Color Selections for Characterization Charts

T.L.V. Cheung and S. Westland
School of Design, University of Leeds, Leeds, United Kingdom

Abstract

Accurate color measurement can be achieved using a trichromatic digital camera if the device is characterized in terms of CIE tristimulus values. A common practical consideration for any characterization method is the choice of the characterization target. The Macbeth ColorChecker DC chart, for example, is widely used for color-characterization tasks. Whereas a great deal of work has been carried out to address which characterization method gives the best performance, rather less work has been carried out to investigate which characterization target is optimum for the characterization process. This paper describes methods to select optimum colors from a large data set of 1269 Munsell colors. The effect of color selection on characterization performance is compared (using a third-order polynomial transform) with performance using the 24 Macbeth ColorChecker chips and 166 chips from the Macbeth ColorChecker DC chart.

Introduction

Digital cameras can effectively be used as tristimulus colorimeters if they characterized in terms of CIE tristimulus values.[1,2] Some researchers[3-5] have suggested that multispectral imaging or, more generally, spectral techniques may be useful for the characterization of imaging devices such as cameras and scanners. Thus, a possible device-characterization method is to try to recover the spectral properties of the surfaces in the scene and then compute the tristimulus values from these estimated reflectances.[4] A common practical consideration for any of these characterization methods is the choice of the characterization target. The Macbeth ColorChecker and Macbeth ColorChecker DC charts are widely used as color characterization tasks.[6-9] Whereas a great deal of work has been carried out to address which characterization method gives the best performance,[2,10-12] rather less work has been carried out to investigate which characterization target is optimum for the characterization process. This paper describes methods to select optimum colors from a set of 1269 Munsell colors.[13] The effect of color selection on characterization performance is evaluated (using a third-order polynomial transform) and compared with performance using the 24 Macbeth ColorChecker colors and 166 Macbeth ColorChecker DC colors.

Methodology

Hardeberg[14] proposed a method to select a set of reflectance samples that would be most suitable for the estimation of camera spectral sensitivity. The approach adopted by Hardeberg was to select spectra from a large set of Munsell samples such that each selected spectrum was as different as possible (in reflectance space) from the other already selected spectra. Hardeberg compared this so-called optimal selection method with a heuristic method (whereby samples with the highest chroma were selected for each hue) and with the samples from the Macbeth ColorChecker. He found that performance (in terms of accuracy of estimation of the camera spectral sensitivities) was almost as good using 20 optimally selected spectra as it was using the full set of 1269 Munsell spectra.[14] It seems clear that the selection of samples for a color chart might be expected to have a substantial effect on the usefulness of that chart for camera characterization. We have considered three methods for the selection of samples to constitute a characterization chart. The simplest method is to randomly select either 24 or 166 samples from the full set of 1269 Munsell samples. These particular numbers were selected so that we could compare performance with the 24 samples of the Macbeth ColorChecker chart and 166 unique samples from the Macbeth ColorChecker DC chart. When the samples were selected randomly the whole set was selected 10 times. Two other methods, called here Method 1 and Method 2, were also investigated. For Methods 1 and 2 the selection of the first sample is arbitrary and in this research a spectral reflectance with the biggest variance has been chosen from the data set as being the first sample. However, subsequent samples are selected according to certain optimal or sub-optimal procedures in a way similar Hardeberg's technique.[14]

In Method 1, *n-1* samples are selected in turn (without replacement) from the pool of Munsell samples such that the *ith* sample ($2 < i \leq n$) is selected to maximize the value of Q_j which is defined thus

$$Q_j = \sum_{i=1}^{i-1} \Delta E_{j,i}^{\;1/2} \tag{1}$$

where $\Delta E_{j,i}$ represents the CIELAB color difference between the sample *j* and the *ith* selected sample for the D65 illuminant.

In Method 2, *n-1* samples are selected in turn (without replacement) from the pool of Munsell samples such that the *ith* sample ($2 < i \leq n$) is selected to maximize the value of P_j which is defined thus

$$P_j = \min_{i=1}^{i-1} \left[\Delta E_{j,i} \right] \tag{2}$$

The main idea is to generate a chart where the samples are as different to each other as possible. The idea is inspired by the earlier work of Hardeberg[14] but uses colorimetric metrics rather than spectral metrics to determine how 'different' samples are from each other.

