Colour by Correlation: A Simple, Unifying Approach to Colour Constancy

G. D. Finlayson and S. D. Hordley
School of Information Systems
University of East Anglia
Norwich, NR4 7TJ

P. M. Hubel
Hewlett-Packard Labs.
Hewlett Packard Inc.
Palo Alto, CA 94306

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? Rather than recovering a single estimate of the illuminant as many previous authors have done, in the first instance we recover a measure of the likelihood that each possible illuminant was the scene illuminant. We do this by correlating image colours with the colours that can occur under each of a set of possible lights. We then recover an estimate of the scene illuminant based on these likelihoods. Computation is expressed and performed in a generic correlation framework which we develop in this paper. We develop a new probabilistic instantiation of this framework which delivers very good colour constancy on synthetic and real images. We show that the proposed framework is rich enough to allow many existing algorithms to be expressed within it; e.g. the grey-world and gamut mapping algorithms. We explore too the relationship of these algorithms to other probabilistic and neural network approaches.

1This work was carried out at the Colour & Imaging Institute, University of Derby. The authors wish to thank that institute for their support of this work.
1 Introduction

In this paper we consider the colour constancy problem; that is, given an image of a scene under an unknown illumination how can we recover an estimate of the unknown illuminant? This is an important problem to solve since various applications (e.g. object recognition, digital photography) require that we be able to disambiguate factors in an image; for example, the surface reflectance properties of the imaged objects, and the spectral power distribution of the incident illumination.

Over many years a number of authors [17, 19, 4, 5, 2, 3, 17, 22, 10, 6] have proposed computational theories and algorithms for colour constancy but to date only limited success has been achieved. At the heart of almost all approaches is the simple Lambertian model of image formation:

$$p_k^x = \int_{\omega} E(\lambda) S^x(\lambda) R_k(\lambda) d\lambda$$

(1)

where $p_k^x$ is the response of the imaging device’s $k^{th}$ sensor at pixel $x$. It is clear from Equation (1) that changing either the surface reflectance function, $S(\lambda)$, or the illuminant spectral power distribution $E(\lambda)$, will change the values recorded by the imaging device. The task for a colour constancy algorithm is then to transform $p_k^x$ to values which are independent of $E(\lambda)$, and hence which correlate with $S(\lambda)$, or equivalently, to recover an estimate of $E(\lambda)$ (since the transform that discounts $E(\lambda)$ depends only on $E(\lambda)$).

The problem is difficult to solve: $E(\lambda)$ and $S(\lambda)$ are continuous functions, yet to estimate them we typically have only 3 measurements: the camera RGB. Even if we define the colour of lights and surfaces by their respective RGBs there are still more unknowns to solve for than there are knowns (with $n$ surfaces in a scene, we have $3n$ unknowns and $3n + 3$ unknowns). It follows that the problem is intrinsically ill-posed.

Many authors [17, 3, 14, 19, 4] have tried to deal with the under-constrained nature of the problem by making additional assumptions about the world. For example Land [17] implicitly assumes that every image contains a white patch, hence there are now only $3n$ unknowns and $3n$ equations. A similar assumption employed by a number of authors [3, 14], is to assume that the average reflectance of all surfaces in a scene is achromatic, so that the average colour of the light leaving the surfaces in a scene will be the colour of the incident illumination. Another approach often adopted [19, 4] has been to model lights and surfaces using low-dimensional linear models and to develop recovery schemes which exploit the features of these models. Other authors have used features such as specularities [18, 22], shadows [7] or mutual illumination [11], to recover information about the scene illuminant. Unfortunately, the assumptions made by all these algorithms are quite often violated in real images, and it has been shown [12] that none of these schemes is capable of delivering adequate colour constancy.

