

Non-iterative Comprehensive Normalisation

Graham Finlayson and Ruixia Xu

School of Information Systems, University of East Anglia

Norwich, United Kingdom

graham@sys.uea.ac.uk

Abstract

The light reflected from an object depends not only on object colours but also on lighting geometry and illuminant colour. As a consequence the raw colour recorded by a camera is not a reliable cue for object based tasks such as recognition and tracking. One solution to this problem is to find functions of image colours that cancel out dependencies due to illumination. While many invariant functions cancel out either dependency due to geometry or dependency due to illuminant colour, only the comprehensive normalisation has been shown (theoretically and experimentally) to cancel both. However, this invariance is bought at the price of an iterative procedure.

In this paper we develop a non iterative log comprehensive normalisation procedure. We begin by reviewing the idea that lighting effects due to geometry and light colour can, under certain reasonable simplifying assumptions, both be modelled using simple scalar multipliers. We now take logarithms and turn geometry and light colour dependency into additive processes. We show how in this log color space two simple projection operators lead to invariance to geometry and light colour. Moreover, because projection operators are idempotent, illuminant invariance is achieved in a single step. Experiments demonstrated that log comprehensive normalisation used as a preprocessing step supports accurate colour based object recognition independent of lighting conditions.

1. Introduction

Colour is an important cue for object recognition and often used as a feature for indexing image databases.^{2,7,18,20} However, object recognition experiments based on the raw RGB values fails when the illumination condition change. For example, the shading observed on objects is a function of the light source position: there is a lighting geometry dependence on colour. Also, as the illumination colour changes so too do the colour of the measured RGBs.

Two approaches are reported in the literature for dealing with illumination problems. First, colour constancy algorithms^{17,14,8,11,22} attempt to recover surface reflectance (or more precisely correlates of surface reflectance) and this is, by definition, independent of lighting conditions. Unfortunately colour constancy has proven to be a very hard problem to solve. The current state of the art can only deliver approximate constancy.

Moreover, it has been shown that the constancy delivered by many of the algorithms¹³ is insufficient to render colour a stable enough cue for object recognition.

An easier way to deal with the illuminant colour problem is the colour invariant approach.^{5,10,4,19,1,15} The goal here is to find functions of proximate image pixels which cancel out lighting dependencies. As an example the chromaticity function, used extensively in colour science and computer vision, removes the intensity from an RGB response vector. Since shading in an image is an intensity artifact, it follows that the chromaticity normalization can be used to remove lighting geometry. Another normalization common in the colour literature is the grey-normalization.²⁴ Here the average RGB pixel in an image is calculated and the other pixels described relative to the average (if (R_p, G_p, B_p) denotes an image pixel and (μ_R, μ_G, μ_B) the average then $(R/\mu_R, G/\mu_G, B/\mu_B)$ is calculated). Under most conditions^{25,23} this simple normalization will remove illuminant colour. Neither the chromaticity nor the grey-world normalization suffices to remove both illuminant colour and geometry dependencies.

Both dependencies are, however, removed in the comprehensive normalisation scheme.¹⁰ Here the chromaticity and grey-world normalizations are applied successively and iteratively. It was shown¹⁰ that this procedure converges to a unique fixed point. Comprehensive normalization for the same scene viewed under most reasonable lighting conditions delivers the same intrinsic image which depends only on surface reflectance. But, is this iteration necessary? In some sense this question has been answered by research which followed the comprehensive normalization result. The methods of Gevers and Smeulders¹⁵ and Berens et al³ remove the illumination geometry and color in a single step. Yet to do this, these normalizations take the 3-component RGB image into a new coordinate frame: the idea of redness, greenness and blueness is removed and one is left with lighting independent scalars which do not have their usual colour meaning. Thus, while the comprehensive normalization is iterative, the iterative cost of the computation delivers an RGB colour image as output (and in this image the RGBs have similar 'meanings' to the input colours).

In this paper we ask whether we can normalize an RGB image, and keep the semantic meaning of the RGBs the same but without the cost of iteration? The answer is yes! We begin with the simple observation that the lighting geometry and lighting colour processes which are multiplicative in RGB space become additive in log RGB space. In a second stage we show that these additive

effects can be removed using projection operators. For example, we show that shading in log RGB space is an additive component in the direction (1,1,1). Projecting RGBs orthogonal to this direction removes shading. A second, simple projector is used to remove light colour. That is, we remove lighting geometry and light colour in log-space in a manner that is analogous to the comprehensive normalisation. However, and this is the important result, we show that the process is non-iterative. This is so, because of the nature of projector operators in linear algebra. A projector multiplied by itself returns the same projector: projectors are idempotent. It follows that if we remove lighting geometry and lighting colour once we remove it for all time.

