Accurate mapping of natural scenes radiance to cone activation space: a new image dataset

C.A. Párraga, R. Baldrich and M. Vanrell; Centre de Visió per Computador/ Computer Science Department, Universitat Autònoma de Barcelona, Edifício O, Campus UAB (Bellaterra), C.P.08193, Barcelona, Spain.

Authors' Biographies

Maria Vanrell is Associate Professor in the Computer Science Department of the Universitat Autonoma de Barcelona and is attached to the Computer Vision Center as a researcher. He received his Phd in Computer Science from the Unviersitat Autonoma de Barcelona in 1996. His research interest is mainly focused in color and texture in computer vision problems, including color constancy, texture description and color and texture grouping.

Ramon Baldrich is an Associate Professor in the Computer Science Department of the Universitat Autonoma de Barcelona and is attached to the Computer Vision Center as a researcher. He received his Phd in Computer Science from the Unviersitat Autonoma de Barcelona in 2001. His research interest is mainly focused in color treatment in computer vision problems, including color segmentation, color constancy, color induction and image shadows.

C. Alejandro Párraga graduated in Physics (UNT, Argentina) in 1993, was awarded his MSc. (Bristol, UK) in 1996 and his PhD (Bristol, UK) in 2003. He worked as a postdoctoral fellow at the Universities. of Cambridge, and Bristol in the UK.. He was awarded both the "Juan de la Cierva" (2005) and the "Ramon y Cajal" (2007) Research Fellowships in Computer Science by the Spanish Ministry of Science and Technology. He is based in the Computer Science Department of the Universitat Autonoma de Barcelona and is attached to the Computer Vision Center as a researcher since 2006.

Accurate mapping of natural scenes radiance to cone activation space: a new image dataset

Abstract

The characterization of trichromatic cameras is usually done in terms of a device-independent color space, such as the CIE 1931 XYZ space. This is indeed convenient since it allows the testing of results against colorimetric measures. We have characterized our camera to represent human cone activation by mapping the camera sensor's (RGB) responses to human (LMS) through a polynomial transformation, which can be "customized" according to the types of scenes we want to represent. Here we present a method to test the accuracy of the camera measures and a study on how the choice of training reflectances for the polynomial may alter the results.

Introduction

The last decade has seen an increasing interest in the interplay between the distinctive characteristics of biological sensory systems and those of the natural environment. In the case of vision, this interest reflects the growing evidence that the statistical properties (both spatial and chromatic) of the visual environment have contributed to shape the way in which our visual system (and that of other species) function. Consequently, much research is based on the analysis of the visual environment (considering the tasks that a living organism needs to perform in order to survive and its biological constraints) with the aim of learning about the statistical regularities that the visual system may have exploited in its development.

In his review work on the relationships between visual perception and the statistical properties of natural scenes, Geisler[1] points out that measuring within-domain statistics is central to testing for "efficient coding" (the hypothesis that the response characteristics of visual neurons can be predicted from the statistics of natural images plus some biological constraints). Only after knowing the probability distribution of the property considered, we can determine which is the most efficient way of coding it. To this respect it is important to point out that there is an exponential relationship between the number of samples required to estimate a probability distribution and the number of properties considered. In other words, the more complex regularities of the visual environment we want to map, the larger the number of scenes we need to gather: this is the main reason why scientists have so far concentrated on only a small group of properties which need few images to compute. Other reason has to do with technical limitations: when natural scene regularities are unrelated to the chromatic responses of the visual system (multiscale analysis, contours, etc.), the use of uncalibrated imagery is justified. However when they involve the chromatic domain, a more sophisticated approach is needed.

Two techniques and methods were initially tried to measure and compute the statistical regularities of nature in the chromatic domain: (a) spectroradiometric devices which measure spectral radiance (radiance as a function of wavelength) from a small patch of image at the time, obtaining information about illuminants and reflective material properties; (b) hyperspectral cameras which measure the same from a whole image at the time but require long exposures, etc. Both these methods are impractical for gathering large databases of in-the-field imagery: spectroradiometric devices do not capture the *spatial* properties of natural images and hyperspectral cameras are only useful for indoor environments or when there is little change in time (long distance shots, man-made structures, landscapes, etc.). A third method has been tried more recently to reach a compromise between speed, portability and accuracy: calibrated trichromatic cameras are fast and portable but do not provide the complete spectral information necessary to fully characterize the reflectance of every patch of the image, however they are the only way to record the statistics of large samples of the visual environment to date.