Furthermore, Methods 1 and 2 are subtly different in that in Method 1 we find the sample that is on average as different as possible from the already selected samples whereas in Method 2 we find the sample whose closest neighbor (in the already selected samples) is as far away as possible. In both methods, however, the *ith* sample selected for the chart is selected so that it is as far away from the already selected samples as possible.

The 24 colors selected from each method (Method 1, Method 2, Random) were used to construct a virtual characterization chart. A linear camera model (Equation 3) with known and fixed camera channel sensitivities and a known illuminant were used and the coefficients of a third-order polynomial were determined to provide the least-square mapping between sample camera responses *RGB* and tristimulus values *XYZ*. The *RGB* values were computed using

$$\begin{aligned} R &= \sum E(\lambda) S_R(\lambda) P(\lambda) \\ G &= \sum E(\lambda) S_G(\lambda) P(\lambda) \\ B &= \sum E(\lambda) S_B(\lambda) P(\lambda) \end{aligned} \tag{3}$$

where $E(\lambda)$ is the spectral power distribution of the illuminant, $S_R(\lambda)$, $S_G(\lambda)$ and $S_B(\lambda)$ are channel spectral sensitivities of the camera system and $P(\lambda)$ is the spectral reflectance of the surface. The 3 × 20 third-order polynomial transform with the following terms,[12]

$[R\ G\ B\ RG\ RB\ GB\ R^2\ G^2\ B^2\ RGB\ R^2G\ R^2B\ G^2R\ G^2B\ B^2R\ B^2G\ R^3\ G^3\ B^3\ 1]$

was used to map from *RGB* to *XYZ*.

The 24 colors selected from each of the methods were used as training sets and characterization performance was evaluated for three testing sets: 1269 Munsell samples, 50 Natural Color System (NCS) samples and 494 natural surfaces contain leaves, petals, grasses and barks. The procedure was repeated using a larger number (166) of selected samples since many typical characterization charts in common use often contain a few hundred samples.

Results

The color distributions for the training and testing samples are visualized in the $L^*a^*b^*$ color space and shown in Figure 1 and Figure 2 respectively.

Table 1 shows the CIELAB color differences for the training performance using different sets of 24 training samples. The median training errors as shown in Table 1 are generally very small for each of the training sets. It is no surprise that it is possible to randomly select a set of samples that would be more accurately fitted using the polynomial transform than the sets selected using Methods 1 or 2. This is because the randomly selected samples may all occupy one small region in color space. This is why performance should be evaluated using the independent testing sets. Tables 2 to 4 show the testing performance for each of the three testing sets.

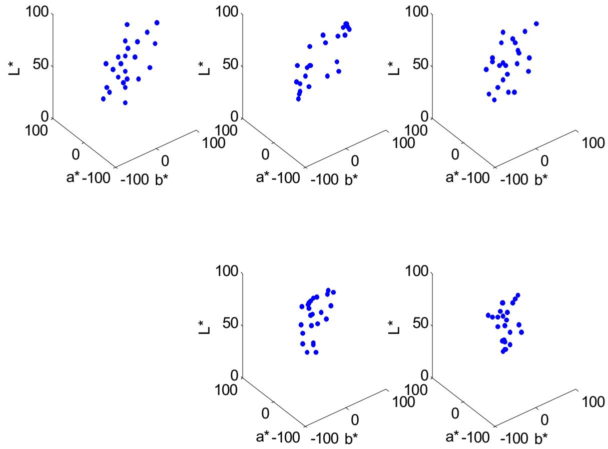

Figure 1. Color distributions in CIELAB space for 24 Macbeth ColorChecker (top left), 24 Method 1 selected (top middle), 24 Method 2 selected (top right), 24 Best Random (bottom middle), 24 Worse Random (bottom right) samples

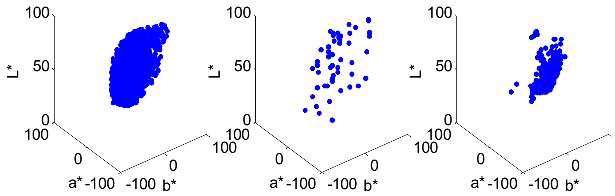

Figure 2. Color distributions in CIELAB space for 1269 Munsell (left), 50 NCS (middle) and 494 natural (right) samples

Table 1. Training performance (3 × 20 polynomial transform) using different 24 characterization samples

	mean	median	min	max
ColorChecker	0.9605	1.2978	0.0562	3.5757
Method 1	0.4094	0.5566	0.0393	1.7665
Method 2	0.5402	0.7893	0.0232	2.4832
Best random	0.2767	0.3731	0.0253	1.0044
Worse random	0.4881	0.5775	0.0424	1.9022

