Corrected-Moment Illuminant Estimation

Graham D. Finlayson
School of Computing Sciences
The University of East Anglia
graham@cmp.uea.ac.uk

Abstract

Image colors are biased by the color of the prevailing illumination. As such the color at pixel cannot always be used directly in solving vision tasks from recognition, to tracking to general scene understanding. Illuminant estimation algorithms attempt to infer the color of the light incident in a scene and then a color cast removal step discounts the color bias due to illumination. However, despite sustained research since almost the inception of computer vision, progress has been modest. The best algorithms – now often built on top of expensive feature extraction and machine learning – are only about twice as good as the simplest approaches.

This paper, in effect, will show how simple moment based algorithms – such as Gray-World – can, with the addition of a simple correction step*, deliver much improved illuminant estimation performance. The corrected Gray-World algorithm maps the mean image color using a fixed (per camera) 3x3 matrix transform. More generally, our moment approach employs 1st, 2nd and higher order moments - of colors or features such as color derivatives - and these again are linearly corrected to give an illuminant estimate. The question of how to correct the moments is an important one yet we will show a simple alternating least-squares training procedure suffices. Remarkably, across the major datasets – evaluated using a 3-fold cross validation procedure – our simple corrected moment approach always delivers the best results (and the performance increment is often large compared with the prior art). Significantly, outlier performance was found to be much improved.*

1. Introduction

There are 3 important physical variables to consider in image formation. First, there are the objects in the scene and their surface reflectance properties. Second we must consider the illumination or illuminations under which a scene is viewed. Lastly, the spectral characteristics of the sensors are an important variable. The interaction of surface, light and sensor can be elucidated in a single simple equation:

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

Equation (1) teaches that the light with spectral power distribution $E(\lambda)$ strikes a surface which reflects light on a per wavelength basis according to its reflectance function $S(\lambda)$. Then, an integrated response is calculated for each of 3 sensor classes $R_k(\lambda)$ (usually, short-, medium- and long-wave sensitive mechanisms or R, G and B). The integral is taken over the visible spectrum ω. Remarkably, this equation though simple, is a pretty accurate – first order – model of image formation; i.e. if you can measure the spectral functions involved then evaluating (1) numerically, synthetically generating an R, G and B, will often predict the actual camera response rather well[48]. Of course (1) does not account for phenomena such as specular highlights[43] or surface roughness[38].

In Figure 1 we illustrate the color constancy problem. The task of an illuminant estimation algorithm is to infer that the prevailing light on the left is bluish and then the blue cast is removed to give the correct (i.e. the image conveys our percept of the color of the scene) image in the middle. Notice the green leaves now look green. Of the two constituent parts of the color constancy problem – removing the color cast and illuminant estimation – illuminant estimation is the most difficult. Indeed, if the illuminant color is known (or correctly estimated) then it is straightforward to discount[16, 47, 20].

In illuminant estimation it suffices to solve for the R, G, B response for the illuminant defined as:

$$\rho_k^E = \int_\omega E(\lambda)R_k(\lambda)d\lambda \tag{2}$$

Often[7], but not always[39], (2) is used in tandem with the RGB model of image formation. Here, the surface response is written as:

$$\rho_k^S = \int_\omega S(\lambda)R_k(\lambda)d\lambda \tag{3}$$

Figure 1. Left shows raw camera image, middle after correcting for the correct illuminant and right when the Gray-World estimate is used. Left and middle panels from http://en.wikipedia.org/wiki/Color_balance

and the response to light and surface combined is calculated as:

$$\rho_k^{E,S} = \rho_k^E \rho_k^S \qquad \text{(4: RGB model of image formation)}$$

Even when (4) is found not to apply directly, it generally holds in some sensor basis (for a linear combination of the sensors[13]). Given the simple structure in (4) it is now apparent why color cast removal is straightforward. Given an accurate estimate of the RGB of the light we divide $\rho_k^{E,S}$ by ρ_k^E to recover ρ_k^S. The color of the surface, by definition, is the color of the surface seen under white light.

Assuming a single illuminant color across a scene and a corresponding image with N pixels then there are $3N + 3$ unknowns (N surfaces and 1 light) but only the $3N$ RGB knowns. Actually, it turns out it is impossible to recover the magnitude of the RGB of the light[17] (a bright scene dimly lit is indistinguishable from a bright light impinging on darker surfaces) so the number of unknowns reduces to $3N + 2$. Even though the degree of ill-posedness for illuminant estimation is markedly less than for some vision problems (e.g. shape from shading and optical flow [4]), it has proven to be a very hard problem to solve.