The latest advances in digital imaging have turned trichromatic cameras into the most common device for estimating/measuring the properties of natural scenes. Commercial digital cameras are relatively cheap and if properly calibrated they can provide photometric information for every region of the scene (i.e. a measure of the radiant power absorbed by each of the camera's sensors, for each pixel). Calibrating a digital camera generally involves converting the image captured in the camera (also called device-dependent) color space into a reference (or device-independent) color-space. There are currently several methods to produce this color space transformation (see Martinez-Verdú *et al* for a more detailed explanation[2]). One of such methods (the *spectroradiometric approach*) consists of obtaining the camera sensor's response to a narrowband monochromatic stimulus which is in turn varied to span the whole spectral sensitivity range of the camera[2-4]. Since the stimulus' radiometric characteristics are known, it is possible to reconstruct the sensor's spectral sensitivity at each wavelength (and for each color sensor). It is also necessary to measure the sensor's output dependency on radiant power to obtain a complete picture of the camera's response to light. Once the sensor's responses are known, it is possible to find an approximate way of transforming the camera's RGB values to any device-independent color space (commonly the CIE 1931 XYZ color space). A second approach to the characterization of digital cameras is based on mathematical models (*mathematical approach*) which estimate the camera's matching functions from the device RGB responses to a set of (known) spectral reflectances, such as the squares of the Macbeth ColorChecker card[5-8]. While the first approach is quite precise, it

is seldom used because of its complexity. The last method is easier to implement but it is very vulnerable to measurement noise. Some intermediate approaches rely on assigning an estimated function to the camera's sensors and performing a mapping of the camera's space by means of a "training set" of RGB responses and radiometric measurements[9-12]. Our approach is a mixture of the two: it consists of measuring the camera sensitivities by means of photographing a white target through a set of spectrally narrowband interference filters (*spectroradiometric approach*) while using a training set to "match" the theoretical camera output to a device independent space (*mathematical approach*)[13].

The problem of converting to cone activation space

Whatever solution is chosen for characterizing the camera output in terms of a device-independent color space, the further transformation of these values into a cone activation space is not without difficulties. Cone activation spaces are physiologically realistic alternatives to the already ubiquitous systems of specifying color adopted by the Commission Internationale de l'Eclairage (CIE), being the best established of these systems the CIE 1931. When the CIE 1931 systems was adopted, the spectral sensitivities of the actual photoreceptors in the human retina were not indisputably known and instead, a set of hypothetical primaries was adopted, based on the experiments of Guild[14] and Wright[15] to determine the human *color matching functions*. The trichromatic values XYZ of the CIE 1931 system can be understood as the photon catches of three arbitrary photoreceptors with spectral sensitivities determined by the so called $\overline{x}, \overline{y}, \overline{z}$ functions. These functions are approximately point-by-point linear transformations of the cone spectral sensitivities of an average human observer (in fact $\overline{z}$ is actually very close to the spectral sensitivity of human short-wavelength -or "S" cones and $\overline{y}$ was chosen to have the same shape as the standard function of luminous sensitivity or V_λ).

Despite the CIE 1931's popularity and some obvious advantages, a chromaticity system that is not based on human physiology (or any other physiology, as in this case) is of limited use for researching the neural properties of a visual system. To amend this situation, a number of physiologically-plausible chromatic systems have been adopted by the neuroscience community, being one of the most popular the MacLeod-Boynton[16] space. In the MacLeod-Boynton space, the axes correspond to two of the chromatic channels identified physiologically by Derrington *et al* [17] in the early visual system. In this space, physiologically significant loci are represented by horizontal and vertical lines. To make the situation more complicated, MacLeod-Boynton system is derived from the Smith and Pokorny[18] human cone sensitivities, which in turn are not exact point-by-point transformations of the CIE $\overline{x}, \overline{y}, \overline{z}$, but of the slightly different set of primaries calculated by Judd in 1951 (and tabulated by Vos in 1978) known as the Judd-Vos response functions[19]. These are favored in visual science because of its better estimate of luminosity at short wavelengths. There is a formula[19] for transforming between the chromaticity coordinates of the CIE 1931 and the Judd 1951 system but it is valid only for monochromatic lights. This means that to use the MacLeod-Boynton system or any other cone activation space derived from

the Smith and Pokorny (1975) sensitivities it is necessary to know the spectroradiometric properties of the stimulus.

