A Primer for Colour Computer Vision

Graham D. Finlayson

Abstract. Still, much of computer vision is predicated on greyscale imagery. There are good reasons for this. For much of the development of computer vision greyscale images were all that was available and so techniques were developed for that medium. Equally, if a problem can be solved in greyscale - and many can be - then the added complexity of starting with 3 image planes as oppose to 1 is not needed. But, truthfully, colour is not used ubiquitously as there are some important concepts that need to be understood if colour is to be used correctly. In this chapter I summarise the basic model of colour image formation which teaches that the colours recorded by a camera depend equally on the colour of the prevailing light and the colour of objects in the scene. Building on this, some of the fundamental ideas of colorimetry are discussed in the context of colour correction: the process whereby acquired camera RGBs are mapped to the actual RGBs used to drive a display. Then, we discuss how we can remove colour bias due to illumination. Two methods are presented: we can solve for the colour of the light (colour constancy) or remove it through algebraic manipulation (illuminant invariance). Either approach is necessary if colour is to be used as a descriptor for problems such as recognition and tracking. The chapter also touches on aspects of human perception.

1 Colour Image Formation

The visible spectrum occupies a very small part of the electromagnetic spectrum. For humans and cameras the visible spectrum lies approximately between 400 and 700 Nanometres[27] (see Figure 1).

School of Computing Sciences,
university of East Anglia,
Norwich
NR4 7TJ
United Kindom
graham@cmp.uea.ac.uk

R. Cipolla et al. (eds.), *Registration and Recognition in Images and Videos*,
Studies in Computational Intelligence 532,
DOI: 10.1007/978-3-642-44907-9_2, © Springer-Verlag Berlin Heidelberg 2014

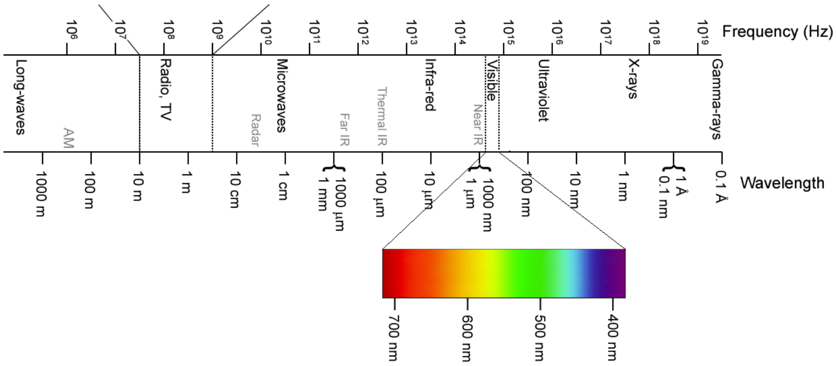

Fig. 1 The Visible Spectrum (Image taken from http://en.wikipedia.org/wiki/Electromagnetic_spectrum)

The spectral power distribution illuminating a scene is denoted $E(\lambda)$. The light strikes an object with surface spectral reflectance $S(\lambda)$ and the light reflected is proportional to the multiplication of the two functions (this product is sometimes called the colour signal). The light is then sampled by a sensor with a spectral sensitivity $R(\lambda)$. The various spectral quantities are shown in Figure 2. The integrated response of a sensor to light and surface is calculated in (1).

$$\rho_k^{E,S} = \int_{\omega} R_k(\lambda)E(\lambda)S(\lambda)d\lambda \quad k \in \{R,G,B\} \tag{1}$$

Where ω denotes the visible spectrum. Immediately, we see that light and surface play, mathematically, the same symmetric role. Each is as important as the other in driving image formation.

Notice that (1) includes no information about either the location of the light sources or the location of the viewer. This is because (1) is an accurate model only for the matte - or Lambertian - aspect of reflectances. Lambertian surfaces scatter the incoming light in all directions equally and they appear to have the same colour viewed from any position[15].

1.1 Colour Correction

Suppose we take a picture with a camera and then we wish to display it on a colour monitor. The raw acquired image cannot be used to directly drive a display. Rather, the image is transformed to a format suitable for display through a process called colour correction.

To understand colour correction, let us assume that the camera samples light like we do (for the purposes of this example, let us assume the camera curves equal those shown in 2c). How then do we transform the RGBs a camera measures to those that

