{p}} = thresh2(binhist3d(RGB_{im})^t M_{For}) RGB_{ill} \qquad (14)$$

where the ij^{th} entry of M_{For} is one if image colour i is possible under illuminant j and is zero otherwise. We note that this is not exactly Forsyth's original algorithm since he used a different method to select an illuminant from the feasible set. We use mean selection here, which has been shown [1] to give very similar performance.

Posing the gamut mapping algorithms in our framework significantly simplifies the computation, which in the original formulation of these algorithms was laborious. The framework also allows us to further improve the algorithms. In both algorithms, a single image colour inconsistent with all possible illuminants caused them to give no solution to colour constancy. In reality we wish to relax things and search for solutions that are consistent with all or almost all image colours. Since, if 50 chromaticities are consistent with illuminant A and 48 with illuminant B then it is likely that both illuminants are possible. This majority consistency is also trivial to implement in our framework; we simply define a new threshold function $thresh3()$ such that:

$$thresh3(x) = \begin{cases} 1/N_{plaus} & \text{if } x \geq m \\ 0 & \text{otherwise} \end{cases} \qquad (15)$$

where m is chosen in adaptive fashion such that $m \leq max(\underline{l})$. The improved gamut mapping algorithm can then be written as:

$$\hat{\underline{p}} = thresh3(binhist3d(RGB_{im})^t M_{For}) RGB_{ill} \qquad (16)$$

3.4. Maloney-Wandell

Many authors [3, 4, 17] have developed algorithms for colour constancy based on a representation of lights and surfaces as low-dimensional linear models. That is, illuminant SPDs $E(\lambda)$, and surface reflectance functions $S(\lambda)$, are expressed as linear combinations of a finite number of basis functions:

$$E(\lambda) = \sum_{i=1}^{n} E_i(\lambda)\epsilon_i, \quad S(\lambda) = \sum_{j=1}^{m} S_j(\lambda)\sigma_j \qquad (17)$$

In this representation lights and surfaces are characterised by their weights $\underline{\epsilon}$ and $\underline{\sigma}$ and solving for colour constancy is posed as recovering these vectors for the light and surfaces in an image. With these models the image formation equation (Eqn. (1)) can be rewritten:

$$\underline{p} = \Lambda_{\epsilon}\underline{\sigma} \qquad (18)$$

where Λ_{ϵ} is called the *lighting matrix* and has kj^{th} entry $\int E(\lambda)S_j(\lambda)R_k(\lambda)d\lambda$. Maloney and Wandell [17] set out conditions under which Eqn. (18) could be solved for $\underline{\epsilon}$ and $\underline{\sigma}$. Specifically they showed that if there were p classes of device sensor, and if illuminants and surfaces could be represented by p and $(p-1)$-d linear models, then colour constancy was soluble. For example, a solution is possible for $p = 3$ classes of sensor if surfaces and illuminants are described by 2 and 3-d linear models. Under such conditions the sensor responses of all surfaces viewed under a particular illuminant must lie in a plane of the sensor space. Furthermore this plane uniquely determines the lighting matrix Λ_{ϵ}. Given a particular image, illuminant recovery in Maloney and Wandell's scheme amounts to finding the plane which best fits the image data and using this plane to determine $\underline{\epsilon}$ and hence the lighting matrix Λ_{ϵ}.

In our framework we begin with a set of illuminants between which we wish to distinguish and a representation of colours in a discretised colour space, for example RGB sensor space. Given a set of reference surfaces imaged under a possible illuminant we can determine the plane which best fits the sensor data. The intersection of this plane with the discretised RGB space tells us the gamut of image colours which can be observed under the illuminant. We can thus determine a correlation matrix M_{M-W} the columns of which are formed by repeating this process for each possible illuminant. Solving for colour constancy can then once again be performed in our framework:

$$\hat{\underline{p}} = thresh2(binhist3d(RGB_{im})^t M_{M-W})RGB_{ill} \qquad (19)$$

