Reproduction Angular Error: An Improved Performance Metric for Illuminant Estimation

Graham D. Finlayson
g.finlayson@uea.ac.uk

Roshanak Zakizadeh
r.zakizadeh@uea.ac.uk

School of Computing Sciences
The University of East Anglia
Norwich, UK

Abstract

Only if we can estimate the colour of the prevailing light - and discount it from the image - can image colour be used as a stable cue for indexing, recognition and tracking (amongst other tasks). Almost all illumination estimation research uses the angle between the RGB of the actual measured illuminant colour and that estimated one as the recovery error. However here we identify a problem with this metric. We observe that the same scene, viewed under two different coloured lights for the same algorithm, leads to different recovery errors despite the fact that when we remove the colour bias due to illuminant (we divide out by light) exactly the same reproduction is produced.

We begin this paper by quantifying the scale of this problem. For a given scene and algorithm, we solve for the range of recovery angular errors that can be observed given all colours of light. We also show that the lowest errors are for red, green and blue lights and the largest for cyans, magentas and yellows. Next, we propose a new *reproduction angular error* which is defined as the angle between the image RGB of a white surface when the actual and estimated illuminations are 'divided out'. Reassuringly, this reproduction error metric, by construction, gives the same error for the same algorithm-scene pair. For many algorithms and many benchmark datasets we recompute the illuminant estimation performance of a range of algorithms for the new reproduction error and then compare against the algorithm rankings for the old recovery error. We find that the overall rankings of algorithms remains, broadly, unchanged - though there can be local switches in rank - and the algorithm parameters provide that the best illuminant estimation performance depend on the error metric used.

1 Introduction

Colour constancy has two parts. First, the illuminant colour is estimated. Second the colour bias due to the illumination is removed (eg. by dividing the image RGBs by the RGB estimate of the colour of the light [2]). Illuminant estimation is a field of great study and there are literally scores of algorithms proposed each year. Only if we can estimate the colour of the prevailing light - and discount it from the image - can image colour be used as a stable cue for indexing, recognition and tracking, etc. [1, 16, 24, 26]. This said it is a matter of some interest to quantify which algorithm works best and where each algorithm is in the league table of performance.

To measure performance we need to have an agreed set of benchmark test images. Here the 'correct' answer is defined to be the RGB of a white tile placed in the scene. Then, we need to measure the error of the estimated light RGB against the known ground truth. Because we cannot determine the absolute brightness of the light, *recovery* error is usually defined in an intensity independent manner. The recovery angular error (or simply the angular error) is the most widely used metric for evaluation of colour constancy algorithm performance [17, 21] and is defined as:

$$err_{recovery} = \cos^{-1}\left(\frac{(\underline{\rho}^E \cdot \underline{\rho}^{Est})}{\|\underline{\rho}^E\|\|\underline{\rho}^{Est}\|}\right) \tag{1}$$

where ρ^E denotes the RGB of the actual measured light, ρ^{Est} denotes the RGB estimated by an illuminant estimation algorithm and '.' denotes the vector dot product. Over a benchmark set, the average angular performance is calculated (including mean, median, and quantiles) and different algorithms are ranked according to these summary statistics [19].

In this paper, we argue that recovery angular error despite its wide spread adoption has a fundamental weakness which casts doubt on its suitability. To illustrate this point we show at the top of Figure 1 four images of the same scene from the SFU Lab dataset [3]. Notice how much the colour (due to illumination) varies from left to right. Now, using the simple gray-world algorithm [4] for illuminant estimation we estimate the RGB of the light (the average image colour is the estimated colour of the light). Dividing the images by this estimate we produce the image outputs shown in the second row. In this case gray-world works reasonably well and the object colours look correct (though, of course this is not always the case). It is easy to show - as shown here - that dividing out by the gray-world estimate (or, indeed the estimates made by most algorithms) that the **same** output reproduction is made. In the 3rd row of the figure we show the recovery angular errors (the plot with open bullets). Even though the same reproduction is produced the recovery angular error varies from 5.5° to 9° (an 80% difference).

