

Colour Matching Function Learning

Luis Romero-Ortega¹ and Antonio Robles-Kelly^{1,2,3}

¹ School of Eng. and Inf. Tech., UNSW@ADFA, Canberra ACT 2600, Australia

² NICTA *, Locked Bag 8001, Canberra ACT 2601, Australia

³ Research School of Eng., ANU, Canberra ACT 0200, Australia

Abstract. In this paper, we aim at learning the colour matching functions making use of hyperspectral and trichromatic imagery. The method presented here is quite general in nature, being data driven and devoid of constrained setups. Here, we adopt a probabilistic formulation so as to recover the colour matching functions directly from trichromatic and hyperspectral pixel pairs. To do this, we derive a log-likelihood function which is governed by both, the spectra-to-colour equivalence and a generative model for the colour matching functions. Cast into a probabilistic setting, we employ the EM algorithm for purposes of *maximum a posteriori* inference, where the M-step is effected making use of Levenberg-Marquardt optimisation. We present results on real-world data and provide a quantitative analysis based upon a colour calibration chart.

1 Introduction

The accurate capture and reproduction of colours as acquired by digital camera sensors is an active area of research. This is not a straightforward task since digital cameras are comprised by three spectral broad-band color sensors which are not colorimetric. This implies that the RGB values yielded by the camera are not a linear combination of the device-independent CIE color matching functions [1]. Further, colours, as acquired by digital cameras, are, in general, device dependent.

Whereas colorimetry focuses on the accuracy of the colours acquired by the camera, spectroscopy has as object of study the spectrum of light absorbed, transmitted, reflected or emitted by objects and illuminants in the scene. In contrast with trichromatic sensors, multispectral and hyperspectral sensing devices can acquire wavelength-indexed reflectance and radiance data in tens of hundreds of bands across a broad spectral range. Recently, there has been renewed interest in multispectral imaging as related to color constancy [2], the analysis of the spectral properties of objects in the scene [3] and the optimal multiplexing of bandpass filtered illumination [4,5].

Moreover, making use of photogrammetry and spectroscopy techniques based upon monochromatic narrow-band illuminants, it is possible to recover the spectral response of the camera under study [6]. Methods which employ calibration

* NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

targets and charts have also been proposed. These methods employ quadratic programming [7], monochromators [8] or spectrophotometers [9].

The methods above require calibration charts and, in many cases, complicated setups and constrained environments. Here we present a method which employs two sets of images, one trichromatic and another one hyperspectral so as to learn the colour matching functions. The method presented here hinges in an inference process based upon a maximum-likelihood formulation which leads to the application of the EM (Expectation-Maximisation) algorithm. In the following section, we provide some background on the relationship between the colour output of the camera, the colour matching functions and the spectral image radiance.

2 Background

To better understand the relation between the spectra in hyperspectral imagery and the color output of trichromatic cameras, we commence by providing some background on the expression of the image radiance at pixel v as given in [10]. Let the image radiance be given by

$$I(\lambda, v) = L(\lambda)P(\varphi_i, \phi_i, \varphi_s, \phi_s)S(v, \lambda) \quad (1)$$

where $P(\cdot)$ is the mean scattered power in the direction (φ_s, ϕ_s) , $L(\lambda)$ is the power spectrum of the light impinging on the object surface in the direction (φ_i, ϕ_i) and $S(v, \lambda)$ is the surface reflectance at wavelength λ .

The expression above is important since it allows the use of the shorthand $R(\lambda, v) = P(\varphi_i, \phi_i, \varphi_s, \phi_s)S(v, \lambda)$ so as to write the image radiance as follows

$$I(\lambda, v) = L(\lambda)R(\lambda, v) \quad (2)$$

This expression has been used widely in the literature [11] and is consistent with reflectance models in the computer vision literature, such as that in [12].

Recall that in a trichromatic camera, a fraction of the light incident on the surface of the object being observed is reflected towards the camera. The light then passes through the camera lens, which focuses the incoming light beam onto the image plane of the camera. After reaching the image plane, the colour channel values for each pixel in the image are determined by the responses to the incoming light of the R, G, B receptors of the camera.