The most straightforward way of avoiding the inconvenience of a two-part chromatic conversion of the stimulus (from device-dependent camera-RGB space to CIE 1931 XYZ and then to Smith and Pokorny LMS (L for *long*, M for *middle* and S for *short* wavelength) cone activation space with the consequent transformation errors, is to characterize the camera directly in terms of the later (LMS) space. This can be done if one already knows the camera's sensor spectral sensitivities by means of finding the best transformation between the two chromatic systems. In this work we have based our analysis in the Smith and Pokorny (1975) cone responses, which are calculated at the cornea (as opposite to cone "pigment" spectral sensitivities). However, an alternative LMS dataset based on the Stockman and Sharpe (2000) cone fundamentals[20] (which are currently adopted by the CIE) is also available for download from our website[13]. The results for both datasets are similar.

A transformation from camera-RGB space to cone-activation LMS space.

Accurate linear transformations between chromatic spaces can be implemented only if the spaces satisfy the Luther condition[21,22] that is, primaries should be linear combinations of each other. This means that for an exact transformation of any triplet representing the quantum catches of the three camera sensors to the equivalent catches of the human LMS cone sensors to exist, there must exist a linear transformation between each point of the camera's spectral sensitivities and the corresponding point of the human cone spectral sensitivities. This condition is extremely hard to satisfy which means that surface metamerism will indeed exist (unless the camera has been specifically designed for that purpose, which is hardly in the interest camera manufacturers, to say the least) however, several approximations can be attempted. One such approximation consists of a *mathematical approach* style and relies on the existence of a calibrated camera where the sensor sensitivities are known, and can be implemented as follows.

Suppose that we know the sensitivities of the camera sensors S_i, and a given sample of the types of reflectances R and illuminations I that we are likely to photograph (which are all function of the wavelength), then we can calculate the camera's output G for each sensor i, surface j and illumination k by using equation 1:

$$G_{i,j,k} = \int_\lambda S_i(\lambda) \cdot R_j(\lambda) \cdot I_k(\lambda) \cdot d\lambda;$$

where

$i = 1, 2, 3:$ red, green and bue sensors

$j = 1, 2, ..., N:$ reflectances most likely to encounter

$k = 1, 2, ..., M:$ illuminants most likely to encounter

(1)

On the other hand we can do the same calculations for the cone activation D values that the same combinations of reflectances and illuminants are likely to elicit when sampled by each of the human cone sensitivities C_i.

$$D_{i,j,k} = \int_\lambda C_i(\lambda) \cdot R_j(\lambda) \cdot I_k(\lambda) \cdot d\lambda;$$

where

$$i = 1, 2, 3 : \text{L, M and S cones} \qquad (2)$$

$$j = 1, 2, ..., N : \text{reflectances most likely to encounter}$$

$$k = 1, 2, ..., M : \text{illuminants most likely to encounter}$$

Our problem is now reduced to finding the best mathematical solution (mapping) to match each set: cone activations to camera sensors activations. To map *D* and *G* there is variety of techniques, from interpolation and lookup tables [23,24] to polynomial regressions [25-27], neural networks [9] and spectral reconstructions [28,29]) which may prove more or less efficient and have various pros and cons (for a comparison of some of these techniques see Cheung et al[9] and Hong et al[27]).

Polynomial solutions

For our transformation, we choose a very simple (least-squares) regression mapping method consisting of finding the optimal polynomial transform for an arbitrary set of reference surface reflectances. The choice of references may depend on the particular problem, e.g. one choice may optimize the colorimetric mapping for natural surface reflectances and give a higher colorimetric error for saturated blues or it may not transform the white appropriately. However, we can turn this multiplicity of choices to our advantage, since it gives us the freedom of deciding which colors are likely to be predominant in the images (or need to be specified precisely) and which ones are likely to withstand a larger colorimetric error. We have used a similar idea to calibrate our camera in the CIE 1931 XYZ color space[13]