The fact that the problem is under-constrained implies that in general the combination of surfaces and illuminant giving rise to a particular image is not unique. So setting out to solve for a unique answer (the goal of most colour constancy algorithms) is not the best way to proceed. This point of view has only been considered in more recent algorithms. For example, the gamut mapping algorithms developed by Forsyth [10], and later by Finlayson [6] and others [8], do not, in the first instance, attempt to find a unique solution to the problem. Rather, the set of all possible solutions are found and from this set “the best” solution is chosen. Other authors [2, 21, 5], recognising that the problem does not have a unique solution, have tried to exploit information in the image to recover “the most likely” solution. Several authors [2, 5] have posed the problem in a probabilistic framework, and more recently Sapiro [21] has developed an algorithm based on the Probabilistic Hough Transform. The neural network approach [13] can similarly be seen as a method of dealing with the inherent uncertainty in the problem.

The first contribution of this paper is to present a new general correlation framework for colour constancy. In this new framework colour constancy is posed as a correlation of the colours in an image and our prior knowledge about which colours can appear under which lights. Intuitively, this correlation idea has merit; if the colours in an image “look” more yellow than they ought to then one might assume that this yellowness correlated with a yellow illuminant. We develop a particular instantiation of this framework, a new correlation algorithm, which has a number of attractive properties. It enforces the physical realizable properties of lights and surfaces, it is insensitive to
spurious image colours, it is fast to compute and it calculates the most likely answer. Perhaps more important than all this is that it delivers good colour constancy.

The second contribution of this paper is to show that the correlation framework we have developed is general and can be used to describe many existing colour constancy algorithms. This leads to a deep understanding of the colour constancy problem, the proposed solutions and their interrelations. Furthermore, solving for colour constancy with any of these algorithms in our framework is always computationally very simple. Yet in comparison algorithms like gamut mapping have previously been criticised for their inherent complexity. Here we see that they are in fact no more complex than the simplest type of colour constancy computation.

In Section 2 of this paper we present our framework (for which a patent has been applied [16]) for solving colour constancy and instantiate a new algorithm within it. We then show, in Section 3, that a number of existing algorithms can be precisely formulated in the same framework. Section 4 details the experiments we performed to test the new algorithm and gives a quantitative and qualitative assessment of its performance. Finally we draw some conclusions from this work in Section 5.

2 Colour by Correlation

We pose the colour constancy problem as that of recovering an estimate of the scene illumination from an image of a scene taken under an unknown illuminant since given this estimate it is relatively straightforward to transform image colours to illuminant independent descriptors [15]. We restrict attention here to the case of an imaging system with three classes of sensors. In such a case it is not possible to recover the full spectral power distribution of the illuminant, so instead an illuminant with spectral power distribution \( E(\lambda) \) is characterised by \( \hat{p}^E \); the response of the imaging device to an achromatic surface under \( E(\lambda) \). An estimate of the illuminant will accordingly be a 3-vector sensor response, \( \hat{p}^E \), or since we cannot recover the overall intensity of the illuminant, but only its orientation, a 2-d chromaticity vector \( \hat{\chi}^E \).

Characterising illuminants in this way we can define the set of all distinguishable illuminants. For example, for a digital camera which gives 8-bit data, the sensor responses will be in the range 0 to 255, and the set of illuminants which can be recovered are the set of all possible sensor responses, \( \hat{p}^E \) whose elements are in this range. For the purpose of our algorithm we represent illuminants by their chromaticity vectors and define an \( N_I \times 2 \) matrix \( C_{IL} \) whose \( i \)th row is the chromaticity of the \( i \)th illuminant. Similarly we use the notation \( C_{im} \) to denote the \( N_{\text{pix}} \times 2 \) matrix of image chromaticities (the \( N_{\text{pix}} \) pixels in an image stretched out). As for illuminants and RGBs we will assume that chromaticity space is discretised. Specifically, we assume chromaticity space is split into \( N \times N \) uniform regions i.e. there are at most \( N^2 \) distinct chromaticities in an image.

We solve for colour constancy in three stages. First we build a correlation matrix to correlate image colours with the set of possible scene illuminants. Then, given a set of image data we determine the relative likelihood that each of the possible illuminants was the scene illuminant. That is, we determine for each illuminant \( E \) a measure of the likelihood that the illuminant was \( E \), given the image data \( G_{im} \). Finally, we use these likelihoods to recover an estimate of the scene illuminant.

