A Multi-Spectral Image Database and its Application to Image Rendering Across Illumination

Steven Hordley, Graham Finalyson and Peter Morovic
School of Computing Sciences, University of East Anglia
NR4 7TJ, Norwich, United Kingdom

Abstract

We present a database of full spatial and spectral resolution images of everyday objects captured with a novel multi-spectral device: the Applied Spectral Imaging © Spectracube *camera. We summarise the operation of the device and present a study of its spectral accuracy before demonstrating the usefulness of the data set by using it to evaluate different* White Point Correction *transforms: transforms which map RGBs under one illuminant to a corresponding set of RGBs under a different light. We find that while the best transforms can often give results close to the results obtained using spectral techniques, the results are both image and illuminant dependent and in a number of cases all transforms tested introduce significant errors.*

1. Introduction

We begin (§2) by providing an overview of the principle of operation of the *Applied Spectral Imaging © Spectracube* camera. This device is an interferometry based, semi-portable digital camera which is able to capture in a single exposure a full 2D spatial array of spectra (i.e. a spectral image). We report the results of a characterisation experiment in which we compared the Spectracube camera to a conventional telespectraradiometer (a Photo Research PR-650). Our experiments show that best results are obtained with the Spectracube when it is used as a device for measuring *reflectance* rather than *radiance* images. Next (§3) we present a database of multi-spectral reflectance images which we have acquired with the Spectracube device. This database consists of a set of high spatial and spectral resolution reflectance images of everyday objects captured under calibrated viewing conditions which we believe will be a useful resource for any researcher interested in the "gap" between conventional (RGB) imaging and multi–spectral capture. We demonstrate (§4) the usefulness of the database by using it to investigate an important problem in image reproduction: *White Point Correction*: the problem of transforming an image captured under an arbitrary illuminant to the corresponding image that would have been recorded under a second illuminant.

2. The Spectracube Camera

The Spectracube camera consists of a lens, an interferometer, a 2D CCD array, and an image processor. Light from a point in an imaged scene enters the lens and subsequently the interferometer where it is first split into two beams. One beam travels through the beam splitter and falls on a fixed mirror while the second beam is reflected at right angles to the first where it falls upon another mirror whose distance from the beam splitter can be varied. Both mirrors reflect the respective beams impinging on them back through the beam splitter and they are eventually reunited at a detector: in the case of the Spectracube, a single element of the CCD array. By varying the distance of the moveable mirror from the beamsplitter, the difference in the length of the paths which the two beams travel can be varied. The total intensity of the reunited beams varies as a function of the optical path difference (OPD). For a given OPD, the CCD of the Spectracube camera measures the intensity of the reunited beams from each point in the scene simultaneously. This 2D array of measurements is called a frame. Many frames are acquired in sequence with the OPD being varied each time. If n frames are captured then for a given pixel in the CCD array there exist n samples for each scene point: each sample corresponds to the intensity of the reunited beams from a given scene point for a given OPD.. This set of measurements is referred to as the *interferogram* and importantly can be shown [1] to be the inverse Fourier Cosine Transform of the spectrum of the original beam (that is, the spectrum of the light at the corresponding scene location). Thus applying a Fourier Cosine Transform to the captured frames results in a spectral image of the scene (more detail can be found in [3]).

We compared the Spectracube data to that captured with a conventional point measurement telespectraradiometer (the Photo Research PR-650). We took as our characterisation data measurements of 180 coloured patches from a Macbeth Digital Colorchecker Chart under two fluorescent daylight simulators (D50 and D75) and a Tungsten source (Ill A). Let us represent measurements from the PR-650 and the Spectracube by the 1×31 vectors $\mathbf{p}_i$ and $\mathbf{s}_i$ respectively. Each measurement corresponds to a uniform 10nm sampling of radiance from a scene point in the 400nm to 700nm wavelength interval.