For our regression mapping, both the outputs of the camera sensors *G* and those of the human sensors (cones) *D* for our reference dataset of *N* surface reflectances and *M* illuminations are treated as a set of *N*x*M* triplets and indexed with a single letter (*j*). The values of *G,* are then mapped to the values of *D* using a polynomial expansion as follows: let $G_{i,j}$ correspond to the jth vector of RGB (R corresponds to *i*=1, G to *i*=2 and B to *i*=3) values obtained by our camera from the combination of surface/illumination *j* (part of a set of *N*x*M* measurements) and let $D_{i,j}$ correspond to the LMS cone capture values (L corresponds to *i*=1, M to *i*=2 and S to *i*=3, to follow a similar notation) of the same combination. In the simplest case, our objective would be to find the matrix *T* that minimizes ε in the following expression:

$$\varepsilon = \sum_{i=1}^{3} \sum_{j=1}^{NxM} \left\| D_{i,j} - G_{i,j}T \right\|^2 \qquad (3)$$

However, in our case, the product *GT* was replaced by a polynomial expansion. Leaving the expression to minimize as:

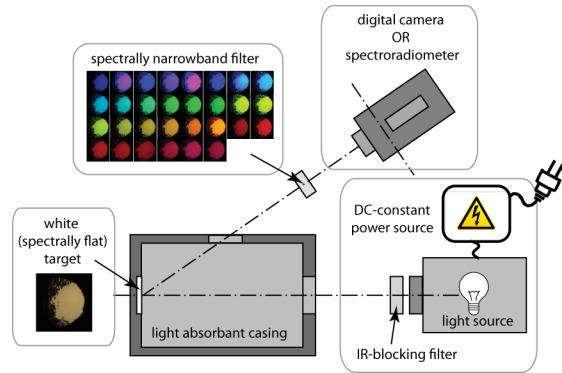$$\varepsilon = \sum_{i=1}^{3} \sum_{j=1}^{NxM} \left\| D_{i,j} - Q_j P \right\|^2 \qquad (4)$$

where *Q* represents the vector $Q_j = [G_{1,j}, G_{2,j}, G_{3,j}, G_{1,j}G_{2,j}, G_{2,j}G_{3,j}, G_{1,j}G_{3,j}]$ for each of the three sensors *i* and NxM measurements *j*. *P* is a matrix of 6x3 coefficients to be determined.

Both sets of NxM triplets can be obtained using equations 1 and 2, applied to our "reference" set of reflectances and

illuminations, the question now resides on how to choose the characteristics of these reference datasets.

Given that we do not know in advance which will be the main use of our characterization to LMS, we tried at first a generalist approach, considering a wide sample of known reflectances: a set of 1269 mate Munsell chips in the 400-700 nm range obtained from the COLORLAB[30] database and illuminated by a standard D65 illuminant.

This choice of training set should give a reasonable mapping over a wide color gamut. However, if we knew in advance that the scenes to be photographed were, say, landscape scenes, it might be possible to get a precise mapping by creating a *P* matrix optimized for reflectances such as those of chlorophyll, bark, sky, etc.

Figure 1: Scheme of the camera calibration set-up. The "target" consisted of a white patch inside a box which could be illuminated form one side and photographed though the other. Light reflected form the patch was either measured or photographed through spectrally narrowband filters. The tungsten-based illumination was supplied by a constant-current power source. The complete camera sensors calibration is described online alongside a calibrated picture dataset. (http://www.cvc.uab.es/color_calibration/).
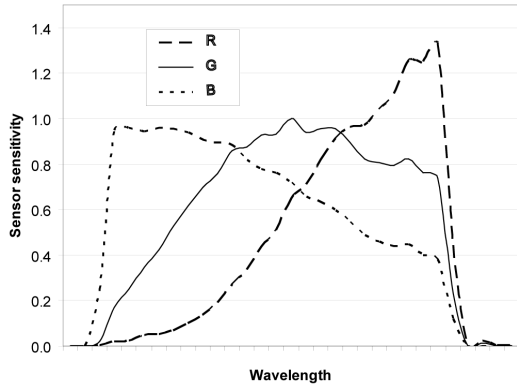
Methods