With only 3 classes of sensor Maloney and Wandell's algorithm does not produce good colour constancy since, illuminants are not in general well described by 2-d linear models as prescribed by their scheme. Furthermore the illuminant recovered by their process is not guaranteed to be physically realisable. The implementation of the algorithm given here does guarantee physical realisablility however it does not account for the dimensionality problem. We predict however, that Eqn. (19) may well deliver some degree of colour constancy because the RGB planes that define an illuminant are based on preprocessing (i.e. a good model of surface reflectance statistics) and not on the image data itself. Moreover, the thresholding step will allow us to quantify the range of possible solutions. It could be that the Maloney Wandell algorithm (viewed in the correlation framework) is indecisive about what light is possible and the fact that the algorithm delivers poor colour constancy is explained by the uncertainty in the illuminant likelihood calculation.

We are currently evaluating the Maloney Wandell algorithm, and other algorithms, in the correlation frameworks. Quantitative results that compare conventional algorithm performance with that delivered by corresponding correlation implementations are currently being carried out. We predict that the latter will significantly outperform the former.

3.5. Maximum RGB

Finally we consider one of the most simple, and most widely used, algorithms for colour constancy [15]. A white patch reflects all light equally at all wavelengths therefore the sensor response to a white patch determines the illuminant colour. So if an image contains a white patch and if this patch can be found then we should be able to solve for colour constancy. Since any other surface in the image must reflect proportionately less light than the white patch, simply finding the surface with maximum RGB should suffice to find the white patch. In fact we do not require there to be a white patch in the image but only that there are surfaces in the image which reflect all light to which the red, green, and blue sensors are sensitive. An estimate of the unknown illuminant can then be found by:

$$\hat{\underline{p}} = max(RGB_{im}) \qquad (20)$$

Before proceeding further we point out to the reader that Eqn. (20) is a highly non-linear function and so it might appear surprising that it can be expressed in the correlation framework.

The sensor response to a white patch under a given illuminant limits the gamut of image colours for that illuminant: any other RGB sensor response must fall within the cube defined by the RGB response to the white patch. Hence we can determine a correlation matrix M_{Max} whose ij^{th} entry is one if the i^{th} RGB response is less than the RGB response of illuminant j to a white patch, and is zero otherwise. Given an image, a set of plausible illuminants can then be found in our framework by:

$$I_{plaus} = diag(thresh(hist(RGB_{im})^t M_{Max}))RGB_{ill} \qquad (21)$$

From I_{plaus} we want to find an estimate of the unknown illuminant. Figure 4 characterises the set of possible illuminants returned by Eqn. (21). For ease of illustration this figure considers 2-d sensor responses. The shaded region in Figure 4a represents the set I_{plaus} - for 3-d sensor responses we have a cuboid rather than a rectangle. In Figure 4b we see what happens when the RGBs are inverted,

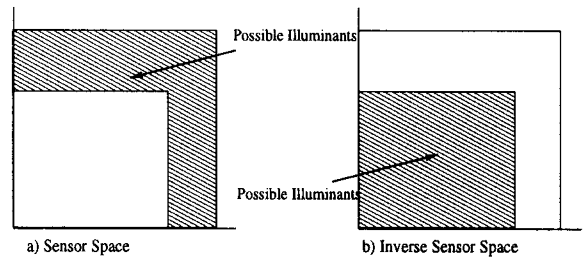

Figure 4: The feasible solutions for the max-RGB algorithm. When inverted they become are cube; the mean of this is the estimate of the unknown illuminant.

the possible illuminants now fall in rectangle (a cuboid in 3-d). If we sample this space finely enough, then taking the average of all the inverse RGBs gives us an inverse RGB which is, to within a scalar, the maximum scene RGB. That is, the maximum RGB in our framework can be calculated as $1/\underline{\hat{p}}$:

$$\underline{\hat{p}} = thresh2(binhist(RGB_{im})^t M_{Max}) RGB_{inv} \quad (22)$$

where RGB_{inv} is an $N_{ill} \times 3$ matrix whose rows are the inverse RGBs of each possible illuminant.