The simplest algorithm for illuminant estimation is the so-called Gray-World approach. Here the illuminant estimate is simply the average[9] (or weighted average [27]) of the Red, Green and Blue image planes. The approach is called Gray-World since if the average surface color in the scene is gray then the mean RGB must – following from the RGB model of image formation – be the same color as the illuminant. Intuitively, the Gray-World algorithm has some merit since if we have an image under a white light (R=G=B=1) and move to a chromatic light (a yellow light, R=1, G=.8, B=.1) then clearly the mean image color $[\mu_R \ \mu_G \ \mu_B]$ is mapped to $[\mu_R \ .8\mu_G \ .1\mu_B]$, i.e. the *first order moments* are strongly biased by illumination). Under a very chromatic yellow light the mean of the image will almost certainly be yellow.

However, typical lights are less chromatic and range from bluish to whitish to yellowish (and on occasion into the oranges). When light is less chromatic it is less obvious that the mean RGB correlates directly with the light color. Indeed, when there a single predominant surface color in a scene – e.g. the leaves in Figure 1 – then this greeness will be reflected in the average RGB. Discounting the illuminant by dividing by the mean must, incorrectly, map green toward gray (as shown in the right panel of Figure 1).

In this paper we propose that the 1st order moments are useful (indeed intrinsic) but that they need to be corrected to provide a useful illuminant estimate. Our corrected Gray-World is simple and it is simply written as:

$$\underline{\hat{\rho}}^E \approx C\underline{\mu} \qquad (5)$$

Here $\underline{\mu}$ denotes the image average RGB vector. C is a 3×3 matrix and $\underline{\hat{\rho}}^E$ is the corrected illuminant estimate. We consider two extensions to (5). First we allow the estimated illuminant to be a correction from 1st, 2nd and higher order moments. In the general case we have M simple statistical measures and these are related to our illuminant estimate using a $M \times 3$ correction matrix. Crucially, 2nd and higher moments include 'cross' color channel terms. In concert with previous work[46] we also find that the moments calculated for a 'color edge' image to be particularly useful for illuminant estimation. Of course, as a second extension, we can also use (5) to 'correct' illuminant estimates of other existing algorithms using a correction matrix. However, this latter modification is not an important focus of this paper (not least because we found it did not deliver as good estimation as the moment-based-approach we develop here).

The question of how we solve for C is important. It is our contention that because we cannot recover the intensity of the light we must incorporate this reality in the solution strategy. Indeed, we must solve for C that best maps the *orientation* of $\underline{\mu}$ toward the orientation of $\underline{\rho}^E$. Significantly, this changes the optimisation from a simple closed form least-squares problem to an alternating least-squares procedure.

Sections 2, 3 and 4 respectively provides the background to our work, describes our corrected moment approach and presents experimental validation. The paper concludes in section 5.

2. Related work

The majority of the statistical moments previously proposed as illuminant estimates can be summarised in a single equation[46]:

$$\left(\int \left| \frac{\delta^n \underline{\rho}(\mathbf{x})}{\delta \mathbf{x}^\mathbf{n}} \right|^p d\mathbf{x} \right)^{1/p} = k\underline{\hat{\rho}}^E_{n,p,\sigma} \qquad (6)$$

Here, we have an RGB color image where the camera response at image location $\mathbf{x}$ is the 3-vector $\underline{\rho}(\mathbf{x})$. The image

can be smoothed with a Gaussian averaging filter with standard deviation σ pixels and the smoothed image is differentiated with an order n differential operator (0 order means no differentiation). We then take the absolute Minkowski p-norm average[22] over the whole image. This results in a singe illuminant estimate $\hat{\underline{\rho}}^E$. Note k is an unknown scalar drawing attention to the fact that it is not possible to recover the true magnitude of the prevailing illuminants. A simple extension to (6) is to allow arbitrary linear filters to be used (rather than differential operators)[11].

Building on top of simple summary statistics we can add additional calibration information such as the gamut of possible camera RGBs under a known reference light[24] and the gamut of typical light colors [17]. In gamut mapping illuminant estimation the idea is to find a plausible illuminant estimate which – after the color cast has been removed e.g. by dividing by the RGB of the light – returns a set of image colors inside the reference gamut. Gamut mapping tends not to return a single estimate but rather a set of plausible lights. A single answer is typically selected by using a statistical argument[18]. Gamut mapping can also work with color derivatives[31].