To compare the measurements from the two devices we used two error measures. First we evaluated difference in the spectral space by calculating the Percent Relative Spectral Error:

$$e_{prse} = 100 \frac{\|\mathbf{p}_i - \mathbf{s}_i\|_2}{\|\mathbf{p}_i\|_2} \tag{1}$$

In addition we evaluated the CIE ΔE error in the near perceptually uniform *CIE Lab* [8] colour space, to assess the visibility of these spectral differences. The results for each illuminant over all 180 patches are summarised in columns 2-5 of Table 1. The errors between the two

	Radiance		Reflectance	
Ill	% Rel. Error	ΔE	% Rel. Error	ΔE
D50	32.67	2.53	6.69	3.06
D75	27.19	2.66	3.99	1.40
A	14.30	7.19	20.22	12.28

Table 1. Summary of median errors between spectra from the PR-650 and the Spectracube.

devices are significant: spectral errors of around 30% for both D75 and D50 and close to 15% for Ill A represent large errors in all cases. The *CIE Lab* errors are also significant: a ΔE of 1 corresponds to a just noticeable difference when two uniform samples are viewed side by side and the mean and median errors for all illuminants are greater than this. However, there is evidence to suggest [11] that errors of up to 6 ΔE units are not detected in complex images. With this in mind we might conclude that the agreement between the two instruments is good enough at least for illuminants D50 and D75. The Spectracube camera is known to have low sensitivity in the wavelength region 380nm to 450nm and since Tungsten sources have little power in this region the data captured in that region under Ill A is likely quite noisy and is the most probable explanation for the relatively worse agreement under Ill A.

Further investigation of the measurements under D50 and D75 revealed that under D55 and D75 differences between Spectracube and PR650 spectra were quite systematic: the spectra are similar but slightly shifted. This is most likely caused by the fact that these sources essentially consist of three sharp peaks of energy. Small differences in how such a spectrum is sampled leads to significant differences in the sampled spectrum. However, in general reflectance spectra are quite smooth so that if we divide an arbitrary radiance measurement by a measurement of the illuminant spectrum any sharp peaks in the captured spectra are removed. Thus we proceed by considering *reflectance* rather than *radiance* measurements. We derive reflectances by capturing a scene under a fixed illuminant and dividing the spectrum at each pixel of that scene by the spectrum of the scene illuminant. Columns 6-9 of Table 1 summarise inter-device agreement in terms of these reflectance measurements and confirm that for both illuminant D75 and D50 much better inter-instrument agreement results.

3. A Multi-Spectral Image Database

Having characterised the Spectracube camera we compiled a database of reflectance images. We chose to capture reflectance rather than radiance since this allows us to light scenes so as to ensure that the reflectance data we record is as accurate as possible and it allows users of the images to synthesise radiance images under arbitrary illuminants. We note that the radiance images obtained in this way will likely be different to those which would be obtained by capturing scene radiance directly but the improved accuracy of the reflectance data and the increased flexibility in rendering options outweigh this drawback. The database currently consists of 22 images. Each imaged scene consists of one or more objects placed in a viewing booth and was lit using D75 illumination which our characterisation study shows gives the best agreement to the PR-650 data. For each scene we took two images using the Spectracube camera. First we imaged the scene itself. Then we placed a uniformly reflecting (white) tile centrally in the scene and took a second picture of the modified scene. We used this second image to obtain (from the white tile) a measure of the spectrum of the illumination in the scene. Our estimate of the scene illuminant was taken by averaging the pixel values of a 5×5 square of pixels from the centre of the white tile. We then divided each spectrum in the first captured image by this estimate of the scene illuminant to obtain a reflectance at each image location. The spectral database can be found online at http://www.cmp.uea.ac.uk/∼pm/chromagenic/msdb in the form of a set of MATLAB data files.

4. White Point Correction

In this section we demonstrate the usefulness of the multi-spectral database by using the images to investigate the relative accuracy of white point correction in a number of different RGB spaces. To begin we define a simple Lambertian [7] model of image formation. Under this model the response of a 3-sensor device with spectral sensitivity functions $Q_k(\lambda)$ $(k = 1, 2, 3)$ to a surface with reflectance function $S(\lambda)$ viewed under an illuminant with spectral power distribution $E^c(\lambda)$ can be written:

$$p_k = \int E^c(\lambda)S(\lambda)Q_k(\lambda)\mathrm{d}\lambda \quad k = 1, 2, 3 \qquad (2)$$