This is important since it permits us to consider two sample-sets. We denote the first of these, which corresponds to trichromatic pixels, as \mathcal{A}^{RGB} . The second sample-set, \mathcal{A}^{HS} , corresponds to the spectra at hyperspectral pixels. If the trichromatic pixel u is a match to the hyperspectral pixel v , i.e. $u \sim v$, then, using the notation above we can write

$$I_k(u) = \kappa_k \sum_{\lambda \in \mathcal{A}} Q_k(\lambda)L(\lambda)R(\lambda, v) \quad (3)$$

where κ_k is a constant that depends on the sensor geometry, $I_k(u)$ is the colour value for the channel $k = \{R, G, B\}$, \mathcal{A} is the visible spectral range,

i.e. $\Lambda = [400nm, 700nm]$, and the colour matching functions for the three colour channels are denoted by $Q_k(\lambda)$, $k = \{R, G, B\}$.

Here, we note that the illuminant power spectrum may be recovered making use of methods elsewhere in the literature such as that in [13]. We can assume it to be in hand. Also, in the following section, we assume that the matches between pixels are available. We will elaborate further on this in Section 4.

3 Maximum Likelihood Formulation

Note that the treatment above permits us to view each pixel value as a product of a per-pixel, per-wavelength factor which applies equally to all the pixels in the sample sets \mathcal{A}^{RGB} and \mathcal{A}^{HS} . This in turn allows a statistical treatment of the problem. In this section, we cast the problem into an Expectation-Maximization setting.

3.1 Log-Likelihood Function

The idea underpinning the EM algorithm is to recover maximum likelihood solutions to problems involving missing or hidden data. To do this, we view the colour matching functions as a set of hidden variables to be estimated. Thus, we cast the problem as a *maximum a posteriori* (MAP) one which aims at maximizing the probability of the colour matching function given the input trichromatic and hyperspectral image pixels. This can be expressed as follows

$$P(Q_k(\lambda) | \Omega, \Theta_k) = P(Q_k(\lambda) | \Omega)P(Q_k(\lambda) | \Theta_k)$$

where Ω is the set of matching spectra-colour pixel tuples and Θ_k is the set of hyperparameters for the colour matching function $Q_k(\lambda)$.

Note that, in the expression above, the first term of the right-hand side is the conditional probability governed by the hyperspectral image radiance value and the corresponding colour value. Thus, the maximization of the probability $P(Q_k(\lambda) | \Omega)$ implies that the colour matching function $Q_k(\lambda)$ should satisfy the relationship between the spectral radiance and the trichromatic colour values. The second term accounts for the dependency of the colour matching function $Q_k(\lambda)$ upon the hyperparameters in Θ_k . Note that these hyperparameters can be viewed as a means to enforcing a cumulative distribution function in a manner akin to histogram equalization methods [14].

We can take our analysis further by considering a probability distribution function for $P(Q_k(\lambda) | \Omega)$ of the form

$$P(Q_k(\cdot) | \Omega) = \frac{1}{\gamma_k \sqrt{2\pi}} \prod_{\substack{u \in \mathcal{A}^{RGB} \\ v \in \mathcal{A}^{HS} \\ u \sim v}} \exp \left\{ -\frac{1}{2\gamma_k^2} \left| I_k(u) - \kappa_k \sum_{\lambda \in \Lambda} Q_k(\lambda) I(\lambda, v) \right|^2 \right\} \quad (4)$$

where γ_k is the variance variable and the second term in the argument of the exponential function arises from Equations 2 and 3.

In a similar fashion, we can consider the colour matching function values to be distributed in accordance to a mixture of N Gaussians. This is a good approximation for the colour matching functions used in practice by camera manufacturers [15]. As a result, we can write

$$P(Q_k(\cdot) | \Theta_k) = \prod_{\lambda \in \Lambda} \sum_{q=1}^N \alpha_{q,k} \frac{1}{\sigma_{q,k} \sqrt{2\pi}} \exp \left\{ - \frac{(\mu_{q,k} - \lambda)^2}{2\sigma_{q,k}^2} \right\}$$

where, $\Theta_k = \{\theta_{q,k}\}_{q=1}^N$ and, as usual, $\theta_{q,k} = \{\mu_{q,k}, \sigma_{q,k}\}$ are the mean and the covariance for the k^{th} colour response and the q^{th} Gaussian in the mixture and $\alpha_{q,k}$ is the mixture weight.