That recovery angular error is disconnected from how the estimates of the illuminant are used, is a serious problem. To illustrate how serious, consider an illuminant that is very red i.e. has an RGB: $(1, \varepsilon, \varepsilon)$. Under such a red light all image colours will be similarly red: $(k_r, \varepsilon k_g, \varepsilon k_b)$ where k_r, k_g and k_b are the fraction of light that is reflected from a surface in each of the colour channels. Clearly, the gray-world estimate in this case will be a vector that is approximately in the same direction as the actual light i.e. the recovery error will be small (and in the limit zero). Yet, these highly chromatic lights are precisely those for which colour constancy algorithms fail.

In this paper we, in effect, propose that the illuminant estimation error should be designed with the knowledge of how the estimate is to be used. Our work builds on that of Forsyth [15], who measured how clustered the RGBs for the same surface were post colour constancy processing: the better the colour constancy the more clustered the RGBs. Like Forsyth we wished to tie illuminant estimation to colour constancy (i.e. post dividing out) but retain the simplicity of the angular error idea. Here we propose the reproduction angular error which is defined as the angle between the image RGB of a white surface when the actual and estimated illuminations are 'divided out'. Crucially we show that for almost all algorithms the reproduction angular error is almost constant as the illumination changes (see the filled circles in Figure 1).

We undertake a large study of existing illuminant estimation algorithms running on common benchmark datasets for both the legacy recovery angular error and our new reproduction
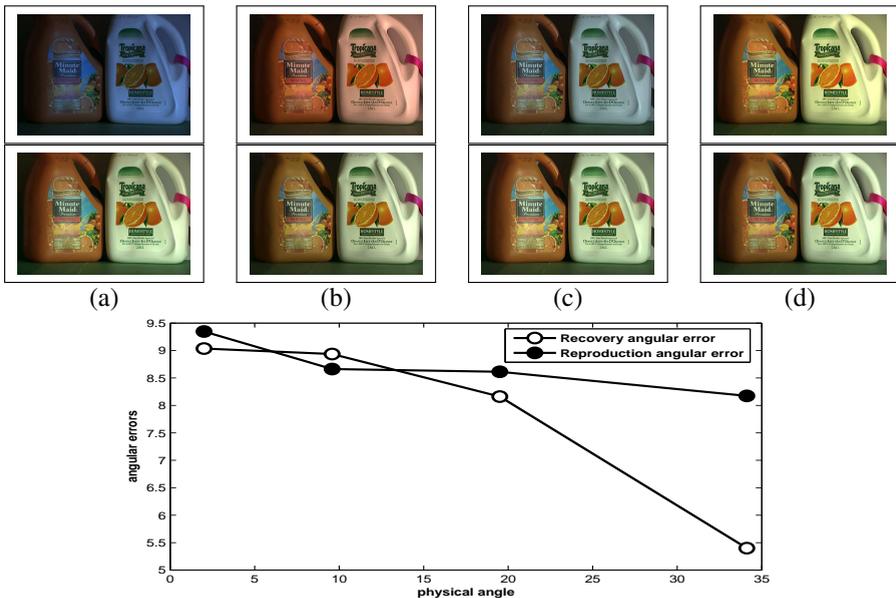
Figure 1: Row 1: four images captured under very chromatic illuminant : (a) solux-4700K+blue filter; (b) Sylvania warm white fluorescent; (c) solux-4700K+3202+blue filter; (d) Philips Ultralume fluorescent. Row 2: corrected images using general gray-world [4] algorithm (Images are from [6]). Row 3: The Recovery angular error (conventional error measure) versus the Reproduction angular error (proposed error measure).

angular error. We are particularly interested in whether the rankings of algorithms changes. The good news is - perhaps because most typical lights are not so chromatic - that the gross ranking of algorithms remains broadly similar whether we use recovery or reproduction angular error. However, there can be local 'swaps' in the rankings of algorithms and there are significant changes in the specific parameters (which tune all algorithms) which work best.

In §2 we present the so-called RGB model of image formation which is useful for analysing illuminant estimation algorithms. The reproduction angular error is presented in §3, where we also bound the range of recovery angular errors that can be found by an algorithm. There we also prove the stability of the reproduction error metric as the light changes. In §4, a large number of illuminant estimation algorithms are benchmarked using reproduction angular error and the algorithms' rankings compared with those found using recovery error. The paper concludes in §5.