where λ denotes wavelength. $E^c(\lambda)$ defines the energy emitted by the light source at each wavelength of the visible spectrum. The responses p_k represent the total energy absorbed by the sensors integrated over all wavelengths. By contrast let us also consider the response of an ideal *spectral* imaging device whose sensor response functions Q'_j are Dirac delta functions $(Q'_j(\lambda) = \delta(\lambda - \lambda_j))$ where the number of sensors is now $M >> 3$ and is sufficiently large to ensure an accurate sampling of lights and surface. In this case the response of the j^{th} sensor becomes:

$$p'_j = E^c(\lambda_j)S(\lambda_j) \;\; j \in [1, M] \qquad (3)$$

Suppose that we capture an image of a scene under a (known) illuminant $E^c(\lambda)$ using either the 3-sensor device or the spectral device. The *White Point Correction* problem can then be stated as follows: How can we predict the responses of the device to the same scene when it is viewed under a second (known) illuminant $E^o(\lambda)$?

In the case of the spectral device, it is easy to make the prediction. Since $E^c(\lambda)$ and $E^o(\lambda)$ are known we can simply divide the sensor responses by $E^c(\lambda)$ and multiply them by $E^o(\lambda)$:

$$p^{o\prime}_j = E^c(\lambda_j)S(\lambda_j) \times \frac{E^o(\lambda_j)}{E^c(\lambda_j)} = E^o(\lambda_j)S(\lambda_j) \;\; j \in [1, M] \qquad (4)$$

In the case of the 3-sensor device, changing the illuminant is more difficult since the illuminant spectrum is confounded in the integral of Eq. (2). In practice 3-sensor images are usually corrected for a change of illuminant by applying an analogous correction to Eq. (4). That is, given the sensor response $\underline{p}^c$ for an arbitrary surface under illuminant $E^c(\lambda)$ we first divide that response by $\underline{p}^c_w$: the sensor response to a white (neutral) surface under $E^c(\lambda)$ and then multiply by $\underline{p}^o_w$: the sensor response to a white (neutral) surface under $E^o(\lambda)$:

$$p^o_k = p^c_k \times \frac{p^c_{w,k}}{p^o_{w,k}} \quad k = 1, 2, 3 \qquad (5)$$

Such a scheme will not in general give the same result as would be obtained by substituting $E^c(\lambda)$ for $E^o(\lambda)$ in Eq. (2) but in practice the results are often quite close. The success of this procedure depends in part on the colour space (the sensor functions $Q_k(\lambda)$) in which it is performed. For example, it is known that if Q_k correspond to the XYZ colour matching functions [8] of our own visual system then such a procedure will give quite poor results. However, it is possible to improve these results by transforming the XYZ responses to a different colour space, applying the correction and then transforming back to the original XYZ space. Mathematically this procedure is written:

$$\underline{p}^o = T^{-1}DT\underline{p}^c \qquad (6)$$

where T is the 3×3 transform which maps XYZ responses into the appropriate space for applying the white point correction and D is a 3×3 diagonal matrix whose k^{th} diagonal element is $r^o_{w,k}/r^c_{w,k}$ and $\underline{r}^?_w = T\underline{p}^?_w$. When T is a transformation of the XYZ colour matching functions of our own visual system the term *chromatic adaptation transform* is used. There is an ongoing debate as to what constitutes an appropriate space (an appropriate choice of T) in which to perform white point correction [4, 5, 9, 10] but as yet no consensus has emerged. Part of the difficulty in determining the most appropriate choice of T is the fact that it is difficult to obtain ground truth data: i.e. spectrally rendered images of the same scene under different illuminants. With the multi-spectral database introduced in this paper we have access to just such data and so we can use the images to evaluate the performance of different approaches.

We evaluate correction accuracy for four different transforms mapping XYZ responses into a different 3-sensor colour space. The *Cone* space transform [8] maps XYZ responses into the space defined by the sensitivities of the cone cells of our own eyes. The Bradford (*BFD*) Transform [9] was derived based on the optimisation of a set of corresponding colour data gathered in a psycho-physical experiment. The *CMC* Transform is the chromatic adaptation transform included in the CIECAM02 [12] Colour Appearance Model and is again derived based on the optimisation of a set of psycho-physical data. Finally the *Sharp* Transform [2] maps to a space defined by theoretical sensor functions with narrow support under the assumption that in the case that sensor sensitivities are perfectly narrow band, a diagonal mapping affords perfect white-point correction. There are a number of different methods for defining a "Sharp" transform. In this paper we use the white-point preserving data-based sharpening approach [5].