Given a set of image data \( G_{im} \) we would like to recover \( \Pr(\hat{C} | G_{im}) \) - the probability that \( E \) was the scene illuminant given \( C_{im} \). Now, if we know the probability of observing a certain chromaticity \( \hat{C} \) under illuminant \( E \); \( \Pr(\hat{C} | E) \), then Bayes’ rule tells us how to calculate the corresponding probability \( \Pr(E | \hat{C}) \):

\[
\Pr(E | \hat{C}) = \frac{\Pr(\hat{C} | E) \Pr(E)}{\Pr(\hat{C})} \tag{2}
\]

It follows that the probability that the illuminant was \( E \) given the image data \( G_{im} \) is given by:

\[
\Pr(E | G_{im}) = \frac{\Pr(G_{im} | E) \Pr(E)}{\Pr(G_{im})} \tag{3}
\]
For a given image \( \Pr(C_{im}) \) is constant, and if furthermore we assume that all illuminants are equally likely, and that image chromaticities are independent then we can re-write Equation (3) as:

\[
\Pr(E|C_{im}) = k \prod_{\mathbf{c} \in C_{im}} \Pr(\mathbf{c}|E)
\]

where \( k \) is a constant. Now, we define a likelihood function:

\[
l(E|C_{im}) = \sum_{\mathbf{c} \in C_{im}} \log(\Pr(\mathbf{c}|E))
\]

and note that the illuminant which maximises \( \Pr(E|G_{im}) \) will also maximise \( l(E|C_{im}) \). The log-probabilities effectively measure the correlation between a particular chromaticity and a particular illumination. Maximising \( l \) in Equation (5) finds the illuminant that correlates most strongly with the image chromaticities. We talk here about correlation rather than just likelihoods since correlation is a more general term. As we shall see in the next section many other algorithms can be cast in the language of correlation.

Usually we think about correlation in terms of a vector dot-product. For example if \( \mathbf{a} \) and \( \mathbf{b} \) are vectors then they are strongly correlated if \( \mathbf{a} \cdot \mathbf{b} \) is large. A similar dot-product definition is used here in order to implement Equation (5). We define a correlation matrix \( M_{Bayes} \) whose \( ij^{th} \) entry is: \( \log(\Pr(\text{image chromaticity } i|\text{illuminant } j)) \) and a vector \( \mathbf{h} \) whose \( th \) entry is one if chromaticity \( i \) occurs in the image and zero otherwise. It then follows that the log likelihood function \( l(E|G_{im}) \) is given by \( \mathbf{h}^T M_{Bayes} \); the correlation of \( \mathbf{h} \) with each column (i.e. illuminant) in \( M_{Bayes} \). Alternatively, we can write this as:

\[
l(E|C_{im}) = \text{thr}(\text{chist}(C_{im}))^T M_{Bayes}
\]

where the operation \( \text{chist}() \) is a histogramming operation returning an \( N_{pix} \times \) 1 vector \( \mathbf{h} \), the \( th \) element of which holds the number of times a chromaticity corresponding to the \( th \) row of \( C_{ill} \) occurs in the image, normalised by \( N_{pix} \), the total number of pixels in the image and \( \text{thr}(\mathbf{h}) \) returns a vector whose \( th \) element is one if \( h(i) > 0 \) and is zero otherwise.

Based on this log likelihood function we want to recover an estimate of the chromaticity, \( \mathbf{c}^E \) of the unknown illuminant. In this algorithm we choose the illuminant which maximises \( l(E|G_{im}) \), that is:

\[
\mathbf{c}^E = \text{thr}^2(\text{thr}(\text{chist}(C_{im}))^T M_{Bayes})C_{ill}
\]

where \( \text{thr}^2() \) returns a vector with entries corresponding to:

\[
\text{thr}^2(h) = h^T h' = \begin{cases} 
1, & \text{if } h_i = \max(h) \\
0, & \text{otherwise}
\end{cases}
\]

Equation (7) defines a well founded maximum likelihood solution to colour constancy. It is important to note that since we have computed likelihoods for all illuminants we can augment the illuminant calculated in (7) with error bars. That is we can inform the user that the illuminant was probably yellow Tungsten but it could have been cool white fluorescent but it was definitely not Blue sky daylight. We believe that this is a major strength inherent in our method and will be of significant value in other computer vision applications.