The new log-normalisation is tested in the context of colour object recognition experiments. Here objects are imaged under different lighting conditions. The images are normalised and then we ask whether the remaining colour content suffices for recognizing the object. We found that the new log-comprehensive normalisation delivered almost perfect recognition on a moderate sized image data set. Moreover, the performance was similar to that delivered by the iterative comprehensive normalisation procedure which in turn has been shown to improve over antecedent methods.

The rest of the paper is organised as follows: the basis of colour image formation is introduced and colour image normalisation methods are reviewed in Section 2. The new log-space comprehensive normalisation is presented in Section 3. In section 4 the object recognition experiments are described and results shown. The paper finishes with conclusions in section 5.

2. Background

In order to develop the theory we will adopt some widely used simplifying assumptions. First, we assume that the response of a colour camera is linear. That is, if we view a surface under a given light and then double the intensity of the light we expect a doubling in the recorded RGB values. If we denote the i th pixel in an image (R_i, G_i, B_i) then as the power of the incident illumination changes then

$$(R_i, G_i, B_i) \rightarrow (\rho_i R_i, \rho_i G_i, \rho_i B_i) \quad (1 \text{ Lighting geometry})$$

where ρ_i is a simple scalar. Note that this scalar has a subscript i indicating that all pixels can have their own individual brightness factors. Brightness changes, or lighting geometry, is a local phenomenon.

Changing the relative position of the light source with the surface introduces shading. Assuming matte Lambertian reflectance and letting $\underline{n}$ denote surface normal and $\underline{e}$ the lighting direction then the power of the light striking a surface is proportional to the scalar $\underline{n} \cdot \underline{e}$ (dot-product). It follows then that a change in shading can also be described according to Equation (1). It is important to note that the Lambertian assumption is important here. Equation (1) cannot account for lighting geometry changes for highly specular surfaces.

Let us now consider a change in lighting colour (assuming lighting geometry is held fixed). In almost all

circumstances, Equation (2) approximately holds²⁵ (or can be made to hold⁹).

$$\begin{pmatrix} R_i \\ G_i \\ B_i \end{pmatrix} \rightarrow \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} R_i \\ G_i \\ B_i \end{pmatrix} \quad (2) \text{ Lighting colour}$$

where α , β and γ are scalars. Note these scalars do not depend on the pixel (there is no subscript i). That is colour change is a global phenomenon affecting the whole image.

Relative to (1) and (2) we can now consider how the scalars ρ_i , α , β and γ can be removed from images.

The chromaticity normalization defined in Ref. 3 cancels lighting geometry:

$$\left(\frac{\rho_i R_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i}, \frac{\rho_i G_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i}, \frac{\rho_i B_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i} \right) \quad (3)$$

(clearly the ρ_i term cancels). We will denote the chromaticity normalisation carried out on the image $\mathcal{I}$ as $C(\mathcal{I})$.

Let $\mu(R)$ denote the mean red pixel value for an image. Assuming N pixels in an image:

$$\mu(R) = \frac{\sum_{i=1}^N R_i}{N} \quad (4)$$

Under a change in illuminant colour (see equation (4)):

$$\alpha \mu(R) = \frac{\sum_{i=1}^N \alpha R_i}{N} \quad (5)$$

It is now easy to see that the grey normalization defined in Ref. 6 cancels light colour:

$$\left(\alpha \frac{R_i}{\alpha \mu(R_i)}, \beta \frac{G_i}{\beta \mu(G_i)}, \gamma \frac{B_i}{\gamma \mu(B_i)} \right) \quad (6)$$

It is clear the scalars α , β and γ cancel. We will denote the image $\mathcal{I}$ post grey normalization as $G(\mathcal{I})$.