Camera calibration

The spectral sensitivities of the sensors of a trichromatic camera (Sigma Foveon SD10) were measured by recording its RGB sensor's responses to light transmitted by a set of 31 spectrally narrowband interference filters. These recordings were later compared to equivalent spectroradiometric measures. The camera's sensors dependency with light intensity and integration time was also measured by means of a Macbeth ColorChecker card. The light was produced by an IR-filtered, tungsten-halogen lamp (Osram HLX 64657FGX-24V, 250W) connected to a constant-current power-supply to ensure illumination stability during the whole process. To minimize measurement noise, the calibration was conducted inside a black room (walls were painted black). All measures were made with a telespectroradiometer (TopCon model SR1, calibrated by the UK's National Physical Laboratory). This instrument was capable of measuring spectral radiance within the 380-760 nm range. Its spectral radiance measurements were within the 4% limits specified by the manufacturer at the time of its calibration. A

complete description of the camera calibration and accuracy tests is currently online [13].

By means of the set-up described in Figure 1 we were able to measure the sensor's dependency with wavelength (including the camera lens at a particular configuration). The dependency with intensity was measured by photographing a Macbeth ColorChecker under fixed illumination several times, varying the camera's shutter speed (the sensor's integration time). The non-linearities of the sensors with regard to intensity were compensated by means of a gamma-correction function similar to that defined for CRT monitors. The sensor's dependency with wavelength is shown in Figure 2 (values have been scaled so that the middle-wavelength sensor's maximum value is equal to 1). In a normal setup, physical information about the camera lens (their aperture and focal length) and shutter speed is extracted from the picture header and incorporated automatically to the lens+sensor capture calculations. These curves are similar to those obtained for the same sensors (in isolated conditions) by Lyon and Hubel[31].

Figure 2: *Camera sensors sensitivity normalized to maximum = 1 for the Green sensor (i = 2).*

Polynomial solutions to equation 4 were found by solving for **P** equation 5, where **D** is the matrix of LMS cone responses calculated from equation 2 and **Q** is the expanded matrix of RGB camera responses calculated from the discrete version of equation 1 for all test reflectances.

$$\mathbf{D} = \mathbf{Q} \cdot \mathbf{P} , \qquad (5)$$

where

$$\mathbf{Q} = \begin{vmatrix} G_{11} & G_{21} & G_{31} & G_{11}G_{21} & G_{21}G_{31} & G_{11}G_{31} \\ G_{12} & G_{22} & G_{32} & G_{12}G_{22} & G_{22}G_{32} & G_{12}G_{32} \\ G_{1j} & G_{2j} & G_{3j} & G_{1j}G_{2j} & G_{2j}G_{3j} & G_{1j}G_{3j} \\ ... & ... & ... & ... & ... & ... \\ G_{1,J} & G_{2,J} & G_{3,J} & G_{1,J}G_{2,J} & G_{2,J}G_{3,J} & G_{1,J}G_{3,J} \end{vmatrix}$$
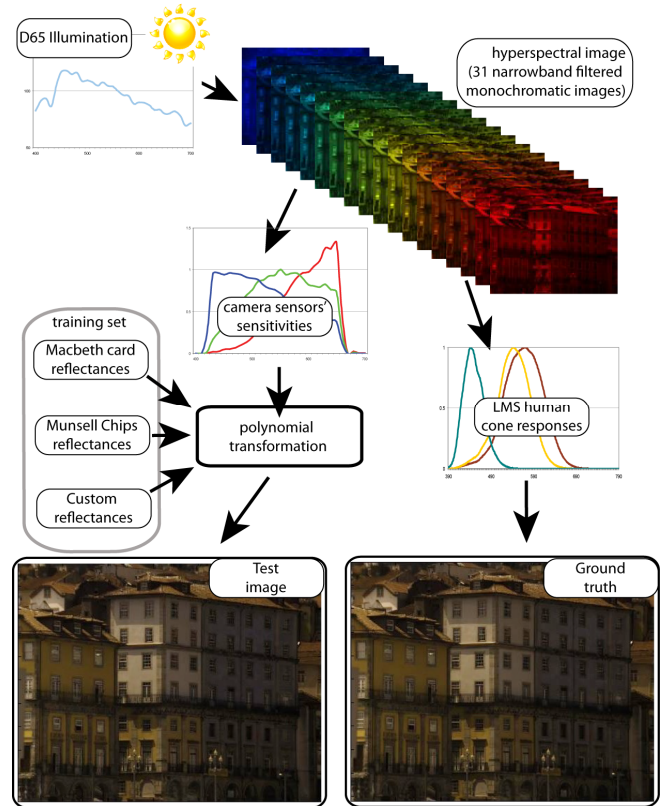
being $J = M \cdot N$, the total number of reflectances and illuminant combinations

The solution for P (calculated for 1269 Munsell chip reflectances) is a 6x3 matrix of coefficients, shown in Table 1.