With these ingredients, the log-likelihood function becomes

$$\begin{aligned} \mathcal{L}(Q_k(\cdot) | \Omega, \Theta_k) = & -\frac{1}{2\gamma_k^2} \sum_{\substack{u \in \mathcal{A}^{RGB} \\ v \in \mathcal{A}^{HS} \\ u \sim v}} \left| I_k(u) - \kappa_k \sum_{\lambda \in \Lambda} Q_k(\lambda) I(\lambda, v) \right|^2 + \\ & \sum_{\lambda \in \Lambda} \log \left\{ \sum_{q=1}^N \alpha_{q,k} \frac{1}{\sigma_{q,k} \sqrt{2\pi}} \exp \left\{ - \frac{(\mu_{q,k} - \lambda)^2}{2\sigma_{q,k}^2} \right\} \right\} \end{aligned} \tag{5}$$

where we have removed the term $\log \left\{ \frac{1}{\gamma_k \sqrt{2\pi}} \right\}$ from further consideration since it does not depend on the colour matching function $Q_k(\cdot)$ or the hyperparameter set and, hence, does not affect the inference process.

3.2 Expectation-Maximization

Note that, in the equation above, only the last term in the right-hand side depends on the hyperparameter-set Θ_k . This is an important observation, since it suggests an iterative update scheme in which the hyperparameters and the colour matching functions $Q_k(\lambda)$ can be recovered using the EM algorithm [16], which we describe in the following.

Expected Log-Likelihood Function. In Equation 5, the two terms on the right-hand side are log-likelihoods in their own right. That is, the first of these is the log-likelihood of $Q_k(\lambda)$ given the colour responses and spectral values. The second term corresponds to the likelihood of $Q_k(\lambda)$ given the hyperparameters Θ_k .

Thus, we can index the expected log-likelihood to iteration number n and write

$$\begin{aligned} \mathcal{L}(Q_k^{n+1}(\cdot) | \Omega, \Theta_k^n) = & -\tau \sum_{\substack{u \in \mathcal{A}^{RGB} \\ v \in \mathcal{A}^{HS} \\ u \sim v}} \left| I_k(u) - \kappa_k \sum_{\lambda \in \Lambda} Q_k^{n+1}(\cdot) I(\lambda, v) \right|^2 + \\ & \sum_{\lambda \in \Lambda} \log \left\{ \sum_{q=1}^N \alpha_{q,k}^n \frac{1}{\sigma_{q,k}^n \sqrt{2\pi}} \exp \left\{ - \frac{(\mu_{q,k}^n - \lambda)^2}{2(\sigma_{q,k}^n)^2} \right\} \right\} \end{aligned} \tag{6}$$

where we have used the shorthand $\tau = \frac{1}{2\gamma_k^2}$.

M-Step. In the M-step, we aim at maximising the expected log-likelihood with respect to the colour matching functions. Note that the maximisation of the log-likelihood can also be cast as a minimisation of the form

$$\operatorname{argmin}_{Q_k^{n+1}(\cdot)} \left\{ \tau \sum_{\substack{u \in \mathcal{A}^{RGB} \\ v \in \mathcal{A}^{HS} \\ u \sim v}} \left| I_k(u) - \kappa_k \sum_{\lambda \in \Lambda} Q_k^{n+1}(\cdot) I(\lambda, v) \right|^2 - \sum_{\lambda \in \Lambda} \log \left\{ \sum_{q=1}^N \alpha_{q,k}^n \frac{1}{\sigma_{q,k}^n \sqrt{2\pi}} \exp \left\{ -\frac{(\mu_{q,k}^n - \lambda)^2}{2(\sigma_{q,k}^n)^2} \right\} \right\} \right\} \quad (7)$$

This observation is important since it allows the M-step to be viewed as a regularised nonlinear least-squares minimisation which can be tackled using the Levenberg-Marquardt algorithm (LMA), which is an iterative trust region procedure [17] aimed at recovering a numerical solution to the problem of minimising a function over a space of parameters.