Using only an illuminant constraint it is also possible to reformulate (6) to find the plausible illuminant RGB which best summarises the colors (or color derivatives) in an image in the Minkoswki sense[23]. More simply, it is useful to know what the average image color actually is if it is not gray. In 'Database Gray world' the illuminant estimate is defined to be a fixed diagonal matrix transform from the mean RGB[1]. We effectively generalise this method in the next section.

A natural extension to the gamut idea is to add probablilities into the mix. If we have access to information about either or both the likelihood of RGBs under different lights or the likelihood of the lights themselves then we can use a maximum likelihood approach to illuminant estimation[8, 19]. Of course it is also possible to formulate illuminant estimation using other parametric formulae such as KL-divegence[41] or in a voting framework[42]

As machine learning has emerged as an important tool for computer vision in general so too it has proven useful in the context of illuminant estimation. Exemplar approaches including selecting algorithms to apply based on an underlying analysis of the spatial characteristics in a scene[29, 5] and the recent idea of finding useful features in images in an illuminant invariant way and then using knowledge of how these features can be mapped to a reference light to solve for the illuminant[35]. Learning based strategies work very well and they are leading the field in terms of the illuminant estimation performance they can deliver.

Wider surveys of color constancy – which include reviews of physics-based approaches[21, 45] are not considered here – can be found in [1, 2, 32]. Equally, our algorithm is developed for 'general' scenes and so do not consider approaches predicated on identifying known objects – e.g. faces – in images[37, 6].

3. Moment based illuminant estimation

The premise of our method, simply stated, is that statistical moments are useful for illuminant estimation. A simple worked example illustrates our approach (and also why it shows promise). Following the protocol set forth in[1] we randomly selected 8 reflectances (from a database of 1995 reflectances) and then chose 1 illuminant from a set of 102 typical lights and with the CIE XYZ color matching curves used (1) to generate an entirely synthetic image. We then repeated this experiment 1000 times. An XYZ triplet correlates with a camera RGB and the x- and y-chromaticities are defined as x=X/(X+Y+Z) and y=Y/(X+Y+Z). This chromaticity representation, because it also relates to how we see[49], is a convenient way to talk about color in an intensity independent way (if we know x and y we can recover [X Y Z] up to an unknown scalar). Usefully, we can assign known colors to each point in the chromaticty diagram as shown in the top panel of Fig. 2.

In the middle panel in blue crosses we show the chromaticities of the Gray-World illuminant estimates (mean colors) for our 1000 Mondrians. The dotted black convex hull, delimits the set of all chromaticities for all 1000 of our Mondrians. The colors outside the black line correspond to 'aperture' colors: colors that can be made from quite narrow band lights but which do not occur in nature as light reflecting from surfaces. In red circles we also plot the chromaticities for the actual 102 illuminants. It is remarkable how given the real-estate that the image colors might occupy – anywhere in the black convex hull – that the means lie so close to the actual plausible illuminants. For this test, the mean appears to correlate *roughly* with the actual illuminant.

Let us now solve for a 3x3 correction matrix C that maps each of our 1000 mean rgb triples as close as possible to the correct rgb in a intensity independent way (see next 3.3 for how we do this). The actual correction matrix we use is shown top right of the bottom panel of Fig. 2. We now draw the same diagram as before but map the means of the Mondrians by our correction matrix. The blue crosses are the corrected-mean chromaticities. Notice how the corrected means now sit almost entirely above the set of plausible answers to illuminant estimation. Even if the corrected Gray-World estimate is wrong it must correspond to a 'plausible' illuminant estimate.

3.1. Color moments which scale with intensity

There are many moments which have been used in illuminant estimation. These include the average of the R, G and B channels or the average of the absolute values of

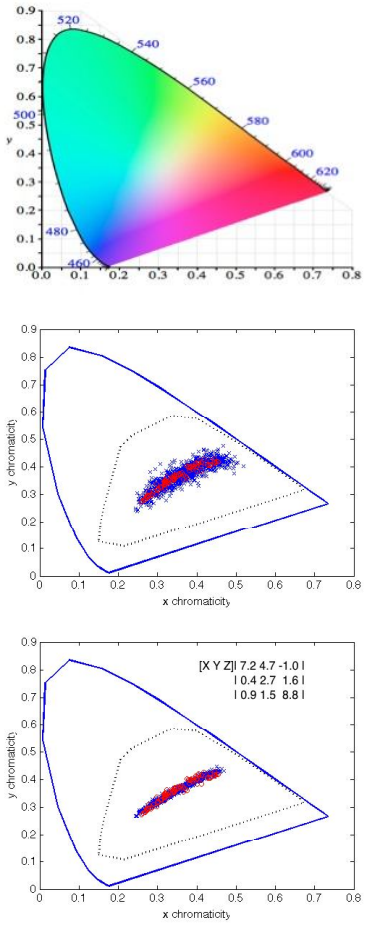

Figure 2. Top: visulization of the colors we can see in the CIE xy chromaticity dugram. Middle and bottom show the Gray-World estimates before and after a 3x3 matrix correction.