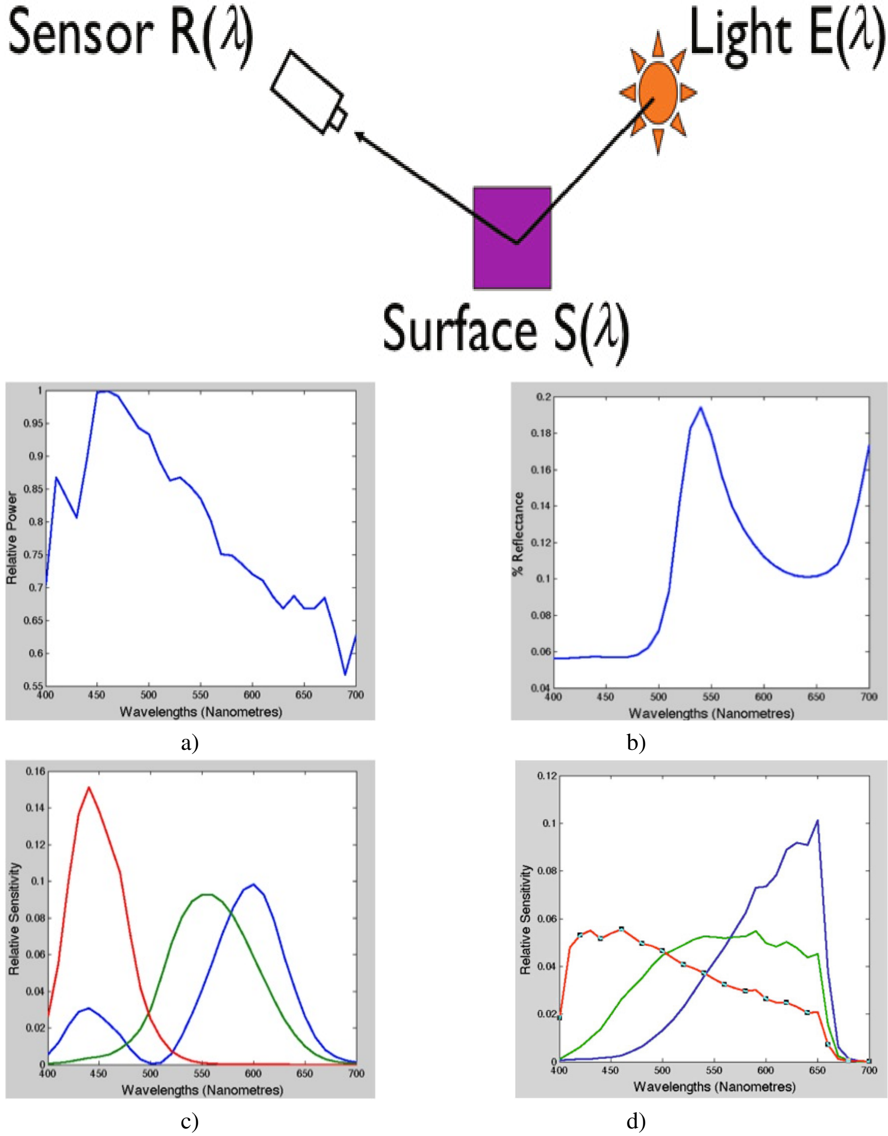

Fig. 2 Middle left shows the spectrum of a bluish light (power concentrated in the shorter wavelengths). The reflectance spectrum of a dark green surface is shown in 2b). Bottom left (2c) we show the XYZ colour matching functions. These are not the sensitivities of an actual camera rather they are reference curves useful for the standard communication of colour[27, 17]. Lastly, in 2d) we show the curves for a commercial camera (Sigms SD9). Notice how differently they sample light compared with the XYZ functions.

drive a display to arrive at the reproduction we seek (i.e. a displayed image that looks like the scene we took a picture of).

To answer this question let us assume that a monitor has 3 colour outputs with spectral power distributions in the short (or blue), medium (green) and long (red) parts of the visible spectrum. The camera response to each display colour is written as:

$$\rho_q^p = \int_\omega R_k(\lambda)P_q(\lambda)d\lambda \quad k \in \{R,G,B\} \quad q \in \{l,m,s\} \tag{2}$$

in (2) $P_q(\lambda)$ denotes the spectral output of the three channels of a colour display. Notice that both equations (1) and (2) are linear systems (double the light double the response). The import of this here is that the response of the camera red sensor to the long and medium display outputs turned on together - e.g. at 50% and 75% intensities - is simply the sum of the individual responses:

$$\begin{aligned}\rho_k^{0.5l+0.75m} &= \int_\omega R_k(\lambda)[0.5P_l(\lambda)+0.75P_m(\lambda)]d\lambda \\ &= 0.5\int_\omega R_k(\lambda)P_l(\lambda)d\lambda + 0.75\int_\omega R_k(\lambda)P_m(\lambda)d\lambda\end{aligned} \tag{3}$$

An implication of (3) is that the camera response to an arbitrary intensity weighting of the display outputs can be written as a matrix equation:

$$\begin{bmatrix}\rho_r \\ \rho_g \\ \rho_b\end{bmatrix} = \begin{bmatrix}\rho_r^l & \rho_r^m & \rho_r^s \\ \rho_g^l & \rho_g^m & \rho_g^s \\ \rho_b^l & \rho_b^m & \rho_b^s\end{bmatrix}\begin{bmatrix}\alpha \\ \beta \\ \gamma\end{bmatrix} \Rightarrow \underline{\rho} = M\underline{\alpha} \tag{4}$$

here α, β and γ vary the intensity of the colour channels (from 0 to 100% or minimum to maximum power). We are now in a position to solve for the display weights i.e. solve for the correct RGBs to drive the display. Denoting the 3-vector of responses in (1) as $\underline{\rho}$ then the correct image display weights $\underline{\alpha}$ (the values recorded in an image pixel) is calculated as:

$$\underline{\alpha} = M^{-1}\underline{\rho} \tag{5}$$

Equation (5) is called colour correction[26]. Note the 3x3 matrix M^{-1} is fixed for a given camera and display. Equation (5) is also the exact solution for colour matching (i.e. how we mix three primary lights to match an arbitrary test light).

However, in reality, it is never the case that a camera samples light like colour matching functions. Thus, the mapping which takes acquired RGBs to display outputs is approximate (and is solved for through regression[26]). We will return to this problem again in section 4 - see equation (15).

For historical reasons, displays typically have a non-linear transfer function. That is, the brightness output is (roughly) the square of the rgb driving the display. Thus the values stored in an image are the square-root of the weights calculated in (5). This process is called gamma correction.

Colour correction is a first order effect. The raw images recorded by a camera are not suitable for display, The effect of colour correction is illustrated in Figure 3.

Fig. 3 Left: image before colour correction. Right shows the corrected colours

2 Colour Constancy and Illuminant Estimation

The colours we see in the world do not depend on the colour of the illuminant. A white T-shirt looks white whether it is viewed in direct sunlight (yellowish colour of light) or in deep shadow (bluish light). Indeed, from an evolutionary point of view such colour constancy is clearly very desirable. As an example, in primate vision it has been proposed that colour is an important cue for judging the ripeness of fruit[24].