E-Step. To estimate the hyperparameter set $\theta_{q,k}^{n+1}$, we introduce the posterior probability

$$P(\theta_{q,k}^n | Q_k^{n+1}(\lambda)) = \frac{P(\theta_{q,k}^n | Q_k^{n+1}(\lambda)) P(Q_k^{n+1}(\lambda))}{P(\theta_{q,k}^n)} \quad (8)$$

so as to write the gradient of the log-likelihood function as follows

$$\nabla_{\theta_k^n} \mathcal{L}(Q_k^{n+1}(\cdot) | \theta_k^n) = \sum_{\lambda \in \Lambda} P(\theta_k^n | Q_k^{n+1}(\lambda)) \nabla_{\theta_k^n} \log \{ P(\theta_k^n | Q_k^{n+1}(\lambda)) \} \quad (9)$$

Recall that the maximum likelihood corresponds to the values of θ_k^{n+1} for which $\nabla_{\theta_k^{n+1}} \mathcal{L}(Q_k^{n+1}(\cdot) | \theta_k^n) = 0$. Since we have assumed $P(Q_k^{n+1}(\lambda) | \theta_k^n)$ to be a mixture of Gaussians, we can recover the maximum likelihood estimates of θ_k^n by differentiating Equation 5 with respect to θ_k^n , substitute the results into Equation 9 and solve $\nabla_{\theta_k^n} \mathcal{L}(Q_k^{n+1}(\lambda) | \theta_k^n) = 0$.

This is a well-known estimation problem [18], which after some algebraic manipulation, yields the following update rules

$$\mu_{q,k}^{n+1} = \frac{\sum_{\lambda \in \Lambda} \lambda h_{q,k}^n(\lambda)}{\sum_{\lambda \in \Lambda} h_{q,k}^n(\lambda)}$$

$$\sigma_{q,k}^{n+1} = \frac{\sum_{\lambda \in \Lambda} (\mu_{q,k}^{n+1} - \lambda)^2 h_{q,k}^n(\lambda)}{\sum_{\lambda \in \Lambda} h_{q,k}^n(\lambda)} \quad (10)$$

$$\alpha_{q,k}^{n+1} = \frac{1}{|\Lambda|} \sum_{\lambda \in \Lambda} h_{q,k}^n(\lambda) \quad (11)$$

where $P(Q_k^{n+1}(\lambda) | \theta_{q,k}^n)$ can be computed, in a straightforward manner, making use of Equation 5 and

$$h_{q,k}^n(\lambda) = \frac{P(Q_k^{n+1}(\lambda) | \theta_{q,k}^n)}{\sum_{r=1}^N P(Q_k^{n+1}(\lambda) | \theta_{r,k}^n)} \quad (12)$$

where we have followed [19] and set

$$P(Q_k^{n+1}(\lambda) | \theta_{q,k}^n) = \frac{1}{\Psi} \sum_{q=1}^N \alpha_{q,k}^n \frac{Q_k^{n+1}(\lambda)}{\sigma_{q,k}^n \sqrt{2\pi}} \exp \left\{ -\frac{(\mu_{q,k}^n - \lambda)^2}{2(\sigma_{q,k}^n)^2} \right\}$$

and Ψ is a normalisation constant.

4 Implementation Issues

Having presented the theoretical background of our approach, we now turn to the implementation of the method. Note that our method is not limited to a particular number of bands and applies equally to each of the three colour channels. Moreover, so far, we have assumed that the colour-spectra matches, i.e. $I_k(u)$ and $I(\cdot, v)$ with $u \sim v$ are available. In practice, this is not the case. Moreover, acquiring spectro-colourimetric image pairs is impractical in many cases due the error that may be introduced by registering the two views, i.e. that captured by a trichromatic camera and that acquired using the hyperspectral imager.

Thus, we opt for a discriminative approach based upon two code books [20] with sufficient amount of samples so as to have statistical relevance. We commence by building two pixel sets. The first of these from images taken using the colour camera for which we aim at learning the colour matching functions and the other one from imagery captured using the hyperspectral imager. Then, we build the two codebooks, one for each of these pixel-sets. We do this making use of k -means clustering [18].