In general, training sets selected using Method 2 outperform the Macbeth ColorChecker and sets selected using Method 1 samples for all different testing sets (Tables 2-4). It is likely that the poor characterization performance using Method 1 occurs because the selection algorithm chose samples located around the boundary for an a^*b^* diagram and the colors have little variation of L^* values. Similarly, in Hardeberg's work the heuristic method of selecting samples of high chroma performed quite poorly.[14]

The finding that 24 samples from the Munsell set can be selected to provide a color chart that outperforms the Macbeth ColorChcker chart is interesting. However, for most practical work on camera characterization a larger set of samples such as the Macbeth ColorChecker DC is used. For such larger training charts, is it possible to use selection algorithms that give enhanced performance?

Table 2. Testing performance (3 × 20 polynomial transform) on 1269 Munsell samples using different 24 characterization samples

	mean	median	min	max
ColorChecker	2.6446	3.1508	0.0870	15.5986
Method 1	7.6956	11.2463	0.0393	45.0489
Method 2	1.6188	2.0417	0.0232	12.7893
Best random	2.0075	3.5099	0.0253	22.6865
Worse random	3.0843	5.8594	0.0037	103.081

Table 3. Testing performance (3 × 20 polynomial transform) on 50 NCS samples using different 24 characterization samples

	mean	median	min	max
ColorChecker	5.1047	6.7939	0.3976	35.9383
Method 1	10.6863	15.3869	1.6286	148.623
Method 2	4.9291	7.4240	0.2016	37.0342
Best random	7.8804	14.5629	0.4484	80.1972
Worse random	14.1560	33.4745	0.5815	211.971

Table 4. Testing performance (3 × 20 polynomial transform) on 494 Natural samples using different 24 characterization samples

	mean	median	min	max
ColorChecker	2.3062	3.1970	0.1213	14.2437
Method 1	7.9100	11.1994	0.5419	54.2463
Method 2	1.1746	1.5611	0.0824	7.7807
Best random	2.2750	3.1302	0.1176	15.6847
Worse random	4.6145	5.3141	0.3752	123.062

In Figure 3 the Macbeth ColorChecker DC samples (using the central 166 samples) are compared with 166 samples selected using Method 2. Tables 5 to 8 show the characterization performances in terms of CIELAB color differences. From the tables it is evident that the chart selected using Method 2 performs better than the Macbeth DC chart for the natural (Table 8) and Munsell (Table 6) samples but a little worse for the NCS (Table 7) samples.

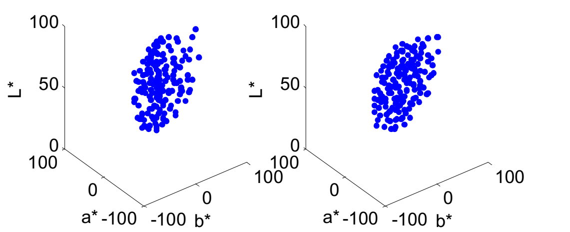

Figure 3. Color distributions in CIELAB space for 166 Macbeth ColorChecker DC (left) and 166 Method 2 selected (right) samples

Table 5. Training performance (3 × 20 polynomial transform) using different 166 characterization samples

	Mean	median	min	max
ColorChecker DC	1.0520	1.3125	0.0662	4.9712
Method 2	1.2811	1.5056	0.1026	7.0685

Table 6. Testing performance (3 × 20 polynomial transform) on 1269 Munsell samples using different 166 characterization samples

	Mean	median	min	max
ColorChecker DC	1.2094	1.5308	0.0778	8.5431
Method 2	1.1110	1.3883	0.0415	8.5179

Table 7. Testing performance (3 × 20 polynomial transform) on 50 NCS samples using different 166 characterization samples

	mean	median	min	max
ColorChecker DC	1.7203	2.8127	0.0266	17.1270
Method 2	2.0447	3.7057	0.2563	21.7783

Table 8. Testing performance (3 × 20 polynomial transform) on 494 natural samples using different 166 characterization samples

	mean	median	min	max
ColorChecker DC	2.2001	2.4257	0.1009	7.7301
Method 2	1.5701	1.6881	0.2563	5.2081

Discussion

Generally, the performance of the standard charts, Macbeth ColorChecker (24) and Macbeth ColorChecker DC (166) is quite good compared to the new methods. This would seem to indicate that the samples for these charts were well selected and are appropriate for the purposes of device characterization. However, some performance gains were evident using charts selected using Method 2 when compared with the Macbeth ColorChecker and Macbeth ColorChecker DC charts. We note, however, that in this work the selection methods were used to select from a relatively small number of Munsell reflectances. It may be that better performance can be obtained if the number of samples from which the methods select is increased and, more importantly, the gamut of these samples is increased. One possible approach, which the authors are currently exploring, is to use a linear model of basis functions rather than a limited pool of samples so that the characterization charts could contain samples that are highly saturated and yet are physically reproducible.