2 Background: Image Formation

The spectral power distribution illuminating a scene is denoted as $E(\lambda)$. The light strikes an object with surface spectral reflectance $S(\lambda)$ and the light reflected is proportional to the multiplication of the two functions. The light is then sampled by a sensor with a spectral sensitivity $R(\lambda)$. The integrated response of a sensor to light and surface is calculated as [28]:

$$\rho_k^{E,S} = \int_\omega R_k(\lambda)E(\lambda)S(\lambda)d\lambda \quad k \in \{R,G,B\} \tag{2}$$

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 image formation.

Simple as eq. (2) is, it is in fact quite complex. It would, for example, be impossible for an illuminant estimation algorithm to recover the full spectrum $E(\lambda)$. Thankfully, the complexity of eq. (2) can, for most practical purposes be simplified and the RGB model of image formation used [6] is:

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

Remarkably, eq. (3) with certain caveats, generally holds to a good approximation [13]. An important interpretation of ρ_k^S is that it is the colour of the surface viewed under a white uniform light $E(\lambda) = 1$.

Colour constancy can be thought of as 'mapping the RGBs measured in an image back to a reference white lighting condition' [15]. 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. We rewrite eq. (3) in vector form:

$$\underline{\rho}^{E,S} = \underline{\rho}^E * \underline{\rho}^S \tag{4}$$

Note, the meaning of eq. (4) is that we multiply the RGB vectors $\underline{\rho}^E$ and $\underline{\rho}^S$ component-wise. If $\underline{\rho}^{Est}$ denotes the estimate of the illuminant made by some algorithm then we recover the colour of the surface (remove colour bias due to the colour of the prevailing light) by calculating:

$$\frac{\underline{\rho}^{E,S}}{\underline{\rho}^{Est}} \approx \underline{\rho}^S \tag{5}$$

where again the division of vectors here is component-wise. In eq. (5), how good the approximation is, depends on how well the illuminant is estimated.

In the context of this paper, two important results follow from the RGB model of image formation. First, we can simulate some other, unknown, light E' by simply multiplying by an arbitrary 3-vector:

$$\underline{\rho}^{E',S} = \underline{d} * \underline{\rho}^{E,S} \qquad \underline{d} = [\alpha \ \beta \ \gamma]^t \quad \alpha, \beta, \gamma \geq 0 \tag{6}$$

For all E we can apply all all-positive vectors $\underline{d}$. Secondly, assuming that the illuminant estimate can be viewed as a kind of statistical *moment* of the image RGBs it follows that for an N-pixel image:

$$\underline{\rho}^{Est} = moment(\{\underline{\rho}^{E,S_1}, \underline{\rho}^{E,S_2}, ..., \underline{\rho}^{E,S_N}\}) \tag{7}$$

Then from eq. (6) and eq. (7),

$$\underline{d} * \underline{\rho}^{Est} = moment(\{\underline{\rho}^{E',S_1}, \underline{\rho}^{E',S_2}, ..., \underline{\rho}^{E',S_N}\}) \tag{8}$$

Eq. (8) teaches that if two lights are related by 3 scaling factors $\underline{d}$ then the corresponding illuminant estimates are similarly related. Eq. (8) holds for most illuminant estimation algorithms including gray-world [4], MaxRGB [22], shades-of-gray [12], 1st and 2nd gray-edge algorithms [22], as well as the gamut-based methods [10, 14, 15, 18].

Many of the statistical moments (eq. (7)) used to estimate the illuminant of the scene can be summarised in a single equation [27]:

$$\left(\int |\frac{\delta^n \rho(x)}{\delta x^n}|^p \mathrm{d}x \right)^{1/p} = k \rho_{n,p,\sigma}^{Est} \qquad (9)$$

Here $\rho(x)$ is the camera response at location x of an RGB image. The image can be smoothed with a Gaussian averaging filter with standard deviation σ pixels and after being smoothed is differentiated with an order n differential operator. We then take the absolute Minkowski p-norm average [12] over the whole image. The unknown value k represents the fact that the true magnitude of the prevailing illuminants cannot be recovered. The tunable parameters of σ and p-norm can be chosen so that the algorithms perform their best. For all statistical moments calculated using eq. (9), eq. (8) is true.