### 3 Other algorithms in this framework

Our framework for solving colour constancy is encapsulated in Equation (7). At the heart of this framework is a correlation matrix which encodes our knowledge about the interaction between lights and image colours. To solve for colour constancy we simply have to correlate the image data with this matrix to recover a measure of the likelihood that each illuminant was the scene illuminant and then use these likelihoods to estimate the scene illuminant. We show in this section that many existing algorithms can be re-formulated in the framework set out above. To do this we simply have to change the entries of the correlation matrix so as to reflect the assumptions about the interactions between lights and image colours made (at least implicitly) by these algorithms.
3.1 Grey-World

We begin with the so-called grey-world algorithm which as well as being one of the simplest is still widely used (despite the fact that it often fails). This algorithm has been proposed in a variety of forms by a number of different authors [3, 14] 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. To recover an estimate of the scene illuminant in the form we require; that is in terms of the sensor response of a device to the illuminant, is trivial; we simply need to take the average of all sensor responses in the image: $p^E = \text{mean}(RGB_{im})$. Equivalently, this can be written in our framework as:

$$\hat{p}^E = \text{hist}(RGB_{im})^t I RGB_{il}$$

where $RGB_{il}$ characterises the set of all possible illuminants in camera RGB space, the operation $\text{hist}()$ is $\text{chist}()$ modified to work on RGBs rather than chromaticities, and the matrix $I$ is the identity matrix. In this formulation $I$ replaces the correlation matrix $M_{Bayes}$ and, as before, can be interpreted as representing our knowledge about the interaction between image colours and surfaces. In this interpretation the columns and rows of $I$ represent possible illuminants and possible image colours respectively. Hence, $I$ tells us that given a sensor response $p$ in an image, the only illuminant consistent with it is the illuminant characterised by the same sensor response.

3.2 Gamut Mapping

Clearly $I$ does not accurately represent our knowledge about lights and surfaces; for example a reddish RGB is consistent with both a red surface under a white light and a white surface under a red light. Forsyth [10] developed an algorithm, called CRULE to exploit this fact. CRULE is founded on the idea of colour gamuts: the set of all possible sensor responses observable under different illuminants. He showed that colour gamuts are closed, convex, bounded, and most importantly that each is a strict subset of the set of possible image colours. The gamut of possible image colours for a light is determined by imaging all possible surfaces (or a representative subset thereof) under that light. We can similarly determine gamuts for each of our possible illuminants, i.e. for each row of $RGB_{il}$ and can code them in a matrix $M_{For}$ such that, if image colour $i$ can be observed under illuminant $j$ then we put a one in the $ij$th entry of $M_{For}$, otherwise we put a zero. The matrix $M_{For}$ more accurately represents our knowledge about the world and can be used to replace $I$ in Equation (9).

Forsyth used the colours present in an image to determine a set of feasible illuminants. An illuminant is taken to be feasible if all image colours fall within the gamut defined by that illuminant. In our framework, the number of image colours consistent with each illuminant can be calculated:

$$l = \text{thr}(\text{hist}(RGB_{il})^t)M_{For}$$

where $\text{thr}()$ and $\text{hist}$ are as defined previously and ensure that each distinct image colour is counted only once.

Once the set of feasible illuminants has been determined, the final step in CRULE is to select a single illuminant from this set as an estimate of the unknown illuminant. Previous work has shown [1] that the best way to do this is to take the mean of the feasible illuminants. In our framework this can be achieved by:

$$\hat{p}^E = \text{thr}2(\text{thr}(\text{hist}(RGB_{il}))^t M_{For})RGB_{il}$$

where $\text{thr}2()$ is defined as before. In CRULE the notion of which image colours can appear under which lights was modelled analytically as closed continuous convex regions of RGB space which leads to a very expensive computation. Equation (11), though equivalent to CRULE (with mean selection) is a much simpler implementation of it.