More generally, we do not expect the colour of the world to change as we move from one environment to another. Indeed, colour is often the primary designator we use in describing objects e.g. the red car or the green door. Yet, physically, the colour signal reflected from a surface may not, is typically not, the same as the object reflectance.

The idea that the colour we see was not a property of the spectrum of light entering the eye (the Newtonian view) is a relatively modern notion. Indeed, Edwin Land (the progenitor of Polaroid corporation) sparked a huge debate in the colour community when, in the 1950s, 60s and 70s, he proposed his Retinex Theory[19] of colour vision (to account for the phenomenon of colour constancy).

Simply, and perhaps somewhat obviously in hindsight, the Retinex theory proposes that the colours we see depend on the context in which we see them. Figure 4 (an example from Beau Lottos lab) illustrates this point. The same physical sample, viewed in two different illumination contexts, looks like it has a different colour.

In Figure 5, we show a colour constancy example from the Computer Vision literature[1]. Here the same object is captured under 4 different lights. It is remarkable how much the colour varies. It is evident then that raw colour does not correlate with object colour. Only if an object's colour *is* independent of illumination can it be used for recognition, indexing or tracking.

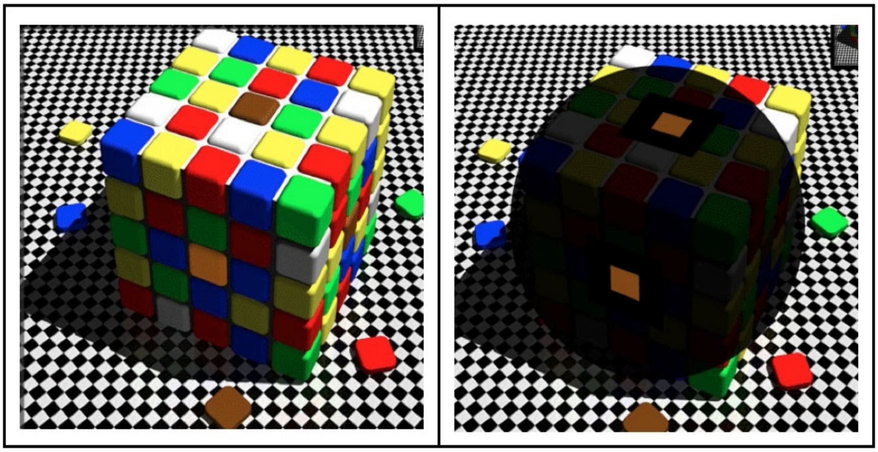

Fig. 4 The brown and orange surface chips on the top and front sides emit physically the same light. We see them as different colours as we perceive them both as a function of other colours in the scene and our physical interpretation of the scene. Clearly, we interpret the front cube face as being in shadow. The right hand panel of the figure (by using a black mask to remove the local context) demonstrates the chips reflect the same identical physical signal (from http://www.lottolab.org/).

Fig. 5 The same object viewed under 4 common lights. It is remarkable how much the colours of the object depends on the colour of the prevailing light[1]

However, getting the colour right is in itself of great interest as a problem in digital photography. We are very attuned (and highly critical) judges of the colours that look right or which look wrong when we look at photographs. Figure 6 shows an example of colour constancy from digital photography. To recover the image on the right we must estimate the illuminant colour and then remove its bias from the image. In photography, this process is often called 'White point adjustment'.

2.1 Estimating the Illuminant

Simple as Equation (1) is, it is in fact quite complex. Even assuming we know the spectral sensitivities of our camera, it is not immediately apparent that we can decouple and recover light $E(\lambda)$ from reflectance $S(\lambda)$. Indeed, each RGB supplies only 3 measurements which is not a propitious starting point for determining how we can solve the colour constancy problem.

To understand how we can, *practically*, solve the colour constancy problem, let us begin by making simplifying assumptions. First, let us assume that rather than recovering the spectrum of the light (or the spectrum of the surface reflectance) we instead wish only to recover the RGB of the light and the RGB of the surface. Second, let us assume that the camera measured RGB is the multiplication of the

Fig. 6 Left shows raw camera image, right after colour constancy (called white balance adjustment in photography). From http://en.wikipedia.org/wiki/Color_balance

RGB of the light and the RGB of the surface (this assumption is commonly made in computer graphics[3]). The **RGB model of colour image formation** is written as:

$$\rho_k^S = \int_\omega S(\lambda)R_k(\lambda)d\lambda \quad \rho_k^E = \int_\omega E(\lambda)R_k(\lambda)d\lambda$$
$$\rho_k^{E,S} = \rho_k^E \rho_k^S \qquad \text{(RGB model of image formation)}$$

Remarkably, these assumptions, with certain caveats, generally hold[6]. An important interpretation of ρ_k^S is that it is the colour of the surface viewed under a white uniform light $E(\lambda) = 1$. Subject to this observation, colour constancy can be thought of as mapping the rgbs measured in an image back to a reference lighting condition. That is, the colour constancy problem involves solving for ρ_k^S . Clearly, if we can estimate the illuminant (solve for the rgb of the light) then by dividing out we can estimate the surface colour.

2.2 The Maloney Wandell Algorithm

In 1986 Maloney and Wandell[22] presented perhaps the first formal treatments of the colour constancy problem. Their idea was that if light and surface were modelled by 3- and 2-dimensional linear models it would be possible to solve for colour constancy at a colour edge (i.e. given the rgb response of just two coloured surfaces).