derivative images. The simple average has the important advantage that as the image data scales – e.g. due to a change of exposure or light brightness – so the moments scale too:

$$E(\alpha\rho_k) = \alpha E(\rho_k) \tag{7}$$

This is a crucial property in the context of illuminant estimation as it implies two images related by a single overall scale factor will return the same illuminant estimate.

Now, let us consider extending the basic corrected Gray-World algorithm to apply to a larger set of moments. While there are many moment-type expansions of multidimensional data (e.g. variance and covariance) it is important to maintain this 'intensity scaling' property. It is natural to use the p-norms as our moments since clearly $E(\alpha\rho_k^p)^{(1/p)} = \alpha E(\rho_k^p)^{(1/p)}$. To these per-channel moments it is natural add 'cross moments'. The six 2nd order cross moments of an image comprise:

$$\begin{bmatrix} E((R^2))^{0.5} & E(G^2)^{0.5} & E(B^2)^{0.5} & \\ \cdots & E(RG)^{0.5} & E(RB)^{0.5} & E(GB)^{0.5} \end{bmatrix} \tag{8}$$

We can of course compute higher order moments too. For a monomial of degree M and 3 variables the number of terms (moments) is equal to $\frac{(M+2)!}{2M!}$ (10 terms of order 3 and 15 order 4). The total number of moments if we calculate all moments of the Mth order and below is equal to

$$\#moments = \Sigma_{m=1}^{M} \frac{(m+2)!}{2m!} \tag{9}$$

and each mth order term is written as

$$\begin{aligned} p_{uvw} &= [\frac{\Sigma_{i=1}^{N} R_i^u G_i^v B_i^w}{N}]^{1/(u+v+w)} \\ u+v+w &= m \ , \ \ u,v,w \geq 0 \end{aligned} \tag{10}$$

In this paper we use the notation $\underline{p}_m^t$ to denote a row vector comprising m moments According to (9) there are respectively 3, 9 and 19 moments for all the monomials of orders less than equal to 1, 2 or 3. We remark that there are other ways of calculating moments which scale with intensity. In analogy to Hu-type spatial moments combinations such as [34], $E(R^4)/E(G^3)$ also scale linearly with intensity.

3.2. Corrected-moment illuminant estimation

The general corrected-moment approach to illuminant estimation is calculated as:

$$[\underline{\hat{\rho}}^E]^t = \underline{p}_m^t C_{m\times 3} \tag{11}$$

where $C_{m\times 3}$ denotes a $m \times 3$ regression matrix. Given Equation (11) it is apparent why we wish moments to scale with image intensity. Since if $\underline{p}_m^t \rightarrow \alpha\underline{p}_m^t$ then our illuminant estimate is clearly $\alpha\underline{\hat{\rho}}^E$, i.e. the magnitude of the estimate changes but its orientation remains fixed. The moments that are used can also be calculated from an RGB image post linear filtering (e.g. using a first and second derivatives). In section 4 we find that the moments of 1st order x- and y- derivative images (we call these **color edge** moments) to be particularly useful.

3.3. Finding an intensity independent correction matrix

Let us assume we have the correct answer, the rgb of the light, for a set of N images. We place each rgb in the row of a $N \times 3$ light matrix L. For each of the corresponding images we calculate their moment vectors and place these moments in the rows of the matrix P. P has N rows and may contain 3, 9 or 19 columns for color or color edge moments Remembering that we cannot solve for the magnitude of the illuminant, we propose to find C by minimizing:

$$\min_{C,d_i} \Sigma_{i=1}^{N} ||d_i P_i C - L_i||^2 \quad (12)$$

where d_i is a scalar and P_i is the ith row of P and L_i is the ith row of L. We solve for the $M \times 3$ correction matrix C. Of course we might achieve a slightly better fit by adding in an additional 3×1 offset and minimizing:

$$\min_{C,d_i} \Sigma_{i=1}^{N} ||d_i P_i C - L_i + \underline{o}||^2 \quad (13)$$

At first glance (13) would seem to imply that we know something about the likely magnitude of our illuminant estimates. However, without loss of generality we could normaize each L_i so it is intensity independent (e.g. divide by the sum of the vector components). Further, d_i could be chosen to make the corrected moments also sum to 1. In this case $\underline{o}$ would be a final offset correction in chromaticity space and would be perfectly valid and sensible even given the caveat of intensity indeterminacy.