Computation aside, our new formulation has another significant advantage over CRULE; rather than saying that an illuminant is possible if and only if it is consistent with all image colours we can instead look for illuminants
that are consistent with most image colours. This subtle change cannot be incorporated easily into the CRULE algorithm yet it is important to do so. In CRULE, if no illuminant is globally consistent then a null set of mappings is returned. That is, there is no solution to colour constancy. What we really are looking for is majority consistency and this is in fact implemented in our new computational framework. The function \( thr^2 \) depends on the maximum count in the histogram. Of course we may wish to make the threshold a little less than the maximum so as not to exclude illuminants that have more or less the same likelihood. We define \( thr^3() \) operating on the vector \( l \) such that:

\[
thr^3(x) = \begin{cases} 
1, & \text{if } x \geq m, \\
0, & \text{(other wise)} 
\end{cases}
\]

Where \( m \) is chosen in an adaptive fashion such that \( m \leq \max(h) \). Thus, we rewrite (11) as:

\[
\hat{\lambda}^E = thr^3(thr(hist(RGB_{im})^t M_{For}) RGB_{ill}) 
\]

3.3 Colour In Perspective

While the formulation of Forsyth’s CRULE algorithm given above addresses some of its limitations there are other problems with CRULE which this formulation doesn’t resolve. First, Finlayson recognised that features such as shape and shading, affect the magnitude of the recovered light but significantly, not its colour. To avoid calculating calculating the intensity of the illuminant (which cannot be recovered) Finlayson carried out computation in a 2-d chromaticity space. If we once again characterise image colours and illuminants by their chromaticities we can define a new matrix, \( M_{Fin} \), whose \( ij^{th} \) element will be set to one when chromaticity \( i \) can be seen under illuminant \( j \) and to zero otherwise. We can then substitute \( M_{Fin} \) in Equation (7)

\[
\hat{\lambda}^E = thr^2(thr(hist(C_{im}))^t M_{Fin})C_{ill} 
\]

Assuming the thresholding operations, \( thr() \) and \( thr^2() \) are chosen as for Forsyth’s algorithm (we could of course use \( thr^3() \) instead of \( thr^2() \) if we wished to implement majority consistency), then the illuminant estimate \( \hat{\lambda}^E \) is the averaged chromaticity of all illuminants consistent with all image colours. Previous work [8] has shown however, that this is not the best estimate of the illuminant, and that the chromaticity transform should be reversed before the averaging operation is performed. This can be achieved here by defining a matrix \( RGB_{ill}^N \), whose \( ij^{th} \) row is the \( i^{th} \) row of \( RGB_{ill} \) normalised to unit length. The illuminant estimate is now calculated:

\[
\hat{\lambda}^E = thr^2(thr(hist(C_{im}))^t M_{Fin}) RGB_{ill}^N 
\]

Another problem with gamut mapping is that not all illuminants in \( G_{ill} \) are plausible (for example purple lights do not occur in practice). This observation is also simple to implement. Suppose the \( k^{th} \) row of \( C_{ill} \) denotes the chromaticity of an illuminant that does not occur in practice. By setting all the entries in the \( k^{th} \) column of \( M_{Fin} \) to 0 we record the fact that no image colours are consistent with the \( k^{th} \) illuminant and so guard against the possibility of ever choosing illuminant \( k \) (which, by assumption, is impossible). Equivalently we can completely remove the \( k^{th} \) row of \( C_{ill} \) and the \( k^{th} \) column of \( M_{Fin} \).