Linear models of light and surface are written as

$$E(\lambda) = \Sigma_{i=1}^3 \varepsilon_i E_i(\lambda) \quad S(\lambda) = \Sigma_{j=1}^3 \sigma_i S_i(\lambda) \tag{6}$$

The intuition bbehind Maloney and Wandells approach is simple equation counting. Given two RGBs we have 6 measurements. Assuming the same light and two reflectances in a scene there are 2*2+3=7 unknowns. However, given the image formation equation (1) it is clear that we cannot distinguish between a bright light illuminating a dim scene or the converse. Thus, there are 6 equations and 6 unknowns to solve for. So, under the linear model assumptions (6), it is plausible we can estimate the RGB of the light given a pair of rgbs (for two different surfaces viewed under the same illuminant). Further, and crucially, 2- and 3-dimensional models for surface and light capture most of the variation found in typical reflectances and illuminations[22, 21].

So, how does plausibility translate into an actual algorithm? Well, here, we do not present their exact solution method (which is very general) but rather the equivalent algorithm that is simpler to implement[5] (which follows from the RGB model of image formation). We begin by observing that if reflectance has two degrees of freedom then this means that the RGB response of any surface under a single light must lie on a 2-dimensional plane. This idea is illustrated by the plane on the left of Figure 7.

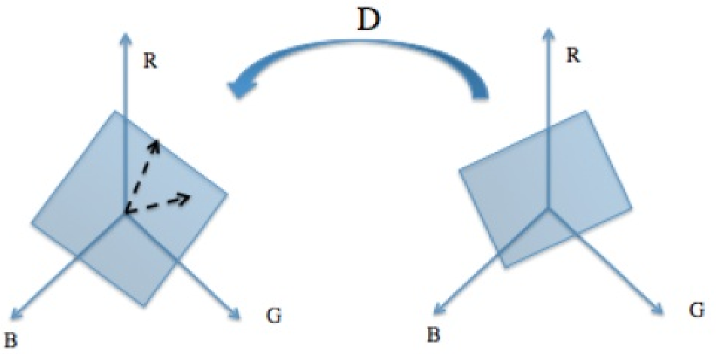

Fig. 7 Left shows the plane of RGBs measured by camera (spanned by the two actual measurements shown as dotted lines) under an unknown light. Right, the set of all camera measurements - also a plane - for a white reference light. The mapping taking right to left defines the colour of the unknown light.

The plane on the right shows the set of all possible camera responses for known white light reference conditions. Let us now rewrite the RGB model of image formation as the matrix equation:

$$\begin{bmatrix} \rho_R^{E,S} \\ \rho_G^{E,S} \\ \rho_B^{E,S} \end{bmatrix} = \begin{bmatrix} \rho_R^E & 0 & 0 \\ 0 & \rho_G^E & 0 \\ 0 & 0 & \rho_B^E \end{bmatrix} \begin{bmatrix} \rho_R^S \\ \rho_G^S \\ \rho_B^S \end{bmatrix} \Rightarrow \underline{\rho} = M\underline{\alpha} \tag{7}$$

Because there is a unique diagonal matrix mapping one plane to any other plane[5] then if we can find the diagonal matrix mapping the plane in the right of Figure 7 (the plane where rgbs lie under a white light) to the plane we observe for our RGB camera (the one on the left) then we have solved for the colour of the light. The diagonal matrix D and the colour of the light are one and the same thing).

Finally, by dividing out, we can solve for the colour of the surfaces.

$$\frac{\rho_k^{E,S}}{\rho_k^E} = \rho_k^S \tag{8}$$

Unfortunately, as elegant as this algorithm is, it actually delivers terrible colour constancy performance (the linear model assumptions do not hold sufficiently well). However, the idea that the colours we observe in an image provide prima facie evidence about the colour of the light is a good one: the reddest red RGB cannot be measured under the bluest light[12]. This idea is a the heart of many modern illuminant estimation algorithms.

Moreover, and more importantly, the tool of linear models has proven to be invaluable both to understanding complex problems and arriving at tractable algorithmic solutions.

2.3 Statistical Illuminant Estimation

Curiously, most illuminant estimation algorithms are based on a much simpler heuristic idea. Specifically, that the colour bias due to the illumination will manifest itself in summary statistics calculated over an image. If the colour of the prevailing light is yellowish then the mean of the image will be more yellow than it ought to be. So, it is reasoned, mapping the mean of the image so it is neutral (the mean of the red, green and blue channels are all equal) should deliver colour constancy. This approach is called grey-world colour constancy. It is easy to show that dividing by the mean is mathematically correct if the expected colour of every scene is gray[13].

In Lands Retinex theory[19] it was (effectively) argued that the maximum red, maximum green and maximum blue channels response is a good estimate of the colour of the light. Should every scene contain a white reflectance then this simple maxRGB approach will work. It would, for example work for the example shown in Figure 6. However, it is easy to find examples of images where neither max RGB nor grey world work very well.

It was observed[11] that the grey-world and maxRGB algorithms are simply the $p = 1$ and $p = \infty$ Minkoswki norms. Minkowski illuminant estimation is, assuming N pixels in an image, written as:

$$\rho_k^E = [\Sigma_{i=1} N [\rho_{k,i}]^p / N]^{(1/p)} \tag{9}$$

Remarkably, across a number of image datasets[10] a p-norm of 4 or 5 returns more accurate estimates of the illuminant than either max RGB or grey-world.

2.4 Evaluating Illuminant Estimation Algorithms

What do we mean if we say that one illuminant estimation algorithm works better than another? (i.e. that a p=4 Minkowski norm approach works best). In answering this question it is common to assume that the measured physical white point (the rgb of a white tile placed in the scene) is the correct answer. The angle between the estimated rgb of the illuminant and the actual true white point (the rgb for the white tile) is taken to be a measure of how accurate an estimate of the illuminant actually is.

The reference[14] provides a broad survey of a large number of illuminant estimation algorithms evaluated on a large number of data sets. The reported experiments convey two important messages. First illuminant estimation is a hard problem and even the best algorithms can fail (sometimes spectacularly). Second, progress on improving illuminant estimation is slow: its taken 30 years to provide a modest increase in performance.