We test the methods as follows. First we calculate XYZ responses for a reflectance image under a reference light (D65) using Eq. (2). Next, we rendered the same reflectance image under a different illuminant. Then, we attempt to re-render this second image so that it matches the first image using Eq. (5) where T is replaced by one of the 4 transforms introduced above. We measure the performance of a given transform T by measuring the difference between the spectrally rendered image (under the second light) and the image rendered in the 3-sensor space defined by the transform T. The results, summarised in Table 2, show performance for four different lights: a daylight (D55), a warm and a cool fluorescent (WWF and CWF) and a tungsten bulb (Ill A). Performance is given in terms of median *CIE Lab* ΔE error (M) and in terms of the number of times a given transform significantly better than all other transforms on a given image pair when assessed according to Wilcoxon's Sign Test [6] (S)s. It is clear that transform performance depends on the illumination: all give acceptable performance when the illuminant is D55 (an illuminant very similar to D65) but for other lights such as CWF performance is not always good: no transform provides acceptable performance for all images under this light. In general the performance of the Cone based transform is poor: it gives acceptable results only under D55. The other three transforms all generally work quite well with the Sharp transform outperforming all other transforms on the majority of images.

	Cone		BFD		CMCCAT		Sharp	
	M	S	M	S	M	S	M	S
D55	1.01	0	0.70	3	0.67	0	0.67	17
Ill A	6.44	0	4.08	2	4.33	0	3.46	20
CWF	6.84	0	5.54	5	5.48	2	5.27	15
WWF	5.65	0	4.08	0	3.91	1	3.23	21

Table 2. Summary of the errors for performing white point correction in four different colour spaces. Results are averaged over 22 images.

5. Conclusions

The study above reveals that none of the state-of-the-art White Point Correction transforms provides an acceptable level of performance in all situations. By using a set of spectral images we have demonstrated that the performance of all transforms is both image and illuminant dependent and in a significant number of cases RGB based corrections are unable to produce a rendering sufficiently close to the spectrally rendered image. Thus we conclude that further research is required in this area to address these shortcomings.

References

[1] J. Chamberlin. *The Principles of Interferometric Spectroscopy.* John Wiley & Sons, 1979.

[2] G. D. Finlayson, M. S. Drew, and B. V. Funt. Spectral Sharpening: sensor transformations for improved color constancy. *Journal of the Optical Society of America, A*, 11(5):1553–1563, 1994.

[3] G. D. Finlayson, S. D. Hordley, and P. M. Morovic. Using the Spectracube to build a multispectral image database. In *2nd Computer Graphics, Imaging and Vision Conference.* IS&T/SID, April 2004.

[4] G. D. Finlayson and S. Susstrunk. Performance of a Chromatic Adaptation Transform based on Spectral Sharpening. In *8th Color Imaging Conference.* IS&T/SID, November 2000.

[5] G. D. Finlayson and S. Susstrunk. Spectral Sharpening and the Bradford Transform. In *CIS2000*, pages 236–243, April 2000.

[6] R. V. Hogg and E. A. Tanis. *Probability and Statistical Inference.* Prentice Hall, 2001.

[7] B. K. Horn. *Robot Vision.* MIT Press, 1986.

[8] R. Hunt. *The Reproduction of Colour.* Fountain Press, 5th edition, 1995.

[9] K. M. Lam. *Metamerism and Colour Constancy.* PhD thesis, University of Bradford, 1985.

[10] C. Li, M. R. Luo, and B. Rigg. Simplification of the CMCCAT97. In *The Eighth Color Imaging Conference*, November 2000.

[11] G. W. Meyer. Wavelength selection for synthetic image generation. *Comp. Vision, Graphics, and Image Proc.*, 79:41–57, 1988.

[12] N. Moroney, M. D. Fairchild, R. W. G. Hunt, C. Li, R. Luo, and T. Newman. The CIECAM02 color appearance model. In *10th Color Imaging Conference.* IS&T/SID, November 2002.