Moreover, this paper raises the question of how to assess the performance of data-driven camera characterization methods. If we wish to characterize a camera to perform well for a certain domain of samples then the ideal data with which to characterize the camera would be a set of samples that possessed the same statistical properties as the samples of the domain. For certain, well-defined, problems it may be possible to ascertain the statistics of the domain and derive an appropriate set of characterization samples. For example, if we wish to develop a camera system to measure the color of bananas or teeth it would be sensible to select a characterization set containing many yellow or white samples respectively. For the development of an optimum characterization set for the general problem of color measurement further work needs to be carried out to ascertain the statistics of natural and man-made colored samples in the world.[15]

References

1. Johnson T., (1996), Methods for characterising colour scanners and digital cameras, *Displays*, **16** (4), 183-191.
2. Hong G., Luo M.R., and Rhodes P.A. (2001), A study of digital camera colorimetric characterisation based on polynomial modelling, *Color Research and Application*, **26** (1), 76-84.
3. Berns R.S., and Shyu M.J. (1995), Colorimetric characterization of a desktop drum scanner using a spectral model, *Journal of Electronic Imaging*, **4** (4), 360-372.
4. Finlayson G.D., and Morovic P.M. (1999), Metamer constrained colour correction, *Proceedings of the 7th Color Imaging Science Conference (Scottsdale, USA)*, 26-31.
5. Shi M. and Healey G., (2002), Using reflectance models for color scanner calibration, *Journal of the Optical Society of America A*, **19** (4), 645-656.
6. Farrell J.E. and Wandell B.A. (1993), Scanner linearity, *Journal of Electronic Imaging*, **2** (3), 225-230.
7. Burns P.D. and Berns R.S. (1996), Analysis multispectral image capture, *Proceedings of the 4th Color Imaging Conference (Scottsdale, USA)*, 19-22.
8. Finlayson G.D., Hubel P.M. and Hordley S.D. (1997), Color by correlation, *Proceedings of the 5th Color Imaging Science Conference (Scottsdale, USA)*, 6-11.
9. MacDonald L.W. and Ji W., (2002), Colour characterization of a high-resolution digital cameras, *Proceedings of the 1st European Conference on Colour in Graphics, Image and Vision (Poitiers, France)*, 433-437.
10. Cheung T.L.V. and Westland S. (2002), Color camera characterisation using artificial neural networks, *Proceedings of the 10th Color Imaging Conference (Scottsdale, USA)*, 117-120.
11. Cheung T.L.V. and Westland S. (2003), An evaluation of multispectral imaging techniques for camera characterization, *Proceedings of the 11th Color Imaging Conference (Scottsdale, USA)*, 193-198.
12. Cheung T.L.V., Westland S., Connah D.R. and Ripamonti C. (2004), A comparative study of the characterization of colour cameras by means of neural networks and polynomial transforms, *in press for Journal of Coloration Technology*.
13. Parkkinen J.P.S., Hallikainen J. and Jaaskelainen T. (1989), Characteristic spectra of Munsell colors, *Journal of the Optical Society of America A*, **6** (2), 318-322.
14. Hardeberg J.Y. (1999), Acquisition and reproduction of colour images: colorimetric and multispectral approaches, *PhD thesis*, Ecole Nationale Superieure des Telecommunications (France).
15. Westland S., Shaw A.J. and Owens H.C. (2000), Colour statistics of natural and man-made surfaces, *Sensor Review*, **20** (1), 50-55.

Biography

Vien Cheung graduated from The Hong Kong Polytechnic University with a BSc degree in Textile Chemistry. She then obtained an MSc in Colour Imaging at the Colour & Imaging Institute at University of Derby. She is currently a postgraduate student at the School of Design at University of Leeds working on methods for device characterization and multispectral imaging with **Professor Stephen Westland**.