Camera manufacturers remain interested in improving their white balance algorithms. Not only do they seek methods which work better, they are interested in identifying images that have multiple lights (sun and shadow) [16]. Modern algorithms have implemented Face detectors to aid estimation[23, 2].

3 Illuminant Invariants

The colours in two pictures of the same object observed from different viewpoints can be quite appear to be quite differen from each other and so, it is not always easy to find corresponding parts from two images that are the same (a necessary step to solve the stereopsis or shape from motion problems). Thus, in carrying out geometric matching it is common to seek geometric invariants i.e. features which do not change with a change in viewpoint. In David Lowes famous SIFT detector[20] features are sought that are invariant to scale and rotation. However, photometric invariance is also a useful property.

In the colour world, Swain and Ballard found that the distribution of colours in an image provides a useful cue for object recognition and object localisation[25]. Unfortunately, a precondition for that method to work is that image colour correlates with object colour. Yet, as we have seen in section 2, the same object will have a different image colour when viewed under differently coloured lights. Indeed, the same physical colour might have a range of intensities (i.e. shading) if the object has shape or the illumination intensity varies across a scene. Equally, if the colour of the light changes then the physical recorded colour will change as well. In either case, matching colour (e.g. to a database of images) without considering this problem can result in very poor recognition performance

Colour change due to changing lighting intensity (due to Lamberts law) and lighting colour is illustrated in Figure 8. The simple test image shown at the top of the figure is imaged under 3 coloured lights and from 3 different light positions. Directly below the image capture diagram we show the corresponding 3x3 image patches for the 9 imaging conditions. It is apparent that there is a remarkable variety of different coloured images resulting from the same physical scene.

If we think of this simple colour edge as a region of interest in the image, then Figure 8 informs us the edge RGBs change when the viewing condition changes. Photometric normalisatiion methods seek simple algebraic formulae or algorithms for canceling out this image variation.

3.1 Intensity Invariance

We can normalise for the lighting geometry of the scene - the intensity variation of the object RGBs due to shading and the position of the light source - simply, by dividing each RGB by its magnitude:

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} R/(R+G+B) \\ G/R+G+B) \\ B/R+G+B \end{bmatrix} \tag{10}$$

Clearly [R G B] and [kR kG kB] have the same normalised output. Note also post-normalisation that b=1-r-g. That is, by removing intensity the colour at each pixel is parameterised by just two numbers. The tuple (r,g) is sometimes called the chromaticity of the RGB.

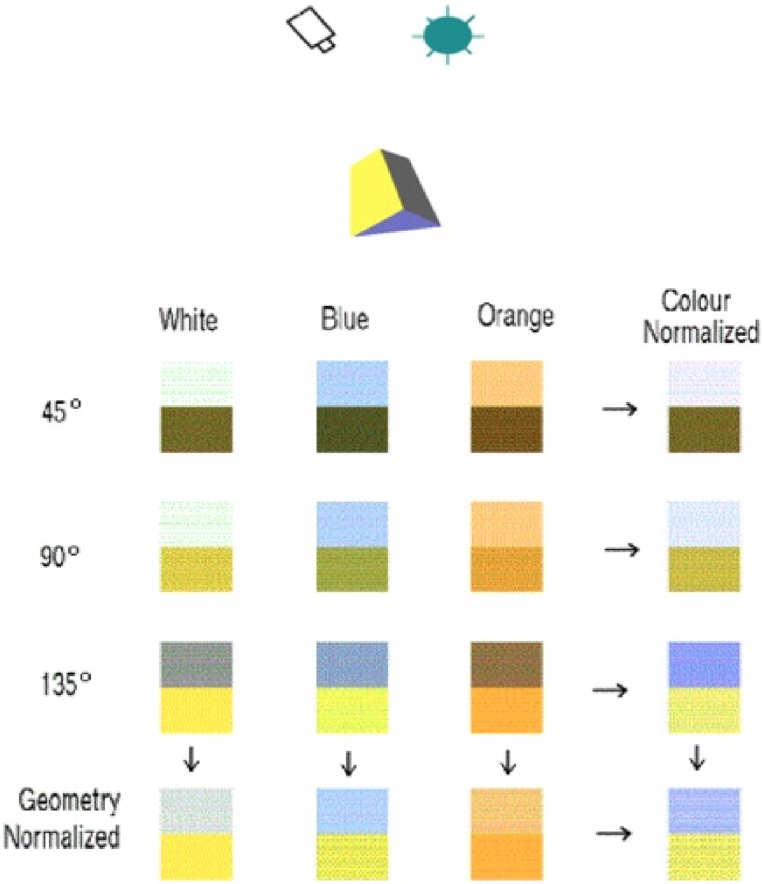

Fig. 8 Top a simple wedge is viewed under a light source. The light source can be one of 3 different colours and be place at 3 different positions. The upper 3x3 image outputs show the range of recorded colours for the wedge for the 3 lighting positions and 3 light colours. The last row shows the output of intensity normalisation and the last column the result of colour normalisation. The patch bottom right is the output of colour and intensity nomalisations carried out iteratively.

The effect of intensity normalisation is shown in the bottom row of Figure 8. It is clear when ony the intensity varies that intensity normalisation suffices to make the colour images the same.

3.2 Colour Invariance

We achieve invariance to the colour of the light in a similar way though now we work not with the RGB at each pixel but rather with all the pixel values in a single colour channel. The RGB of the ith pixel is made invariant to the colour of the light by calculating:

$$\begin{bmatrix} R_i \\ G_i \\ B_i \end{bmatrix} = \begin{bmatrix} R_i/\mu(R) \\ G_i/\mu(G) \\ B_i/\mu(B) \end{bmatrix} \tag{11}$$

In (11) we divide each pixel value by the average of all the pixels (in the same colour channel). Because we are adopting the RGB model of image formation (from section 2) the illuminant colour must appear in both the numerator and denominator of the right hand side of (11) and, so, must cancel.