To minimize (12) we propose a simple alternating least-squares solution strategy. Below, D is the $N \times N$ diagonal matrix where the ith diagonal component is d_i.

1. initialise: $D^0 = \mathcal{I}$ (the identity matrix) and $i = 1$
2. $C^i = [D^{i-1}P]^+ L$ (where $^+$ denotes the 'pre multiplying' pseudo inverse $A^+ = [A^t A]^{-1} A^t$)
3. $D^i_{jj} = L_j [P_j C^i]^\dagger$ (here we find the best scalar using the 'post-multiplying' pseudo inverse $[\underline{v}^t]^\dagger = \frac{\underline{v}}{\underline{v}^t \underline{v}}$)
4. i = i+1, goto 2 until convergence

While this solution strategy is simple it does not guarantee an optimal global solution. However, it empirically appears to work well. Further, allthough the pseudo inverse in step 2, for 3rd order moments, requires the inversion of a 19x19 matrix, we found all our data to be well enough conditioned that regularisation was not needed.

Alternating least-squares is only one way to solve our optimisation. In fact we can solve for all unknowns simultaneously if (12) is recast as a homogeneous regression (though, with much poorer estimation results). We also note that we might find C through search and perhaps also minimize an alternate error measure (e.g. the mean angular error). Further robust solution strategies could be applied.

3.4. Corrected moments vs committee-based color constancy

If we compiled row vector $\underline{p}^t$ using the outputs from 2 or more illuminant estimation algorithms then equation (11) already appears in the literature[10]. However, that work differs our own in 4 important respects. First, in [10], the correction matrix C is solved by direct least-squares i.e. it is assumed that the magnitude of the illuminant vectors estimated by the various algorithms are in sync and are directly relateable to the correct answer (this is not the case). Interestingly, the authors acknowledge this point and set forth a method of mapping chromaticity estimates to the chromaticity of the correct answer. In doing so they lose one degree of freedom per algorithm and for the 3-algorithm case reduce 9 numbers to 6. This loss of information speaks to the second important difference. Because all our moments scale linearly with intensity, when we are given a 9-vector of moments, only 1 degree of freedom is lost due to light intensity indeterminacy. Third, our moment approach has access to 'cross terms' and these are neither used nor enviaged in[10]. Yet, employing them is important to achieve the best estimation results. Last, and dependent on these first three observations, the antecedent work only delivers a modest improvement for illuminant estimation. In contrast our corrected moment approach delivers a step change in performance.

4. Experiments

In Figure 3 we show images drawn form 4 image data sets often used to evaluate lluminant estimation algorithms. Our experiments focus mainly of the so-called 'Color Checker' data set[26]. Not only is this image set the one most commonly used it is large (568 images) and so it is plausible we have enough data to solve for our correction matrices. The Color checker dataset also has the advantage that, as reprocessed by Shi and Funt[44], it is linear. The more processed[40] (i.e. the more pleasing an image looks) the more unknown variables affect an image and so the harder illuminant estimation becomes.

Figure 3. Top left an image from the Color Checker data set, Top right from the gray-ball data set, bottom left an example from a set of HDR images and bottom right from the SFU object set of images.

To corroborate our results we also evaluate our method on the hdr datasets[25] , the grayball[15] image set and the SFU object data set[3]. For each data set, we try and give a summary of the performance data that is available and always include the best prior-art result known to us.