4. Conclusions

In this paper we have considered the colour constancy problem; that is how we can find an estimate of the unknown illuminant in a captured scene. We have presented here a correlation framework in which to solve for colour constancy. The simplicity, flexibility and robustness of this framework makes solving for colour constancy easy (in a complexity sense). Moreover, and this is the main result of the paper, we have shown that a number of existing algorithms for colour constancy can be expressed in this framework. These include the commonly used grey-world and max-RGB algorithms and also the Maloney-Wandell algorithm.

By casting all algorithms in a single computational framework we are able to see how one algorithm relates to another. Not only is this useful in explaining differential algorithm performance but, it also helps us understand how these algorithms might be improved.

This algorithm complements a companion work [7] where we have shown that other algorithms, including the Bayesian [2] and Neural Net [11] approaches, can also be placed in the correlation framework. These works taken together, show that all colour constancy algorithms which are based solely on the colour statistics of an image, can be placed in, or are closely related to, the correlation framework.

Acknowledgements

We would like to thank the EPSRC and Hewlett-Packard Company for their support of this work.

References

[1] K. Barnard. Computational Color Constancy: Taking Theory Into Practice. Master's thesis, Simon Fraser Univ., School of Computing Science, 1995.

[2] David H. Brainard and William T. Freeman. Bayesian color constancy. *Journal of the Optical Society of America, A*, 14(7):1393–1411, 1997.

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

[4] M. 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(10):2148–2165, 1993.

[5] M. M. D'Zmura and G. Iverson. Probabilistic Color Constancy. In M. M. D'Zmura, D. Hoffman, G. Iverson, and K. Romney, editors, *Geometric Representations of Perceptual Phenomena: Papers in Honor of Tarow Indow's 70th Birthday*. Laurence Erlbaum Associates, 1994.

[6] G. D. Finlayson. Color in Perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):1034–1038, 1996.

[7] G.D. Finlayson, S.D. Hordley, and P.M. Hubel. Colour by correlation: A simple unifying framework for colour constancy. In *ICCV99*, September 1999.

[8] Graham Finlayson and Steven Hordley. A theory of selection for gamut mapping colour constancy. In *Computer Vision and Pattern Recognition '98*, pages 60–65. IEEE, June 1998.

[9] Graham D. Finlayson, Paul M. Hubel, and Steven Hordley. Color by Correlation. In *Fifth Colour Imaging Conference*, pages 6–11. IS&T/SID, November 1997.

[10] D. A. Forsyth. A Novel Algorithm for Colour Constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990.

[11] Brian V. Funt, Vlad Cardei, and Kobus Barnard. Learning color constancy. In *Proceedings of the Fourth Color Imaging Conference*, pages 58–60, November 1996.

[12] Ron Gershon, Allan D. Jepson, and John K. Tsotsos. From [R,G,B] to Surface Reflectance: Computing Color Constant Descriptors in Images. *Perception*, pages 755–758, 1988.

[13] Paul M. Hubel, Jack Holm, Graham D. Finlayson, and Mark S. Drew. Matrix calculations for digital photography. In *Proceedings of the Fifth Color Imaging Conference*, pages 105–111. IS&T/SID, November 1997.

[14] P.M. Hubel and G.D. Finlayson. Whitepoint determination using correlation matrix memory. U.S. Patent Application, Submitted.

[15] Edwin H. Land. The Retinex Theory of Color Constancy. *Scientific American*, pages 108–129, 1977.

[16] Edwin H. Land. Recent advances in retinex theory and some implications for cortical computations. *Proc. Natl. Acad. Sci. USA*, 80:5163–5169, 1983.

[17] Laurence T. Maloney and Brian A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *Journal of the Optical Society of America, A*, 3(1):29–33, 1986.