The effect of this illuminant normalisation is shown in the right hand column of Figure 8. If only the illuminant colour changes then (7) suffices to normalise the colours (all the images in the same row have the same output colours).

However, by dividing by the mean is similar to the grey-world colour constancy algorithm discussed in section 2 (we divide by the p=1 norm of Eq. (9)). The only difference is that in colour constancy research we wish the normalised colours to look correct. The bar is set lower for colour invariance: it suffices that same object viewed under different lights is normalised to the same (albeit often false) image colours.

3.3 Comprehensive Normalisation

Remarkably, in[8] it was shown that if we iteratively calculate (10) (intensity invariance) and then (11) (colour invariance) then this process converges to an output that is independent of lighting geometry and light colour. The 9 input images in Figure 8 all converge to the same single output shown bottom right. Importantly, colour normalisation (intensity, colour and comprehensive) has been shown to be useful for object recognition and image indexing[8].

3.4 Colour Constancy at a Pixel

Let us suppose that we could calculate intensity and colour invariance at a pixel i.e. at an image containing a single RGB pixel. We cannot do this using comprehensive normalisation. Indeed, any input pixel will, by iteratively applying (10) and (11), result in the triple (1,1,1) i.e. we get invariance in a trivial sense (all input RGBs map to the same output colour).

In fact under typical illuminant conditions it is, remarkably, possible to find a single scalar value that is independent of the intensity and is independent of the colour of the light. To see this, we begin by adopting an alternative chromaticity definition:

$$\begin{bmatrix} r \\ b \end{bmatrix} = \begin{bmatrix} R/G \\ B/G \end{bmatrix} \tag{12}$$

Note for all RGBs G/G=1 and so we ignore this term ($[r \;\; b]^t$] encodes RGB up to an unknown intensity) The formula in (12) is useful because it implies that[4]:

$$\begin{bmatrix} r^{E,S} \\ b^{E,S} \end{bmatrix} = \begin{bmatrix} r^E & 0 \\ 0 & b^E \end{bmatrix} \begin{bmatrix} r^S \\ b^S \end{bmatrix} \tag{13}$$

i.e. the chromaticity response is a simple multiplication of the chromaticity of the light and the chromaticity of the reflectance. The diagonal model of (3) for RGBs holds for spectral band ratios too.

In[7] the following experiment is carried out. A picture of a Macbeth colour checker, shown in the top of Figure 9, is captured. There we mark 7 basic colours: Red, Orange, Yellow, Green, Blue, Purple and White patches. We now take pictures of these patches under 10 different typical lights ranging from indoor yellow tungsten to white cloudy day light to blue sky i.e. a range of typical lights For each of the 7 patches we plot the spectral band ratios on a log-log scale. We plot these results in the graph shown at the bottom of Figure 9.

Note, that as the illumination changes the spectral band ratios sweep out a line on the log-log plot. More importantly, the slope of each line is the same for all surface colours but the intercept varies. Clearly then, the intercept can be used as an scalar measure of reflectance which, empirically at least, does not vary with illumination.

Suppose we take a picture of the world where there are cast shadows. The light colour in and out of the shadow are different (the light for the shadow region is much bluer). Assuming that the data shown in Figure 9 holds in general (implies that there is an intrinsic reflectance invariant calculable at a pixel) then we should be able to simply - trivially - remove shadows.

First, we remember we know the slope of the lines in Figure 9. Second, for a given RGB we calculate its log spectral band ratio coordinates. Then we can calculate the intercept (with either the x- or y- axis). We then take the scalar image of intercepts and recode as a greyscale image. In Figure 10 we show the outputs of applying this methodology. The shadows magically disappear.

In general, the grey-scale invariant image, that is independent of the colour of the light and intensity, conveys salient information and has been shown to be useful in applications ranging from scene understanding[9], to object recognition[7] to tracking[18].

The reader will, no doubt, be curious as to why spectral band ratios on a log log plot look as they do in Figure 9. Well, it turns out that most lights (at least in

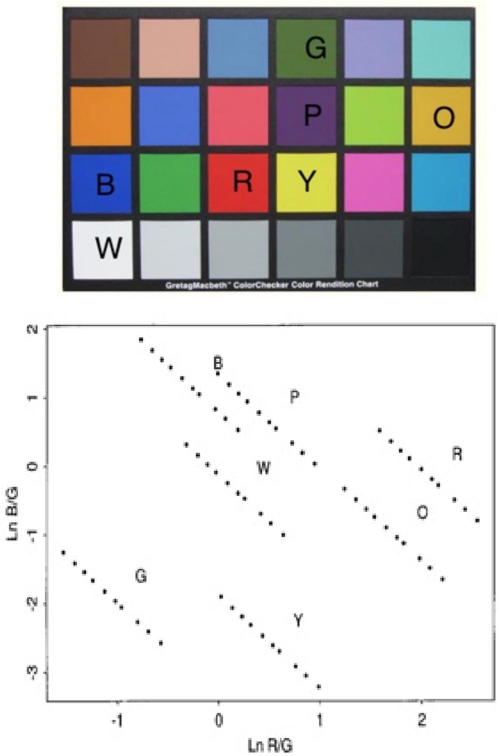

Fig. 9 Top, Macbeth colour checker. Bottom, log spectral band ratios for marked surfaces for 10 lights (the dots shown).

terms as how they project to form RGBs) can be thought of as Planckian black-body radiators. The colour of black-body radiators is parameterised by one number: temperature. Because of the form of the mathematical equation that models black-body radiation, it turns out that the log-log plot must look like Figure 8. For a description of why this is the case, the reader is referred to [7].