Neither (3) nor (6) by itself suffices to remove both lighting geometry and lighting colour change. Comprehensive normalisation, which can remove both dependencies, is defined as:

1. $\mathcal{I}_0 = \mathcal{I}$ Initialization
 2. $\mathcal{I}_{i+1} = G(C(\mathcal{I}_i))$ Iteration step
 3. $\mathcal{I}_{i+1} = \mathcal{I}_i$ Termination condition
- (7)

The comprehensively normalised counterpart of the image $\mathcal{I}$ is denoted $CN(\mathcal{I})$.

Finlayson et al¹⁰ proved that this iterative process converged to a unique fixed point. Thus if $\mathcal{I}$ and $\mathcal{J}$ denote images of the same scene where only the lighting conditions have changed then $CN(\mathcal{I}) = CN(\mathcal{J})$. Moreover, this method was shown to outperform the individual normalization functions $C()$ or $G()$ in object recognition experiments.

The problem with (7) (unlike (3) or (6)) is that it is iterative. Can we carry out a comprehensive normalisation without iteration?

3. Comprehensive Normalisation in Log-space

It is clear in RGB space that as the lighting conditions change, the effect on a pixel is multiplicative. Comprehensive normalisation removes multiplicative dependencies using division. Here we propose to work in log RGB space since multiplication in log space is turned into addition. As we shall see this simple conceptual step eventually leads us to an idempotent (non-iterative) normalization. Let $r = \ln R$, $g = \ln G$ and $b = \ln B$. Then we can rewrite (1) and (2) as:

$$(r_i, g_i, b_i) \rightarrow (\rho_i + r_i, \rho_i + g_i, \rho_i + b_i) \quad (8 \text{ Lighting geometry in log space})$$

$$\begin{pmatrix} r_i \\ g_i \\ b_i \end{pmatrix} \rightarrow \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} r_i \\ g_i \\ b_i \end{pmatrix} \quad (9 \text{ Lighting colour in log space})$$

Let us now consider how we might remove lighting dependencies. From equation (9), we have:

$$(\rho_i + r_i, \rho_i + g_i, \rho_i + b_i) = (r_i, g_i, b_i) + (1, 1, 1)\rho_i \quad (10)$$

Equation (10) tells us in log RGB space, lighting geometry changing only affects the length of the log rgb vector in the direction of $U = (1, 1, 1)$. That is, the directions orthogonal to $(1, 1, 1)$ are unaffected by brightness change. It follows that we can normalize a log rgb to remove brightness by projecting it onto the 2-dimensional space which is orthogonal to the line that spanned by U . We can do this by applying some simple results from linear algebra. We define a 3×3 projection matrix Pr for the space spanned by U , and a complementary projection matrix $[I - Pr]$ for the space which is orthogonal to the space spanned by U (where I denotes the 3×3 identity matrix). By definition these matrices have the property that $Pr^*(1, 1, 1)^T = (1, 1, 1)^T$ and $[I - Pr]^*(1, 1, 1)^T = (0, 0, 0)^T$. Pr is defined as^{16,6}:

$$Pr = U'(UU')^{-1}U = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad (11)$$

and so

$$I - Pr = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \quad (12)$$

If we write an rgb log-vector as $\underline{v} = x * (1, 1, 1) + y * (-1, 1, 0) + z * (1, 1, -2)$, it is straightforward to show that $Pr\underline{v} = x * (1, 1, 1)$ and $[I - Pr]\underline{v} = y * (-1, 1, 0) + z * (1, 1, -2)$. Looking at the structure of matrix $[I - Pr]$, we can see that the meaning of the matrix multiplication is that we subtract the mean log rgb from the r , g , and b values:

$$\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \frac{2a}{3} & -\frac{b}{3} & -\frac{c}{3} \\ \frac{2b}{3} & -\frac{a}{3} & -\frac{c}{3} \\ \frac{2c}{3} & -\frac{a}{3} & -\frac{b}{3} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a - \frac{a+b+c}{3} \\ b - \frac{a+b+c}{3} \\ c - \frac{a+b+c}{3} \end{pmatrix}$$

That is, we can remove dependency on lighting geometry by subtracting the mean log response (a 'brightness' correlate) from each pixel.