3 Reproduction versus Recovery Angular Error

3.1 The range of recovery angular error

Figure 1 shows that even though there is the same image reproduction, the recovery angular error for different light found using the same algorithms varies. How large is this variation?

Theorem 1. *Given a white reference light (the RGB of the light is $\underline{U} = [1\ 1\ 1]^t$) and denoting the illumination estimate made by a 'moment type' illuminant estimation algorithm as $\underline{\mu}$ then the illuminant that maximises recovery angular error is an illuminant with 0 in exactly one of the either R, G or B channels.*

Proof. From eq. (6) and without the loss of generality we assume that the starting illuminant is $\underline{U}$ (if it is not, we can map to $\underline{U}$ using 3 scaling factors). From eq. (8) the new illuminant estimate from eq. (7) and eq. (8) is $\underline{d} * \underline{\mu}$ and the recovery error (eq. (1)) can be written as:

$$err_{recovery}(\underline{d}, \underline{d} * \underline{\mu}) = \cos^{-1}\left(\frac{(\alpha^2 \mu_r + \beta^2 \mu_g + \gamma^2 \mu_b)}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}\sqrt{(\alpha\mu_r)^2 + (\beta\mu_g)^2 + (\gamma\mu_b)^2}} \right) \qquad (10)$$

Remembering $\underline{d} = [\alpha\ \beta\ \gamma]$ without loss of generality - as, in illuminant estimation, we are only interested in the orientation of $\underline{d}$ - let us set $\alpha = 1$. Assume we are given β and γ and we ask the question: "does the error vary if we hold γ fixed and we solve for the optimal β". Observing that we maximize eq. (10) by minimizing the function within the parentheses, we find β by minimizing $f(\beta)$:

$$f(\beta) = \left(\frac{(\mu_r + \beta^2 \mu_g + \gamma^2 \mu_b)}{\sqrt{1 + \beta^2 + \gamma^2}\sqrt{(\mu_r)^2 + (\beta\mu_g)^2 + (\gamma\mu_b)^2}} \right) \qquad (11)$$

The derivative of $f(\beta)$ is:

$$\frac{\partial f}{\partial \beta} = \frac{-\beta \cdot (\mu_r + \mu_g \beta^2 + \mu_b \gamma^2) \cdot (\mu_r^2 + \mu_g^2 \beta^2 + \mu_b^2 \gamma^2)}{(1 + \beta^2 + \gamma^2)^{\frac{3}{2}} \cdot (\mu_r^2 + \mu_g^2 \beta^2 + \mu_b^2 \gamma^2)^{\frac{3}{2}}}$$
$$+ \frac{2\mu_g \beta \cdot (1 + \beta^2 + \gamma^2) \cdot (\mu_r^2 + \mu_g^2 \beta^2 + \mu_b^2 \gamma^2)}{(1 + \beta^2 + \gamma^2)^{\frac{3}{2}} \cdot (\mu_r^2 + \mu_g^2 \beta^2 + \mu_b^2 \gamma^2)^{\frac{3}{2}}} - \frac{\mu_g^2 \beta \cdot (1 + \beta^2 + \gamma^2) \cdot (\mu_r + \mu_g \beta^2 + \mu_b \gamma^2)}{(1 + \beta^2 + \gamma^2)^{\frac{3}{2}} \cdot (\mu_r^2 + \mu_g^2 \beta^2 + \mu_b^2 \gamma^2)^{\frac{3}{2}}} \qquad (12)$$

In eq. (12), (β) is the common factor in all three numerators and the solution to minimizing f is $\beta = 0$. β is 0 for all γ (including the γ that maximizes eq. (11)).   □

Lemma 1.1. *Assuming $\alpha = 1$ and $\beta = 0$, the recovery angular function has at most 3 stationary values.*