Algorithm	Mean	Median	95% quantile
GrayWorld	6.4	6.3	11.3
Shades of Gray (p=4)	4.9	4	11.9
GrayEdge (n=2,σ=1,p=1)	5.1	4.4	11
Gamut Mapping	4.2	2.3	14.1
Spatio-Spectral Statistics	3.4	2.6	9.52
Natural Image Statistics	4.2	3.1	11.7
Exemplar-Based	3.1	2.3	-
3 Color Moments	4.0	3.3	8.9
3 Edge Moments	3.0	2.2	7.2
9 Color Moments	3.6	2.8	9.1
9 Edge Moments	2.9	2.1	7.1
19 Color Moments	3.5	2.6	8.6
19 Edge Moments	2.8	2.0	6.9

Table 1. Performance statistics for the Color Checker dataset

4.1. Colour checker data set

In this dataset every one of the 568 images contains a Macbeth Color checker[36], from which the ground truth illuminant color is measured (the checker is removed when estimating the illuminant). In Table 1 we recapitulate a range of experimental results reported by Gijensij[32, 28] and for which the illuminant estimates are also available to the community. Results are presented for Gray-World, Shades of Gray[22], Gray-Edge[46], Spectral-spatio correlations[12], and Natural Image statistics[30]. To these we add the results of Exemplar based color constancy[35] since it is, to our knowledge, the leading current algorithm. We use the angular error in degrees between the actual RGB of the light as our error measure. For this dataset we show the median[33], mean and 95% quantile errors summary statistics.

For all our data sets we use 3-fold cross validation and in all cases we minimize (12). For all datasets we have the option- in analogy to the Minkowski norm approach[23] - of raising our image data to a power. We find that a linear power term works well for the Color Checker dataset (i.e. not gamma encoded).

We remark that for this dataset and the others evaluated later, color edge moments seem more useful in estimating the illuminant than moments calculated for colors alone. We also note that while there seems to be significant benefit in accuracy in moving from 1st to 2nd order moments there is little benefit to using 3rd order moments. Notice also that the 95% quantile statistics are substantially improved compared to the prior art.

Remarkably a 3x3 matrix correction of the simple grey-edge algorithm provides illuminant estimation that out performs all other algorithms we are aware of.

The reader will be curious to learn the significance of modelling the unknown intensity of the light in the optimisation set forth in (12) since, in section 3.4, we claimed it was a matter of some importance. Well, to check, we re-ran our experiments and fixed the scaling terms as 1 throughout (the method becomes closed form and no iteration is needed). For the 9 color edge moments, the mean, median and 95% quantile error were found to respectively equal 3.6, 2.8 and 9.1 degrees of angular error (compared with 2.9, 2.1 and 7.1 before). Ignoring the variability of intensity leads to respectively 24%, 33% and 28% higher errors. A significant performance impediment was found for whatever order of moments were employed.

Algorithm	Mean	Median	Max
GrayWorld	7.9	7.3	30
MaxRGB (post-blur)	6.3	3.9	28
Constrained Minkowski	5.8	3.6	–
3 Edge Moments	4.0	3.2	17.3
9 Edge Moments	3.5	2.7	12

Table 2. Performance statistics for the HDR dataset

4.2. HDR data set

This data set has been compiled by Funt and Shi[25] and contains 105 scenes in HDR format. Like the color checker data set a reference chart is placed in every scene so the light color can be measured (again, this is removed for testing the illuminant estimation algorithms). The HDR data set was designed specifically so that image data was not clipped as it is known that clipping introduces uncertainty into illuminant estimation. Indeed, the MaxRGB algorithm (where the estimate is the max response in each color channel) is, in particular, known to perform less well in the face of extensive clipping. Max RGB works surprisingly well on the clip free dataset[25] and Max RGB calculated post blurring the image delivered very good performance. The idea of linear filtering to help illuminant estimation is also proposed in [14]. A slight increment in performance from post-blur MaxRGB was recently reported for this dataset[23]. .

In Table 2 we show the results (mean, median and max) for Gray-World, post-blur MaxRGB, the constrained Minkowski method and our algorithm (running directly on the linear data and evaluated using a 3-fold cross-validation procedure). Notice that our corrected 3 Color Edge moments method works much better than the antecedent algorithms and our corrected 9 moment approach works better still. Significantly the maximum error is very much improved – it is reduced by almost 2/3).

4.3. Grayball

This data set is so-called as every image has a gray ball, bottom right, from which the ground truth RGB of the light is measured (and then removed when illuminant estimation

Algorithm	Median
GrayWorld	7.3
GrayEdge	4.1
Constrained Minkowski	3.81
3 Color-Edge Moments	3.8
9 Color- Edge Moments	3.3

Table 3. Performance statistics for the grayball dataset

Algorithm	Mean	Median	Max
GreyWorld	9.8	7.0	37.3
GreyEdge (n=1,σ=5,p=7)	5.6	3.2	31.6
Gamut Mapping	3.6	2.1	27.1
3 Edge-Moments	4.1	3.6	14.1
9 Edge-Moments	2.6	2.0	12.9

Table 4. Performance statistics for the SFU dataset

algorithms are tested). Here we use the 10-image per clip (150 image) data set compiled by Van der Weijer et al[46]. As before, we use a 3-fold cross validation of the image set. Here, we raise all images to the power of 2 (to approximately un-do the display gamma) before we estimate our moment correction matrix.