The effect of illuminant colour can be removed in a similar way. However, rather than dealing with log rgb vectors we must operate on the vector of all log red responses (or log green responses or log blue responses). From (9) we can write

$$(\alpha + r_1, \alpha + r_2, \dots, \alpha + r_n) \rightarrow \alpha(r_1, r_2, \dots, r_n) + (1, 1, \dots, 1)$$

It follows that the following projection matrices (which are $n \times n$ for an n pixel image) will respectively project a colour channel in the direction of all ones $(1, 1, \dots, 1)$ or the space orthogonal to this direction.

$$Pc = \begin{pmatrix} \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{pmatrix} \quad (13)$$

$$I - Pc = \begin{pmatrix} \frac{n-1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{n-1}{n} & \dots & -\frac{1}{n} \\ \vdots & \vdots & & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \dots & \frac{n-1}{n} \end{pmatrix} \quad (14)$$

Again, by inspecting the structure of the projection matrix, we can see that to implement this normalisation we only need to subtract the mean red log value from all log red pixel values and subtract the mean log green and log blue values from the log green and log blue pixel values.

To remove lighting geometry and illuminant colour both at the same time we just apply the projectors (12) and (14) to the log image. This operation is easy to write down mathematically if we think of an n pixel image as an $n \times 3$ matrix of log rgbs. Denoting this image Y , we can write down an explicit equation for the application of the lighting geometry and light colour normalizations:

$$Y_{\text{normalised}} = (I - Pc)Y(I - Pr) \quad (15)$$

here $Y_{\text{normalised}}$ represents the normalised image. Though, it is important to remember that to implement (16) we

simply subtract row means from rows and then column means from columns.

From projection theory, we know that matrix $[I - Pr]$ and $[I - Pc]$ are both idempotent. That is to say that $[I - Pr][I - Pr] = [I - Pr]$ and $[I - Pc][I - Pc] = [I - Pc]$. It follows then that removing shading or light colour once, removes it for all time:

$$\begin{aligned} Y_{normalised} &= (I - Pc)(I - Pc)Y(I - Pr)(I - Pr) \\ &= (I - Pc)Y(I - Pr) \end{aligned} \quad (16)$$

In summary, the projection of image matrix in log space onto two spaces which are orthogonal to either the lighting geometry change or the illuminant colour change are invariant to lighting geometry or illumination colour change. In practice, post multiplying the log image matrix by $[I - Pr]$ is equivalent to subtracting the mean of the row from each element of log image matrix. Similarly, premultiplying the log image matrix by projection matrix $[I - Pc]$ is equivalent to subtracting the mean of the column from each element of log image matrix.

4. Object Recognition Experiments

We now wished to test the new non iterative comprehensive normalisation procedure for colour object recognition. To do this take an image, normalize it and then build its colour histogram. This histogram is compared with normalized histograms stored in a database and the closest overall is found. Because, the database contains images (and their histograms) of the same objects for which we have query images, the closet histogram match can be used to identify the query.

We applied this work flow for the composite dataset that was used to test the original comprehensive normali-

zation procedure.¹⁰ This dataset is composed of Swain and Ballard²¹ image set (66 database images 32 queries), the *legacy* Simon Fraser image set¹² (11 database images, 22 queries) and the Berwick and Lee image set⁴ (10 database images and 10 queries). A second larger Simon Fraser image dataset (www.cs.sfu.ca/~colour/data) with 20 objects imaged under 11 illuminations is also used as a test set.

Tables 1 and 2 summarize the indexing performance for the two normalisation operations and operation on composite and the new larger Simon Fraser dataset. We show the average percentile match and the % of objects in the 1st, 2nd and worse than 2nd ranks. If the closet database histogram to the query is the correct answer (both corresponding images are of the same object) then the correct answer is found in rank 1. If the correct answer is the k th closest then the correct answer has rank k .

We see that the comprehensive normalization delivers excellent recognition and this confirms previous reported results. The log-normalization performs if anything slightly better. It is particularly pleasing to see that the worst case for the large Simon Fraser data set has diminished from 16th for comprehensive normalization to just 5th for the log normalization method.

5. Conclusion

We have developed a new closed form log space comprehensive normalisation method which can cancel out the colour dependencies due to lighting geometry and illuminant colour both. Experimental results on image datasets showed that this new normalisation method performed at least as well as the iterative (non closed form) comprehensive normalisation method.