Proof. Since $\alpha = 1$ and $\beta = 0$, this leaves the function within the parentheses in eq. (10) with γ as a variable and $f(\gamma)$ is written as:

$$f(\gamma) = \frac{(\mu_r + \gamma^2 \mu_b)}{\sqrt{1+\gamma^2}\sqrt{(\mu_r)^2 + (\gamma\mu_b)^2}} \tag{13}$$

The derivative of $f(\gamma)$ is calculated as:

$$\frac{\partial f}{\partial \gamma} = \frac{(\mu_r - \mu_b)^2 \cdot \gamma \cdot (\mu_b \gamma^2 - \mu_r)}{(\gamma^2 + 1)^{\frac{3}{2}} \cdot (\mu_b^2 \gamma^2 + \mu_r^2)^{\frac{3}{2}}} \tag{14}$$

Which if it is set to zero leads to:

$$\gamma = \pm\sqrt{\mu_r/\mu_b} \quad \gamma = 0 \tag{15}$$

when $\gamma = 0$ the angle is a global minimum and equals to 0. When $\gamma = +\sqrt{\mu_r/\mu_b}$ (the lights are all positive) then the function in eq. (13) is a local minimum (which one can show by applying the standard second derivative test). This implies that $cos^{-1}(f(\gamma))$ at this point is a local maximum.   □

Of course we can repeat the above argument, setting $\alpha = 0$ or $\gamma = 0$. So, it follows that there are 3 possible maximums, one of which is the global maximum.

It follows that lights with one wavelength set to zero (eg. [1 0 $\sqrt{\mu_r/\mu_b}$]) maximise the recovery angular error for a given illuminant estimation algorithm applied on a given scene(lemma 1.1). It follows from the theorem that lights which are cyan, purple and yellow result in the highest angular error. Conversely, as we saw in the introduction pure red, green and blue lights induce the lowest error. In Figure 2, the two magenta curves represent cumulative probability distribution function of the analytical maximum errors for the two algorithms - gray-world [4] (solid lines in Figure 2 (a)) and pixel-based gamut mapping [13] (dashed lines in the same Figure) - for the 321 images of the SFU Lab dataset [3].

We are also interested in the maximum error for more typical 'real' lights (eg. for lights that are a convex combination of the measured lights for the SFU Lab dataset [3]). There is a companion theorem (which space limitation does not allow us to reproduce here) which allows us to find the maximum error of a convex set (in Figure 2, shown in blue). The actual recovery angular error for the ground-truth illuminant provided by SFU Lab dataset [3] is also included in this Figure (red lines). We can see that there is a great difference between the maximum angular errors and the actual angular errors, specially for the gray-world algorithm. Note that the maximum illumination error performance is similar for gray-world and pixel-based gamut mapping (even though the average performances are quite different). Also, even when typical lights are considered, the actual error is still significantly worse for gray-world. With the advent of LED lighting, which have narrow emission spikes, it is now more likely to encounter a light which induces a worst case error.
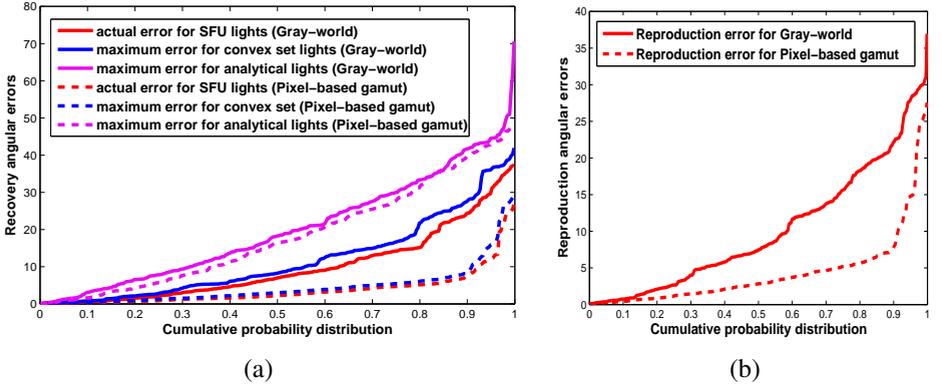
Figure 2: (a) Cumulative probability distribution of recovery angular errors for: ground-truth light of SFU Lab dataset [8] (in red) , convex-set lights (in blue), analytical lights (in magenta) of gray-world [4] (solid lines) and pixel-based gamut mapping [13] (dashed lines) for the 321 images of SFU Lab dataset; (b) Cumulative probability distribution of Reproduction angular errors of the same algorithms for the same dataset.