The codebook for the hyperspectral sample is then converted into RGB values making use of the current estimate of the colour matching functions. This permits matches between the two codebooks to be recovered making use of nearest-neighbours in the RGB chromaticity colour space. These codebooks are, hence, used as an alternative to \mathcal{A}^{RGB} and \mathcal{A}^{HS} , where the elements indexed u and v in the respective sets are a match to each other if they are the nearest neighbour to one another.

5 Experiments

Here we show results on real-world multispectral and trichromatic imagery. To this end, we have used two trichromatic commercial cameras, i.e. Nikon D80 and Nikon D5100, so as to acquire 210 pictures, 105 images with each camera. For each of the color cameras, we have used approximately 5.2 Mega pixels sampled

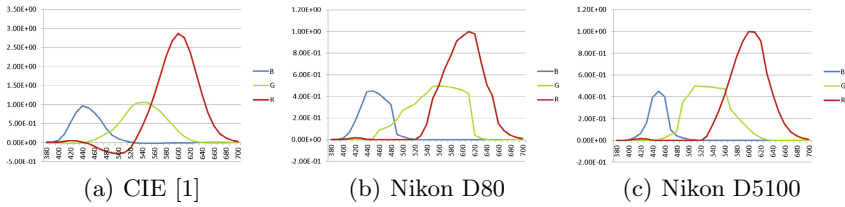


Fig. 1. Colour matching functions for the CIE 1955 standard and those recovered by our algorithm for the Nikon D80 and Nikon D5100 camera models



Fig. 2. From left-to-right: imagery rendered with the CIE colour matching functions and those recovered by our algorithm for the Nikon D80 and Nikon D5100 camera models

over the 105 images in a grid like fashion. All our trichromatic imagery has been acquired in raw output mode with manual exposure calibration.

For our hyperspectral data, we have used 1 Mpixel taken from 131 images. Similarly to our trichromatic data, these have been sampled in a grid like fashion with 7.6 Kpixels selected from each image using 16×16 pixel tiles. All our hyperspectral imagery was acquired with a Liquid Crystal Tunable Filter (LCTF) at 2 Mpixel resolution and 33 spectral bands in the visible range in $10nm$ steps. In all our experiments, the number of mixtures N is set to two and initialised the colour matching functions and hyperparameters making use of the CIE1955 standard [1]. Here we have used $\tau = 1$ and iterated until the L_2 -norm between the hyperparameter set Θ^n and Θ^{n+1} is below a user-provided threshold, which we set to $1e^{-15}$. In average, the algorithm converged in 10 iterations.

In Figure 1, we show the colour matching functions for the CIE 1955 standard [1] and those recovered using our approach. Note that the CIE standard does contain negative values, whereas the colour matching functions for commercial cameras, by definition, should be positive. In Figure 2, we show example results for one of our input images. In the figure, we show the image rendered with the CIE colour matching functions and those recovered by our algorithm. Note the differences between the images. In particular, with respect to the CIE colour matching functions.

To provide a more quantitative result, we have acquired imagery, hyperspectral and trichromatic, for the XRite Color Checker chart. This is a tiled

Table 1. Mean and standard deviation for the error when using the CIE color matching functions and those learnt by our method

	Nikon D80	Nikon D5100		Nikon D80	Nikon D5100
CIE Matching Functions [1]	0.048 ± 0.025	0.040 ± 0.019	Our Method	0.032 ± 0.017	0.036 ± 0.0215

colorimetric calibration board containing 12 colours and 6 shades of gray. We have rendered the hyperspectral image of the colour chequer with the CIE colour matching functions and those learnt by our method. Once the RGB images are generated, we compare the colours on the chart with those on the imagery acquired with the trichromatic cameras. To do this, we have performed white balancing using the shades of gray and computed the mean-squared differences between the trichromatic imagery and that yielded by the colour matching functions applied to the hyperspectral image.

In Table 1, we show the mean and standard deviation for the CIE colour matching functions and those learnt by our method. In the table, the mean and standard deviation have been normalised to be between zero and unity. This is so as to allow scale variations between the two sets of colour matching functions. This also permits comparison with colour difference measures often used in colorimetry. Note that our learning method outperforms the CIE color matching functions for both cameras.