-0.03642	0.158517	-0.0680	-1.63e-08	-1.52e-08	3.71e-08
-0.08864	0.192649	0.07204	-1.02e-08	-8.73e-09	2.23e-08
0.043564	0.13908	0.14357	-9.69e-09	-9.27e-09	2.17e-08

*Table 1: Exemplary set of coefficients obtained by solving equation 5 for **P** in Matlab. The coefficients applied to double products (columns 4, 5 and 6) are usually much smaller than those applied to single sensor outputs (columns 1,2 and 3), indicating that a simple 3x3 transformation may be adequate for many applications.*

The following section shows the "RGB to LMS" characterization errors obtained for several training sets of reflectances, tested on a dataset of eight calibrated hyperspectral images, available online from Manchester University [32] in the UK.

Figure 3: *Schematics of the experiment to test the accuracy of the LMS cone characterization of our trichromatic (RGB) camera. The workings of the camera were simulated by sampling a hyperspectral set of scenes (synthetically illuminated by a D65 illuminant) through the camera sensors' sensitivities. The results were then transformed to LMS by the polynomial described in equation 5 ("test image") and compared to the actual samplings of the same hyperspectral imagery by the human cone responses ("ground truth"). The pictures inside the figure show gamma-corrected versions of the actual results for an exemplary image (scene 7, taken from the hyperspectral database gathered by Foster et al[32]).*

Experiments and results

The quality of our camera characterization in terms of LMS-cone capture was measured for several different polynomial values, each one obtained for a different training set of reflectances. The training sets used here were in the 400-700 nm interval, illuminated by a simulated D65 (CIE standard daylight illuminant) and can be described as follows.

TS1: Munsell training set. It consists of the reflectances of 1269 mate Munsell chips sampled every 10 nanometers and interpolated to 1nm. They were obtained from the COLORLAB database[30].

TS2: Macbeth ColorChecker training set. It consists of a set of 24 reflectances obtained from the Macbeth ColorChecker sampled every 1nm intervals.

TS3: NE reflectances training set. It consists of a set of 219 Northern European natural reflectances[33] sampled in 1nm intervals.

TS4: Chlorophyll training set: it consists of reflectance samples obtained from scene 2 of the dataset (see Figure 4 for a thumbnail set showing all the scenes used in this analysis). The samples were obtained by probing the hyperspectral scenes, extracting the spectral reflectance of every other pixel in both dimensions (341,030 samples in total). 10 nm sampling was interpolated to 1 nm.

TS5: Urban reflectances training set. It consists of samples obtained from both, scene 6 and scene 7 (see Figure 4), extracting the spectral reflectance of surfaces every three pixels in both dimensions (302,911 samples).

The different polynomial values were tested on the dataset by calculating the corresponding LMS values using both, the camera-sensors-polynomial transformation and a simulation of the LMS cones (see Figure 3). The first transformation was the "test" and the second was used as "ground truth" for this comparative analysis. The resulting pairs of LMS images were scaled to a maximum value of 1 and the "test" was subtracted from the "ground truth". The difference between the images gives us an idea of the error (*err*) arising from the calibration method. TS1, TS2 and TS2 were tested on the complete dataset of 8 images. TS4 was tested on scenes 1, 3, 4, 5, 6, 7 and 8, skipping the scene which generated the learning set (scene 2). TS5 was tested in scenes 1, 2, 3, 4 and 5 (the ones containing a higher proportion of natural objects, which were not used in the training).

Table **2** and Figure 5 show a summary of our results for all training sets and tests scenes. The histograms show how the errors are distributed in terms of numbers of pixels. All histograms show peaks close to 0, which corresponds to their small mean errors (shown in the plots). Surprisingly, the errors produced by one of the "generalist" methods (the polynomial generated by TS1 -Munsell chips) are among the smallest.