In Table 3 we report the median results from Gray-world, Gray-edge (with optimal parameters) – only median statistics are in the journal reference[46] – and the constrained Minkowski algorithm[23]. We also show the results for our 3- and 9-color edge moment corrected approach.

4.4. SFU object dataset

The SFU image set contains linear images of 31 objects viewed under, up to, 11 lights (321 images in total). In rows 1 and 2 of Table 4 we show the results for Gray-World and Gray-Edge (with optimal parameters selected). Gamut mapping is the leading algorithm for this dataset. Its performance statistics are reported in the third row of the table.

For the 'Color checker', 'HDR' and 'Gray Ball' datasets we found good performance was delivered when moments were calculated from the linear (checker and HDR) or linearized (Grey Ball) images. Here we found it advantageous to first raise each image to the power of 2 to give brighter colors slightly greater weight. The performance of our 3- and 9-color edge moment approach (where again using a 3-fold cross-validation procedure is used) is shown in the last two rows of the table. The 9-color edge moments delivers the best performance observed to date on this data set.

5. Conclusion

This paper unveils a very surprising result: linearly correcting simple Gray-World or Gray-Edge moments leads to illuminant estimates which are more accurate than those delivered by almost all other illuminant estimation algorithms including those which use much more sophisticated and complex reasoning. Further, the extended method - that corrects a larger set (say 9) statistical moments - delivers the very best estimation performance we are aware of. The novelty of our approach lies not only in the idea of correcting moments but also in the moments we use (and that they scale with intensity) and how we solve for the correction matrix. It is universally accepted that we cannot recover the absolute intensity of the light and so, we argue, this intensity indeterminacy must be incorporated in any optimisation scheme.

That a simple corrected moment approach works so well presents an opportunity for the wider field. Indubitably, some of the innovations developed over the last few decades to improve upon Gray-World might also further extend the corrected-moment approach to illuminant estimation.

References

[1] K. Barnard, V. C. Cardei, and B. V. Funt. A comparison of computational color constancy algorithms. i: Methodology and experiments with synthesized data. *IEEE Transactions on Image Processing*, 11(9):972–984, 2002.

[2] K. Barnard, L. Martin, A. Coath, and B. V. Funt. A comparison of computational color constancy algorithms. ii. experiments with image data. *IEEE Transactions on Image Processing*, 11(9):985–996, 2002.

[3] K. Barnard, L. Martin, B. Funt, and A. Coath. A data set for color research. *Color Research and Application*, 27(3):147–151, 2002.

[4] M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision, 1987. MIT AI Lab Memo 924.

[5] S. Bianco, G. Ciocca, and C. Cusano. Color constancy algorithm selection using CART. In *Computational Colour Imaging Workshop*, pages 31–40, 2009.

[6] S. Bianco and R. Schettini. Color constancy using faces. In *IEEE Conference on computer vision and Pattern recognition*, pages 65–72, 2012.

[7] C. Borges. Trichromatic approximation for computer graphic illumination models. *Computer Graphics*, 25:101–104, 1991.

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

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

[10] V. C. Cardei and B. V. Funt. Committee-based color constancy. In *7th IS&T/SID Color Imaging Conference*, pages 311–313, 1999.

[11] A. Chakrabarti, K. Hirakawa, and T. Zickler. Color constancy beyond bags of pixels. In *IEEE Conference on Computer Vision and Pattern Recognitions*, 2008.

[12] A. Chakrabarti, K. Hirakawa, and T. Zickler. Color constancy with spatio-spectral statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1508–1517, 2012.

[13] H. Y. Chong, S. J. Gortler, and T. Zickler. The von kries hypothesis and a basis for color constancy. In *International Conference on Computer Vision*, pages 1–8, 2007.

[14] A. Choudhury and G. G. Medioni. Color constancy using denoising methods and cepstral analysis. In *IEEE International Conference on Image Processing*, pages 1637–1640, 2009.

[15] F. Ciurea and B. V. Funt. A Large Image Database for Color Constancy Research. In *11th IS&T/SIDColor Imaging Conference*, pages 160–164, 2003.

[16] M. Drew and B.V. Funt. Natural metamers. *CVGIP:Image Understanding*, 56:139–151, 1992.