Table 1. Indexing Performance of Composite Dataset (ranks are % of the dataset)

Methods	Av.Percentile(%)	Rank 1	Rank 2	Rank > 2	Worst Rank
log comprehensive	99.71	95.38	1.15	3.47	4th out of 87
comprehensive	99.71	92.31	2.3	5.39	3th out of 87

Table 2. Indexing Performance on Large Simon Fraser Dataset (ranks are % of the dataset)

Methods	Av.Percentile(%)	Rank 1	Rank 2	Rank > 2	Worst Rank
log comprehensive	99.13	91.00	3.5	5.5	5th out of 20
comprehensive	98.84	91.00	3.5	5.5	16th out of 20

References

1. Donald A. Adjeroh and M. C. Lee. *On Ratio-Based Color Indexing*. IEEE transaction on image processing, vol(10), 2001.
2. J Bach, C. Fuller, A. Gupta, and A. Hampapur. *The virage image search engine: An open framework for image management*. In SPIE Conf. on Storage and Retrieval for Image and Video Databases, volume 2670, pages 76-87, 1996.
3. J. Berens, G. D. Finlayson, and G. Qiu. Image indexing using compressed colour histograms. *Vision, Image and Signal Processing*, 147(4):349-355, 2000.
4. D. Berwick and S. W. Lee. *A chromaticity space for specularly for specularly-, illumination color- and illumination pose-invariant 3-d object recognition*. ICCV98, Session2.3, 1998.
5. G. Buchsbaum. *A spatial processor model for object colour perception*. Journal of the Franklin Institute, 310:1-26, 1980.

6. F. Chatelin. *Eigenvalues of Matrices*. Wiley, 1993.
7. G. D. Finlayson. *Colour Object Recognition*. Simon Fraer Univ. School of Computing Science, 1992.
8. G. D. Finlayson. *Color in perspective*. IEEE Trans. PSML, 18(10):1034-1038, 1996.
9. G. D. Finlayson, M. S. Drew, and B. V. Funt. *Spectral sharpening: Sensor transformations for improved color constancy*. J. Optical Society of America A, 11(5):1553-1563, 1994.
10. Graham D. Finlayson, Bernt Schiele, and James L. Crowley. *Comprehensive Colour Image Normalisation*. In 5th European Conference on Computer Vision, page 475-490, 1998.
11. M. D. Zmura and G. Iverson. *Color constancy, basic theory of two-stage linear recovery of spectral description for lights and surfaces*. JOSA, 10:2148-2165, 1993.
12. B. V. Funt, K. Barnard, and L. Martin. *Is machine colour constancy good enough*. In *The Fifth European Conference on Computer Vision (Vol II)*, pages 445-459. European Vision Society, 1998.
13. B. V. Funt and G. D. Finlayson. *Color constant color indexing*. PAMI, 17(5):522-529, May 1995.
14. Ron Gershon, Allan D. Jepson, and John K. Tsotsos. *From [R,G,B] to Surface Reflectance: Computing Color Constant Descriptors in Images*. In International Joint Conference on Artificial Intelligence, pages 755-758, 1987.
15. G. Gevers and A. W. M. Smeulders. *Color based object recognition*. Pattern Recognition, 32:453-464, 1999.
16. G. H. Golub and C. F. van Loan. *Matrix Computations*. John Hopkins U. Press, 1983.
17. E. H. Land. *The retinex theory of color constancy*. Scientific American, pages 108-129, 1977.
18. M. J. Swain and D. H. Ballard. *Color indexing*. International Journal of Computer Vision, 7(11):11-32, 1991.
19. S. Lin and S. W. Lee. *Using chromaticity distributions and eigenspace analysis for pose-, illumination- and specularly-invariant recognition of 3d object*. In CVPR97, pages 426-431, 1997.
20. W. Niblack and R. Barber. *The QBIC project: Querying images by content using color, texture and shape*. In Storage and Retrieval for Image and Video Databases I, volume 1908 of SPIE Proceedings Series, 1993.
21. M. J. Swain and D. H. Ballard. *Color indexing*. International Journal of Computer Vision, 7(11):11-32, 1991.
22. B. V. Funt, K. Barnard, and L. Martin. *Is colour constancy good enough?* In 5th European Conference on Computer Vision, pages 445-459, 1998.
23. B. V. Funt and Graham D. Finlayson. *Color constant color indexing*. IEEE Trans. Pattern Recognition and Machine Intelligence, 17(5):522-529, 1991.
24. R. W. G. *The Reproduction of Color*. Fountain Press, 5th edition, 1995.
25. J. A. Worthey and M. H. Brill. *Heuristic analysis of von Kries color constancy*. Journal of The Optical Society of America A, 3(10):1708-1712, 1986.