3.4 Illuminant Color by Voting

Sapiro [21] has recently proposed an algorithm for estimating scene illumination which is based on the Probabilistic Hough Transform. In this work lights and surfaces are represented as low-dimensional linear models and probability distributions are defined from which surfaces are drawn. Given a sensor response from an image, a surface is selected according to the defined distribution. This surface, together with the sensor response is used to recover an illuminant. If the recovered illuminant is a feasible illuminant (in Sapiro’s case an illuminant on the daylight locus) a vote is cast for that light. For each sensor response many surfaces are selected and so many votes are cast. To get an estimate of the illuminant the cumulative votes for each illuminant are calculated by summing the votes from all sensor responses in the image. The illuminant with maximum votes is selected as the scene illuminant.
The votes for all illuminants for a single sensor response, \( p \), represent an approximation to the probability distribution: \( \Pr(E|p) \) - the conditional probability of the illuminant given the observed sensor response. Sapiro chooses the illuminant which maximises the function: \( \sum_{i \in RGB} \Pr(E|p) \). Since we know the range of possible image colours, rather than compute the probability distributions \( \Pr(E|p) \) on a per image basis, we could instead, using Bayes rule, compute them once for all combinations of sensor responses and illuminants. We can then define a matrix \( M_{\text{Sapiro}} \) whose \( ij^{th} \) entry is \( \Pr(\text{illuminant } j | \text{image colour } i) \) - which, by Bayes rule, is proportional to the probability of observing image colour \( i \) under illuminant \( j \). It then follows that Sapiro’s estimate of the illuminant can be found in our framework by:

\[
P^E = \text{thr}3(\text{thr}(\text{hist}(\text{RGB}_{\text{im}})^t \cdot M_{\text{Sapiro}})) \cdot \text{RGB}_\text{light}
\]

and we note that the matrix \( M_{\text{Sapiro}} \) is equal to \( k_{\text{hist}} \cdot M_{\text{Page}} \).

### 3.5 Probabilistic Algorithms

Brainard et al [2] have recently given a Bayesian formulation of the colour constancy problem. Their approach is again founded on a linear models representation of lights and surfaces so that lights and surfaces are characterised by low-dimensional vectors of weights. They used principal component analyses of collections of surfaces and illuminants to define probability distributions for these weights and then used Bayesian decision theory to recover estimates of the weights for the surfaces and illuminant in an image. If \( \vec{x} \) represents the combined vector of weights for surfaces and lights, the problem is to find an estimate \( \hat{\vec{x}} \) of \( \vec{x} \).

If there are \( N \) surfaces in the image then the vector to be recovered is \((3N + 3)\)-dimensional. Finding \( \hat{\vec{x}} \) is therefore computationally extremely complex. The authors have implemented the algorithm as a numerical search problem and shown results for the case \( n = 8 \). However, since typical images contain many more surfaces than eight, as a practical solution for colour constancy this approach is far too complex. A precise formulation of their algorithm is not possible within our framework, however we can use their prior distributions on surfaces and illuminants when constructing our correlation matrix. If we then restrict the problem to that of recovering an estimate of the unknown illuminant, then the two approaches should produce similar results.

An approach which is closer to the algorithm we have presented was proposed by D’Zmura et al [5]. They also adopted a linear models representation of surfaces, but they used these models to derive a likelihood distribution \( \Pr((x,y)|E(\lambda)) \) that is the probability of observing a given CIE-xy chromaticity [24] co-ordinate, under an illuminant \( E(\lambda) \). This is done by first defining distribution functions for the weights of their linear model of surfaces. They then generated a large number of surfaces by selecting weights according to these distributions and calculated the corresponding chromaticity co-ordinates for these surfaces. By selecting a large number of surfaces, a good approximation to \( \Pr((x,y)|E(\lambda)) \) can be found. If we put likelihoods corresponding to these probabilities in a correlation matrix, then this algorithm can be formulated in the framework we have developed. We point out that this algorithm like grey-world takes no account of the relative frequency of individual chromaticities: the function \( \text{thr} \) is not used. As such, the algorithm is highly sensitive to large image areas of uniform colour and so can suffer serious failures (similar to grey-world).