6 Conclusions

In this paper, we have introduced an approach aimed at learning the colour matching functions from hyperspectral and trichromatic imagery. We do this based upon a probabilistic formulation where the EM algorithm is employed so as to recover the colour matching functions directly from trichromatic and hyperspectral pixel pairs. We have derived a log-likelihood function which is governed by both, the accordance of the spectra-to-colour equivalence and a generative model for the colour matching functions. The method is quite general in nature, being data driven and devoid of constrained setups. Our results on real-world data show that our method is capable of learning colour matching functions which deliver colours in close accordance to those acquired by sample trichromatic cameras.

References

1. Stiles, W.S., Burch, J.M.: Interim report to the Commission Internationale de l'Éclairage Zurich, 1955, on the National Physical Laboratory's investigation of colour-matching. *Optica Acta* 2, 168–181 (1955)
2. Wandell, B.A.: The synthesis and analysis of color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9(1), 2–13 (1987)

3. Schechner, Y., Nayar, S.: Uncontrolled modulation imaging. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. II:97–II:204 (2004)
4. Schechner, Y., Nayar, S., Belhumeur, P.N.: A theory of multiplexed illumination. In: Int. Conference on Computer Vision, pp. II:808–II:816 (2003)
5. Chi, C., Yoo, H., Ben-Ezra, M.: Multi-spectral imaging by optimized wide band illumination. *Int. Journal of Computer Vision* 86, 140–151 (2010)
6. Vora, P.L., Farrell, J.E., Tietz, J.D., Brainard, D.H.: Linear models for digital cameras. In: Proceedings of the IS&T's 50th Annual Conference, pp. 377–382 (1997)
7. Finlayson, G., Hordley, S., Hubel, P.: Recovering device sensitivities with quadratic programming. In: Proceedings of the IS&T/SID Color Imaging Conference, pp. 90–95 (1998)
8. Urban, P., Desch, M., Happel, K., Spiehl, D.: Recovering camera sensitivities using target-based reflectances captured under multiple led-illuminations. In: Proceedings of the Workshop on Color Image Processing, pp. 9–16 (2010)
9. Martin, R., Arno, Z., Reinhard, K.: Practical spectral characterization of trichromatic cameras. In: SIGGRAPH Asia (2011)
10. Horn, B.K.P., Brooks, M.J.: The variational approach to shape from shading. *CVGIP* 33(2), 174–208 (1986)
11. Kimmel, R., Elad, M., Shaked, D., Keshet, R., Sobel, I.: A variational framework for retinex. *International Journal of Computer Vision* 52(1), 7–23 (2003)
12. Finlayson, G.D., Schaefer, G.: Solving for colour constancy using a constrained dichromatic reflection model. *International Journal of Computer Vision* 42(3), 127–144 (2001)
13. Huynh, C.P., Robles-Kelly, A.: A solution of the dichromatic model for multispectral photometric invariance. *International Journal of Computer Vision* 90(1), 1–27 (2010)
14. Gonzalez, R.C., Wintz, P.: *Digital Image Processing*. Addison-Wesley Publishing Company Limited (1986)
15. Vora, P., Farrell, J., Tietz, J., Brainard, D.: Image capture: Simulation of sensor responses from hyperspectral images. *IEEE Transactions on Image Processing* 10(2), 307–316 (2001)
16. Dempster, A., Laird, N., Rubin, D.: Maximum-likelihood from incomplete data via the EM algorithm. *J. Royal Statistical Soc. Ser. B Methodological* 39, 1–38 (1977)
17. Nocedal, J., Wright, S.: *Numerical Optimization*. Springer (2000)
18. Bishop, C.M.: *Pattern Recognition and Machine Learning*. Springer (2006)
19. Barber, D.: *Bayesian Reasoning and Machine Learning*. Cambridge University Press (2012)
20. Dalai, N., Triggs, B.: Histogram of oriented gradients for human detection. In: *Computer Vision and Pattern Recognition*, pp. I:886–I:893 (2005)