[17] G. Finlayson. Color in perspective. *IEEE transactions on Pattern analysis and Machine Intelligence*, pages 1034–1038, October 1996.

[18] G. Finlayson and S. Hordley. Selection for gamut mapping colour constancy. *Image and Vision Computing*, 17(8):597–604, June 1999.

[19] G. Finlayson, S. Hordley, and P. Hubel. Color by correlation: A simple, unifying framework for color constancy. *IEEE Transactions on pattern analysis and machine intelligence*, 23(11):1209–1221, November 2001.

[20] G. Finlayson and P. Morovic. Metamer constrained colour correction. *Journal of Image Science & Technology*, pages 295–300, 2000.

[21] G. D. Finlayson and G. Schaefer. Solving for colour constancy using a constrained dichromatic reflection model. *International Journal of Computer Vision*, 42(3):127–144, 2001.

[22] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *12th IS&T/SIDColor Imaging Conference*, pages 37–41, 2004.

[23] G. D. Finlayson, P. A. Troncoso-Rey, and E. Trezzi. General $^{\mathrm{p}}$ constrained approach for colour constancy. In *ICCV Workshopon Color and Photoemtry in Computer Vision*, pages 790–797, 2011.

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

[25] B. Funt and L. Shi. The rehabilitation of maxrgb. In *18th Color Imaging Confernce*, pages 256–259. 2010.

[26] P. V. Gehler, C. Rother, A. Blake, T. P. Minka, and T. Sharp. Bayesian color constancy revisited. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

[27] R. Gershon, A. Jepson, , and J. Tsotsos. Ambient illumination and the determination of material changes. *J. Opt. Soc. Am. A*, 3:1700–1707, 1986.

[28] A. Gijsenij. Color Constancy: Research Website on Illuminant Estimation, access from http://www.colorconstancy.com.

[29] A. Gijsenij and T. Gevers. Color constancy using natural image statistics and scene semantics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):687–698, 2011.

[30] A. Gijsenij and T. Gevers. Color constancy using natural image statistics and scene semantics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):687–698, 2011.

[31] A. Gijsenij, T. Gevers, and J. van de Weijer. Generalized gamut mapping using image derivative structures for color constancy. *International Journal of Computer Vision*, 86(2-3):127–139, 2010.

[32] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, 2011.

[33] S. D. Hordley and G. D. Finlayson. Re-evaluating colour constancy algorithms. In *International Conference on Pattern Recognition*, pages 76–79, 2004.

[34] M.-K. Hu. Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, 8:179–187, 1962.

[35] H. R. V. Joze and M. Drew. Exemplar-based colour constancy. In *Proceedings of the British Machine Vision Conference*, pages 26.1–26.12. BMVA Press, 2012.

[36] C. McCamy, H. Marcus, and J. Davidson. A color-rendition chart. *J. App. Photog. Eng.*, pages 95–99, 1976.

[37] J. Montojo. Face-based chromatic adaptation for tagged photo collections, university of toronto, 2009.

[38] M. Oren and S. K. Nayar. Generalization of the lambertian model and implications for machine vision. *International Journal of Computer Vision*, 14(3):227–251, 1995.

[39] M. S. Peercy. Linear color representations for full spectral rendering. *Computer Graphics*, 27:191–198, August 1993.

[40] R. Ramanath, W. Snyder, Y. Yoo, and M. Drew. Color image processing pipeline. *Signal Processing Magazine, IEEE*, 22(1):34–43, 2005.

[41] C. R. Rosenberg, M. Hebert, and S. Thrun. Color constancy using kl-divergence. In *International Conference on Computer Vision*, pages 239–246, 2001.

[42] G. Sapiro. Bilinear voting. In *IEEE International Conference on Computer Vision*, pages 178–183, 1998.

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

[44] L. Shi and B. Funt. Re-processed version of the Gehler color constancy database of 568 images, accessed from http://www.cs.sfu.ca/colour/data/.

[45] R. T. Tan, K. Nishino, and K. Ikeuchi. Illumination chromaticity estimation using inverse-intensity chromaticity space. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 673–682, 2003.

[46] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *IEEE Transactions on Image Processing*, 16(9):2207–2214, 2007.

[47] M. Vrhel and H. Trussel. Color correction using principal components. *Color Research and Application*, 17:328–338, 1992.

[48] B. Wandell. The synthesis and analysis of color images. *IEEE Trans. Patt. Anal. and Mach. Intell.*, PAMI-9:2–13, 1987.

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