### 3.6 Neural Networks

The final algorithm we consider here is the neural network approach of Funt et al [13]. Computation proceeds in three stages and is summarised below:

\[
\begin{align*}
\text{output}_1 &= \text{thr}4(\text{thr}(\text{hist}(C_{\text{im}})^t \cdot M_{\text{Funt}})) \\
\text{output}_2 &= \text{thr}4(\text{output}_1 \cdot M_{\text{Funt},1}) \\
\text{output}_3 &= \text{output}_2 \cdot M_{\text{Funt},2}
\end{align*}
\]

where \( \text{thr}4 \) is similar to \( \text{thr}3 \) but exactly what it is has not been specified [13]. In the parlance of Neural Nets \( \text{output}_1 \) is the first stage in a 3 layer perceptron calculation. The 2nd, hidden layer computation, is modelled by the second equation and the final output of the Neural net is \( \text{output}_3 \). The correlation matrix \( M_{\text{Funt},1} \) typically
has many fewer columns than $\mathbf{M}_{\text{unt}}$. Moreover, $\mathbf{M}_{\text{unt},2}$ only has two columns and so the whole network only outputs two numbers. These two numbers are trained to be the chromaticity of the actual scene illuminant. As such we can replace $\mathbf{M}_{\text{unt},2}$ by $\mathbf{C}_{\text{ill}}$ (though it is important to realise that here $\mathbf{C}_{\text{ill}}$ is discovered as a result of training and does not bear a one to one correspondence with actual illuminants).

In the context of this paper output $k$ is very similar to Equation (7) albeit with a different correlation matrix and slightly different threshold functions. The other two stages address the question of how a range of possible illuminants is translated into a single illuminant chromaticity estimate. As one might imagine the Neural net approach, which basically fits a parametric equation to model image data, has been shown to deliver reasonably good estimates. However, unlike the approach advocated here, it is not possible to give certainty measures with the estimate nor is it possible to really understand the nature of the computation that is taking place.

4 Results

We conducted a number of experiments to assess the performance of our new correlation algorithm and to compare it to existing algorithms. We present the results of two experiments here. To get some of idea of the algorithm’s performance over a large data set we tested it first on synthetically generated images. In a second experiment we tested the algorithm on a number of real images captured with a digital camera. We show exemplar results of these tests for a small number of images here.

Before running the algorithm we must make the correlation matrix; this could be $\mathbf{M}_{\text{fin}}$, $\mathbf{M}_{\text{Bayes}}$, or some other matrix $\mathbf{M}$ depending on the algorithm we wish to test. In the case of $\mathbf{M}_{\text{fin}}$ we want to determine which image chromaticities are possible under each of the illuminants between which we wish to distinguish. For these experiments we chose a set of 37 illuminants, representing a wide range of commonly occurring indoor and outdoor lights; this set includes a number of daylights of different correlated colour temperatures, tungsten lights, and a variety of fluorescent sources. To determine the range of possible chromaticities we used Equation (1) to calculate the responses of a digital camera to a large set of surface reflectance functions (a combination of the Munsell chip set [24] and a collection of object surface reflectance functions measured by Vrhel et al [23] were used in this experiment) under a given light. From these sensor responses we calculated the corresponding chromaticity co-ordinates, and considered any chromaticity within the convex hull of this set to be a possible chromaticity. The columns of the matrix $\mathbf{M}_{\text{fin}}$ were formed by repeating this process for all of the 37 illuminants. Rows of $\mathbf{M}_{\text{fin}}$ were formed by dividing the chromaticity space into a grid of discrete bins. A one in the $i^{th}$ entry of $\mathbf{M}_{\text{fin}}$ implies that a chromaticity in the $i^{th}$ bin of the chromaticity space is observable under illuminant $j$.