3.2 Reproduction Angular Error

The RGB model of image formation presented in section 2 teaches that as the illumination changes (eq. (6)), then all the RGBs in the images are scaled by the same three multiplicative factors [11]. Usually the estimated illuminant is 'divided out' to remove the colour bias due to illumination. Eq. (5) shows how we can remove the bias of illumination by dividing out the colour of the light. Let us rewrite eq. (5) for the specific example of a white surface $\underline{\rho}^{E,W} = [1\ 1\ 1]^t$. We see that:

$$\frac{\underline{\rho}^{E,W}}{\underline{\rho}^{Est}} \approx \underline{U} = \frac{\underline{\rho}^{E,W}}{\underline{\rho}^{E,W}} \qquad (16)$$

Remembering we cannot recover the absolute brightness of the light, we define the **Reproduction Angular Error** - our new metric for assessing illuminant estimation algorithms - as:

$$err_{reproduction} = \cos^{-1}\left(\frac{(\underline{\rho}^{E,W}/\underline{\rho}^{Est}).\underline{U}}{|(\underline{\rho}^{E,W}/\underline{\rho}^{Est})|\sqrt{(3)}} \right) \qquad (17)$$

Theorem 2. *Given a single scene viewed under two lights. The reproduction error of the estimated light by a 'moment type' illuminant estimation algorithm is the same.*

Proof. For a chromatic light defined with $\underline{d} = [\alpha\ \beta\ \gamma]^t$ (eq. (6)), using the fact presented in eq. (7), the reproduction angular error (eq. (17)) can be written as:

$$err_{reproduction} = \cos^{-1}\left(\frac{(\frac{\alpha}{\alpha\mu_r} + \frac{\beta}{\beta\mu_g} + \frac{\gamma}{\gamma\mu_b})}{\sqrt{(\frac{\alpha}{\alpha\mu_r})^2 + (\frac{\beta}{\beta\mu_g})^2 + (\frac{\gamma}{\gamma\mu_b})^2}\sqrt{(3)}} \right) \qquad (18)$$

It can be seen easily in eq. (18), that the scaling factors α, β and γ are cancelled. The reproduction error is stable regardless of the colour of the light. □

In Figure 1 we show the same object under 4 chromatic lights together with the corresponding images where the light colour is divided out. In the third row are the recovery and reproduction angles for the illuminant estimates. Notice the large variation for the recovery errors. Whereas the reproduction errors are much more stable. They are not exactly constant as, although the RGB model of light formation is good, it is still an approximation [13].

4 Experiments and Results

We evaluate a representative selection of illuminant estimation algorithms which includes MaxRGB [22], gray-world [4], shades-of-gray [12], 1st and 2nd gray-edge algorithms [27], as well as the gamut-based methods [10, 14, 15, 18]. These algorithms were applied on the well-known illuminant estimation datasets such as SFU Lab dataset [3], Gray-ball set [7] and Colour-checker (by Shi) [23]. In addition, other colour constancy algorithms such as spatial-correlations-based algorithms (denoted here as heavy tailed-based [5] and weighted gray-edge [20]) are also included. The algorithms are previously evaluated using the recovery angular error by Gijsenij *et al.* [19]. Here we repeat this experiment for our new reproduction angular error.

Method	Recovery error				Reproduction Error			
	p	σ	Median	Rank	p	σ	Median	Rank
Gray-world	-	-	7°	11	-	-	7.5°	11
MaxRGB	-	-	6.5°	10	-	-	7.4°	10
Shades of gray	7	-	3.7°	**9**	7	-	3.9°	**8**
1st Gray-edge	7	4	3.2°	**7**	14	4	3.58°	**6**
2nd Gray-edge	14	10	2.7°	4	15	10	3°	4
Pixel-based gamut	-	4	2.26°	**2**	-	4	2.8°	**3**
Edge-based gamut	-	2	2.27°	**3**	-	2	2.7°	**2**
Intersection-based gamut	-	4	2.1°	1	-	3	2.5°	1
Union-based gamut	-	2	3°	5	-	2	3.4°	5
Heavy tailed-based	-	-	3.5°	**8**	-	-	4.1°	**9**
Weighted gray-edge	2	1	3.1°	**6**	2	1	3.62°	**7**

Table 1: Recovery and Reproduction median errors of several colour constancy algorithms for SFU Lab dataset [7]. The ranks given to each algorithm are bold and underlined if they have changed. The optimal parameters are also shown where applicable and different ones are highlighted.