4 Computer Vision and Colour Perception

There are many applications where we would like a machine vision to see like we do. Unfortunately, we do not have good operational models of our own vision system so we, instead, seek to equip machine vision systems with simpler -though, still useful - competences. In the applied colour industries there are specialised measurement devices that attempt to numerically gauge the similarity of colour pairs. Figure 10 illustrates the industrial colour difference problem. Perceptual relevance in machine vision is sometimes taken to mean that the vision system might be used for colour difference assessment.

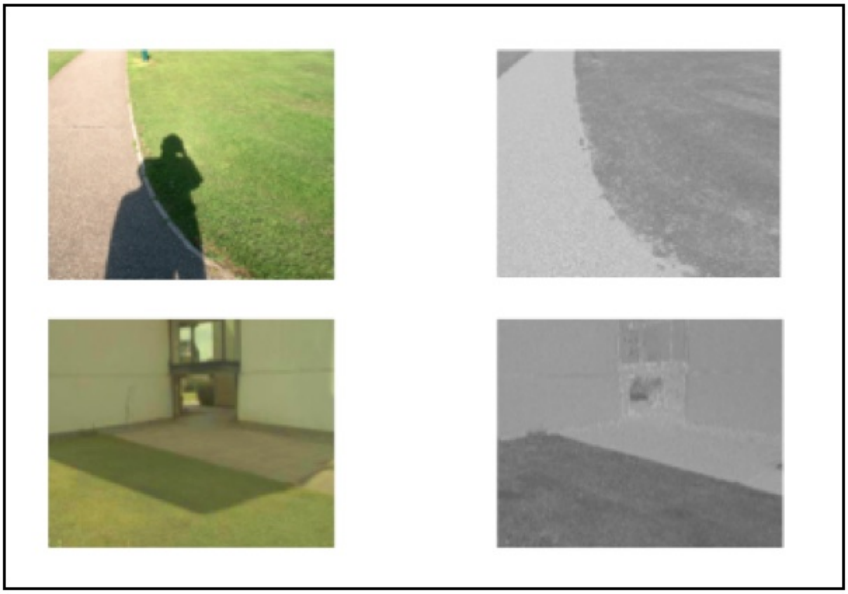

Fig. 10 Left, Raw camera image. (As described in the text), illuminant invariant calculated in the right: the shadows magically disappear.).

Assuming we had camera sensors with sensitivities the same as those shown in 2c (not generally the case) and if we also knew the colour of the light then there are standardised formula[27] for mapping camera measurements to, so called, Lab values. Euclidean distance in Lab space approximately account for the perceived difference between stimuli. Specifically, a distance of 1 correlates with a just noticeable difference. We recapitulate the CIE Lab[27] equations below:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x/x_n \\ y/y_n \\ z/z_n \end{bmatrix}$$

$$\begin{bmatrix} L \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 & 1160 & \\ 500 & -500 & 0 \\ 0 & 200 & -200 \end{bmatrix} \begin{bmatrix} x^{1/3} \\ y^{1/3} \\ z^{1/3} \end{bmatrix} + \begin{bmatrix} -16 \\ 0 \\ 0 \end{bmatrix} \tag{14}$$

Here x, y and z are the responses of a camera with sensitivities 2c. The triple (x_n, y_n, z_n) is the camera response to the illuminant. The Lab formula was derived [27] by fitting psychophysical data (real colour difference judgements made by people). In Figure 11, the 'Delta E' colour difference is about 9 indicating a visually significant colour difference.

To use a vision system for colour grading when the sensitivities are not like those in 2c (the actual sensitivities of a commercial camera are shown in 2d) then this

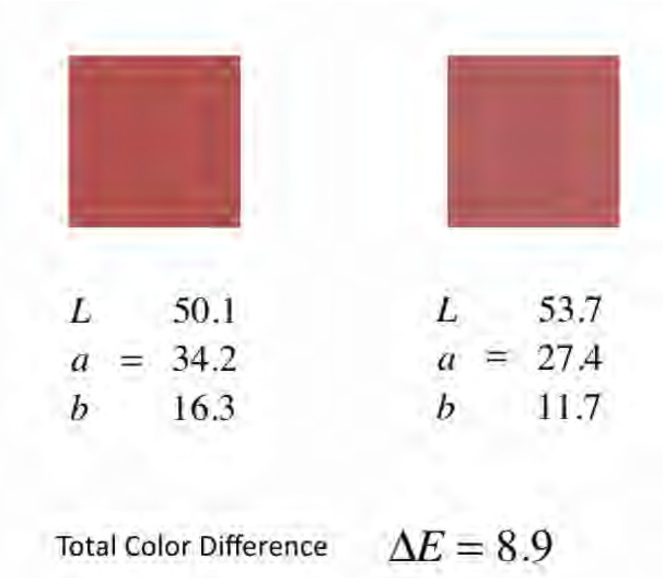

Fig. 11 Two similar colour patches are measured and their colours summarised according to three (L, a and b) coordinates. The Euclidean distance between the triples correlates with perceived colour difference.

means that the actual camera colours must be mapped to approximate corresponding xyzs. This mapping is often solved for as a simple linear transform:

$$\min_{T} ||R_{N\times 3}T_{3\times 3} - X_{N\times 3}|| \tag{15}$$

where, $R_{N\times 3}$ above is a set of measured RGBs for a calibration target (e.g. of the kind shown on the left of Figure 9). $X_{N\times 3}$ are the corresponding measured XYZs. Once we have solved for the best regression matrix T we can use a camera to measure arbitrary scenes and calculate Labs according to the above formulae.