The entries of the matrix $\mathbf{M}_{\text{Bayes}}$ are found in a similar manner, except now, rather than recording only whether a chromaticity is possible or not, we want to record the relative frequency with which it occurs. To do this we took the same set of surface reflectances and calculated chromaticity co-ordinates as before, then, when we discretised the chromaticity space we simply counted the number of chromaticities falling in each bin. The entries of $\mathbf{M}_{\text{Bayes}}$ are the log of these raw counts normalised by the total number of chromaticities. The matrix $\mathbf{M}_{\text{Sapiro}}$ is similarly created except that we put actual probabilities rather than log probabilities in the matrix.

To create the synthetic images on which we tested the algorithms we randomly selected between 2 and 64 surfaces from a set of surface reflectances, and a single illuminant, drawn from the set of 37. To make the test more realistic we used reflectances from a set of natural surface reflectances measured by Parkkinen et al [20] rather than using the same reflectances on which the correlation matrices were created. We calculated the sensor response for a surface and then weighted each surface by a random factor chosen to ensure that the number of pixels in each image was $512 \times 512$.

To run the algorithm, we simply built an image histogram, and calculated the likelihood for each illuminant, by multiplying this histogram by the correlation matrix. These likelihoods were then used to select a single illuminant from the set as an estimate of the scene illuminant. For matrix $\mathbf{M}_{\text{fin}}$ we used the mean selection method [8], whereas for $\mathbf{M}_{\text{Bayes}}$ and $\mathbf{M}_{\text{Sapiro}}$ we chose the illuminant with maximum likelihood.

In total we tested 4 algorithms: grey-world, colour in perspective (with mean selection), Sapiro’s algorithm, and
our new algorithm. Each of these algorithms was used to generate an estimate of the unknown scene illuminant. We used this estimate to re-render the image as it would have appeared under standard daylight D65 illumination by mapping the image data by a $3 \times 3$ diagonal matrix whose entries map an algorithm’s estimate of the scene white to white under D65 illumination. We also corrected the image using the actual (measured) scene illuminant. We then calculated the root mean square error in chromaticity space between these two images. Figure 1 shows the relative performance of these four algorithms. This figure shows the average root mean square error in chromaticity for images containing between 2 and 64 surfaces. In each case the average was taken over 500 synthetic images. We can draw a number of conclusions from these results. First, accurately encoding information about the world leads to improved colour constancy performance; the gamut mapping algorithm, and the two algorithms exploiting probability information all perform considerably better than grey-world. Further, we can see that adding information about the probabilities of image colours under different illuminants significantly improves performance. It is important though, that this information is encoded correctly. Our new algorithm, which correctly employ Bayes’s rule, gives significantly lower average RMSE than the second best algorithm. However, Sapiro’s algorithm which does not correctly encode probability information performs only slightly better than the gamut mapping algorithm.

Previous work [8] has demonstrated that 2-d gamut mapping produces better results than most other algorithms [9, 8]. So, it is significant that our new approach delivers much better constancy. Moreover, the Neural Net approach (for which there insufficient information for us to implement) has also been shown to perform similarly to 2-d gamut mapping. Thus, to our knowledge, our new algorithm significantly outperforms all other algorithms.

The second experiment we ran was to test the performance of the algorithm on real images. To this end we captured a number of images with a digital camera under a variety of different lights. We then used our algorithm to find an estimate of the scene illuminant. As before, we then used this estimate to re-render the captured image under D65 illumination. Figure (2) gives typical examples of the algorithm’s performance for two images. While two images do not represent an exhaustive test of the algorithm they do illustrate its typical performance. The figure also shows that the new algorithm works well even on images on which other commonly used algorithms, such as grey-world, perform badly.

---

1 A colour version of this paper is provided on the CDROM supplied with the conference proceedings.
5 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 seen that existing constancy algorithms are inadequate for a variety of reasons. For example, many of them make unrealistic assumptions about images, or their computational complexity is such that they are unsuitable as practical solutions to the problem. 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, we have shown how a particular Bayesian instantiation of the framework leads to excellent colour constancy (and which is significantly better than other algorithms tested). Many other previously proposed algorithms are also placed within the correlation framework.

Acknowledgements

This work was supported by Hewlett-Packard Incorporated and the EPSRC.

References