In Table 1, the recovery and reproduction median angular errors for the SFU Lab dataset (321 images) [3] are shown. Where applicable the optimal parameters-Minkowski norm and smoothing value (eq. (9))- are also determined. In this table we can already notice the ranking of several algorithms (ranks shown in bold and underlined) have moved. To study to what extent the ranking of these algorithms has changed using the new reproduction error compared to the old recovery error, we performed the Kendall test [8, 25]. A change in the ranking of a selection of illuminant estimation algorithms can be considered a permutation problem. The Kendall test is a method to compare two permutations (it correlates to the number of exchanges needed in a bubble sort to convert one permutation to the other [9]).
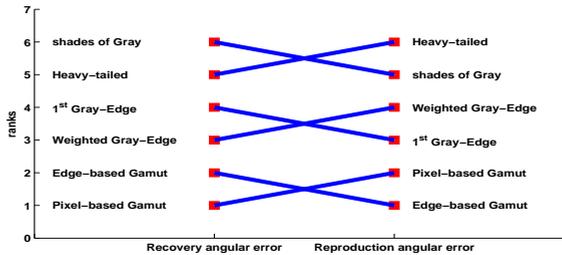
Figure 3: The pictorial scheme of Kendall test for the changed rank algorithms in Table 1.

The Kendall test statistic T can give us a measure of correlation between pairs of ranks where there are no ties. A pair of unique observations (x_1, y_1) and (x_2, y_2) are said to be discordant if the ranks of the two elements (x_1, x_2) and (y_1, y_2) do not agree, otherwise the pair are concordant. In case of no ties, T is defined as:

$$T = C - D, \tag{19}$$

where C is the number of concordant pairs and D is the number of discordant pairs.

Figure 3 depicts a pictorial scheme of the Kendall test for the 6 algorithms with changed ranks, the total number of pairings is $6(6-1)/2 = 15$ and the number of the crossings represents the number of discordant pairs. For the 6 algorithms with changed ranks in Table 1, the Kendall test statistic is $T = 12 - 3 = 9$. For our data we conclude the rankings are discordant at 97% significance level (i.e. with type I error (P) < 0.03) [8].

Notice also that in Table 1 the optimal parameters (eq. (9)) chosen for the tunable algorithms such as 1st and 2nd gray-edge algorithms [27] differ for recovery versus reproduction angular error.

We compute similar ranking results for different summary error measures - e.g. the mean and 95% quantiles - and come to the same conclusion: broadly the rankings change a little using the new reproduction error and optimal algorithm parameters also change. We come to the same conclusion using the Colour-checker [23] and the Gray-ball [7] data sets[1].

5 Conclusion

In this paper we propose a reproduction angular metric to assess the performance of illuminant estimation algorithms and argue that it improves upon the widely used recovery angular error. The latter measure was shown - for the same scene and algorithm pair - to result in a huge range of angular errors (which we solved for this error interval) even though the images reproduced (after dividing out the light) are the same. In contrast, the proposed reproduction angular error (which measures how well a white surface is reproduced) is by construction much more stable. We show that using the new measure, the rankings of algorithms reported in the literature while broadly similar, can change. The ranks of local pairs of algorithms (e.g. pixel-based and edge-based gamut mapping [18]) can switch and this switching results in rankings which are statistically different. Also, using the new reproduction error, the op-

---

[1]http://colour.cmp.uea.ac.uk/datasets/illuminant-estimation-errors.html

timal parameters used to tune a given illumination estimation algorithm can change.