4.1 Colour Difference Formulae and Computer Vision (a Cautionary Remark)

That we might carry out a simple calibration and recover approximate Lab values is all well and good if we wish to carry out colour measurement. But, the reader should be aware that Euclidean distance on CIE Lab values only models small colour differences. If a pair of colours are compared and found to be (say) 20 units apart, this means almost nothing at al. i.e. we cannot measure the perceived closeness of red and green using Lab colour differences.

Unfortunately, in computer vision researchers sometimes assume that once we transform to Lab then we have somehow carried out a 'perception transform'. It is naively proposed that, simply, by transforming to Lab space we can claim perceptual relevancy. One cannot. It is often quite inappropriate to claim that a given tracking, recognition, object finding algorithm in Lab space says anything much about our own perception or how we ourselves solve these problems.

5 Conclusion

Colour is a huge field and is studied in physics, computer science, psychology and neuroscience (among other fields). While great progress has been made in the last 100 years, colour is still far from a solved problem. That this is so, accounts, in part, for colour sometimes being used wrongly in computer vision.

In this short primer we have tried to introduce the reader to colour in computer vision. We have explained how camera RGBs are mapped to image colours that drive the display (colour correction). Removing colour bias due to illumination (colour constancy) is perhaps the most studied aspect of colour in computer vision. Solving for colour constancy is essential if colour is to be used as an absolute correlate to reflectance. However, relative measures of colour - functions of proximate pixels - can be used to cancel illumination effects (colour invariance). Remarkably, we can calculate a grey-scale invariant at a pixel which cancels the colour and intensity of the light (with respect to which shadows, magically, disappear).

The assumption that a camera system might easily play a surrogate role for our own vision system is a seductive idea. The good news is that, yes, colour cameras can be used for colour measurement. The bad news is that colour measurement does not really say anything very profound about how we see

Acknowledgements. The author gratefully acknowledges the support of EPSRC grant H022236,

References

1. Barnard, K., Martin, L., Funt, B., Coath, A.: A data set for color research. Color Research and Application 27(3), 147–151 (2002), `http://dx.doi.org/10.1002/col.10049`, doi:10.1002/col.10049
2. Bianco, S., Schettini, R.: Color constancy using faces. In: CVPR, pp. 65–72 (2012)
3. Borges, C.: Trichromatic approximation for computer graphic illumination models. Computer Graphics 25, 101–104 (1991)
4. Finlayson, G.: Color in perspective. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1034–1038 (1996)
5. Finlayson, G., Drew, M., Funt, B.: Color constancy: Generalized diagonal transforms suffice. J. Opt. Soc. Am. A 11, 3011–3020 (1994)
6. Finlayson, G., Drew, M., Funt, B.: Spectral sharpening: Sensor transformations for improved color constancy. J. Opt. Soc. Am. A 11(5), 1553–1563 (1994)
7. Finlayson, G., Hordley, S.: Color constancy at a pixel. JOSA-A 18(2), 253–264 (2001)
8. Finlayson, G.D., Schiele, B., Crowley, J.L.: Comprehensive colour image normalization. In: Burkhardt, H.-J., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1406, pp. 475–490. Springer, Heidelberg (1998)

9. Finlayson, G.D., Hordley, S.D., Lu, C., Drew, M.S.: On the removal of shadows from images. IEEE Trans. Pattern Anal. Mach. Intell. 28(1), 59–68 (2006)
10. Finlayson, G.D., Rey, P.A.T., Trezzi, E.: General p constrained approach for colour constancy. In: ICCV Workshops, pp. 790–797 (2011)
11. Finlayson, G.D., Trezzi, E.: Shades of gray and colour constancy. In: Color Imaging Conference, pp. 37–41 (2004)
12. Forsyth, D.: A novel algorithm for color constancy. Int. J. Comput. Vision 5, 5–36 (1990)
13. Gershon, R., Jepson, A., Tsotsos, J.: Ambient illumination and the determination of material changes. J. Opt. Soc. Am. A 3, 1700–1707 (1986)
14. Gijsenij, A., Gevers, T., van de Weijer, J.: Computational color constancy: Survey and experiments. IEEE Transactions on Image Processing 20(9), 2475–2489 (2011)
15. Horn, B.: Robot Vision. MIT Electrical Engineering and Computer Science Series. MIT Press (1986)
16. Hubel, P.M.: The perception of color at dawn and dusk. In: Color Imaging Conference, pp. 48–51 (1999)
17. Hunt, R.: Measuring Colour, 3rd edn. Fountain Press (2001)
18. Jiang, H., Drew, M.S.: Shadow resistant tracking using inertia constraints. Pattern Recognition 40(7), 1929–1945 (2007)
19. Land, E.: The retinex theory of color vision. Scientific American, 108–129 (1977)
20. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)
21. Maloney, L.: Evaluation of linear models of surface spectral reflectance with small numbers of parameters. J. Opt. Soc. Am. A 3, 1673–1683 (1986)
22. Maloney, L., Wandell, B.: Color constancy: a method for recovering surface spectral reflectance. J. Opt. Soc. Am. A 3, 29–33 (1986)
23. Montojo, J.: Face-based chromatic adaptationfor tagged photo collections (2009)
24. Regan, B., Julliot, C., Simmen, B., Vinot, F., Charles-Dominique, P., Mollon, J.: Frugivory and colour vision in alouatta seniculus, a trichromatic platyrrhine monkey. Vision Research 38(21), 3321–3327 (1998), `http://www.sciencedirect.com/science/article/pii/S0042698997004628`, doi:10.1016/S0042-6989(97)00462-8
25. Swain, M., Ballard, D.: Color indexing. International Journal of Computer Vision 7(11), 11–32 (1991)
26. Vrhel, M.J., Trussell, H.J.: The mathematics of color calibration. In: ICIP (1), pp. 181–185 (1998)
27. Wyszecki, G., Stiles, W.: Color Science: Concepts and Methods, Quantitative Data and Formulas, 2nd edn. Wiley, New York (1982)