References

[1] Alaa E Abdel-Hakim and Aly A Farag. Csift: A sift descriptor with color invariant characteristics. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:1978–1983, 2006.

[2] Kobus Barnard, Vlad Cardei, and Brian 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.

[3] Kobus Barnard, Lindsay Martin, Brian Funt, and Adam Coath. A data set for color research. *Color Research & Application*, 27(3):147–151, 2002.

[4] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980.

[5] Ayan Chakrabarti, Keigo Hirakawa, and Todd Zickler. Color constancy with spatio-spectral statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(8):1509–1519, 2012.

[6] Hamilton Y Chong, Steven J Gortler, and Todd Zickler. The von Kries hypothesis and a basis for color constancy. *IEEE 11th International Conference on Computer Vision (ICCV)*, pages 1–8, 2007.

[7] Florian Ciurea and Brian Funt. A large image database for color constancy research. *11th IS & T/SID Color and Imaging Conference*, 2003(1):160–164, 2003.

[8] William Jay Conover and WJ Conover. *Practical nonparametric statistics, Third Edition*. John Wiley & Sons, New York, 1999. ISBN 0-471-16068-7.

[9] Ronald Fagin, Ravi Kumar, and D Sivakumar. Comparing top k lists. *SIAM Journal on Discrete Mathematics*, 17(1):134–160, 2003.

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

[11] Graham D Finlayson. Corrected-moment illuminant estimation. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1904–1911. IEEE, 2013.

[12] Graham D Finlayson and Elisabetta Trezzi. Shades of gray and colour constancy. *12th IS & T/SID Color and Imaging Conference*, 2004(1):37–41, 2004.

[13] Graham D Finlayson, Mark S Drew, and Brian V Funt. Spectral sharpening: sensor transformations for improved color constancy. *Journal of the Optical Society of America A*, 11(5):1553–1563, 1994.

[14] Graham D Finlayson, Steven D Hordley, and Ruixia Xu. Convex programming colour constancy with a diagonal-offset model. *Proc. IEEE Int. Conf. Image Processing*, 3: 948–951, 2005.

[15] David A Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990.

[16] Theo Gevers and Arnold WM Smeulders. Color-based object recognition. *Pattern recognition*, 32(3):453–464, 1999.

[17] Arjan Gijsenij, Theo Gevers, and Marcel P Lucassen. Perceptual analysis of distance measures for color constancy algorithms. *Journal of the Optical Society of America A*, 26(10):2243–2256, 2009.

[18] Arjan Gijsenij, Theo Gevers, and Joost Van De Weijer. Generalized gamut mapping using image derivative structures for color constancy. *International Journal of Computer Vision*, 86(2-3):127–139, 2010.

[19] Arjan Gijsenij, Theo Gevers, and Joost Van De Weijer. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, 2011. URL http://colorconstancy.com/.

[20] Arjan Gijsenij, Theo Gevers, and Joost Van De Weijer. Improving color constancy by photometric edge weighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5):918–929, 2012.

[21] Steven D Hordley and Graham D Finlayson. Reevaluation of color constancy algorithm performance. *Journal of the Optical Society of America A*, 23(5):1008–1020, 2006.

[22] Edwin Herbert Land et al. The retinex theory of color vision. *Scientific American*, 237 (6):108–128, 1977.

[23] Lilong Shi and Brian Funt. Re-processed version of the gehler color constancy dataset of 568 images. *Simon Fraser University*, 2010. URL http://www.cs.sfu.ca/~colour/data/.

[24] David Slater and Glenn Healey. The illumination-invariant recognition of 3d objects using local color invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(2):206–210, 1996.

[25] P. Sprent and N.C. Smeeton. *Applied Nonparametric Statistical Methods, Fourth Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 2007. ISBN 9781584887010. URL http://books.google.co.uk/books?id=arn4eBAWodwC.

[26] Koen EA Van De Sande, Theo Gevers, and Cees GM Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 2010.

[27] Joost Van De Weijer, Theo Gevers, and Arjan Gijsenij. Edge-based color constancy. *IEEE Transactions on Image Processing*, 16(9):2207–2214, 2007.

[28] Brian A Wandell. The synthesis and analysis of color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1):2–13, 1987.