The next "generalist method", the polynomial derived from TS2 (Macbeth ColorChecker samples) is the worst case, both in terms of mean error and absolute maximum error. This is not surprising giving the small number of samples. The polynomial generated by TS3 (NE reflectances database) has not improved on the results of TS1, which might be due to its bias towards saturated colors, which are relatively unusual in nature. The best results were produced by the polynomial generated by TS4, applied to the rest of the scenes. The sampling generated by the pixels of scene 2 (predominantly chlorophyll) might be a more representative model of the reflectances encountered in nature. The polynomial generated by TS5 (based on urban surface reflectances) was also not bad in terms of mean errors, coming on top of those based on NE reflectances and Macbeth ColorChecker samples.
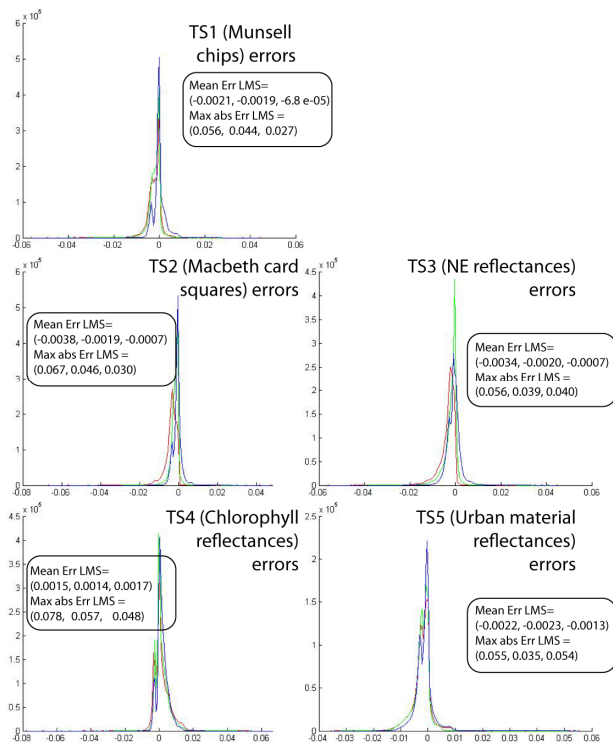

Figure 4: *Thumbnail representations of the hyperspectral reflectance scenes used to test the camera RGB to LMS characterization in this work. Each scene consists of 1018 x 1339 pixels sampled along the visible spectra in 10 nm steps (33 planes per image).More information can be obtained from http://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_04.html*

The distribution of errors in terms of chromaticity was also explored to see if there was any bias towards specific colors. Figure 6 shows the distribution of errors in the CIE 1931 *x,y* chromaticity diagram. Although the gray-levels in Figure 6 represent mismatches between the chromatic channels in LMS cone space, we found more convenient to visualize the position of these errors in a more familiar chromaticity space such as the CIE 1931. Given the disparity between the maximum values and the mean ones, and the presence of negatives, we plot the logarithm of the absolute error (log(|err|)) in Figure 6. All plots show a

Training set	errors	L plane	M plane	S plane
TS1	mean err	-0.0021	-0.0019	-0.00007
	max(|err|)	0.056	0.044	0.027
TS2	mean err	-0.0038	-0.0019	-0.0008
	max(|err|)	0.067	0.046	0.030
TS3	mean err	-0.0034	-0.0020	-0.0007
	max(|err|)	0.056	0.039	0.040
TS4	mean err	0.0015	0.0014	0.0017
	max(|err|)	0.078	0.057	0.048
TS5	mean err	-0.0022	-0.0023	-0.0013
	max(|err|)	0.055	0.035	0.054

Table 2: *Summary of the average errors measured for each L,M, and S plane for each of the training sets considered.*

concentration of errors along the yellow-greenish frontier. This can be explained by the presence of large shadowy areas among the leaves, which are prone to noise. The model that produces the most even distribution of errors (small errors = darker gray areas in the figures) is the one based on TS1 (the Munsell chips training set). The other remarkable model (it produces a large "peak" of errors mostly in the reddish-pink part of the diagram) is the one based on TS5 (urban scenes). This is to be expected, since this model should not be particularly well suited for the general "naturalistic" scenery.

Another measure of the characterization error is shown in Table 3. The columns show, for each LMS plane considered, the relative error, i.e. the averaged difference between the values obtained by subtracting the *test* to the *ground truth* solution (the err), divided by the largest of the two.
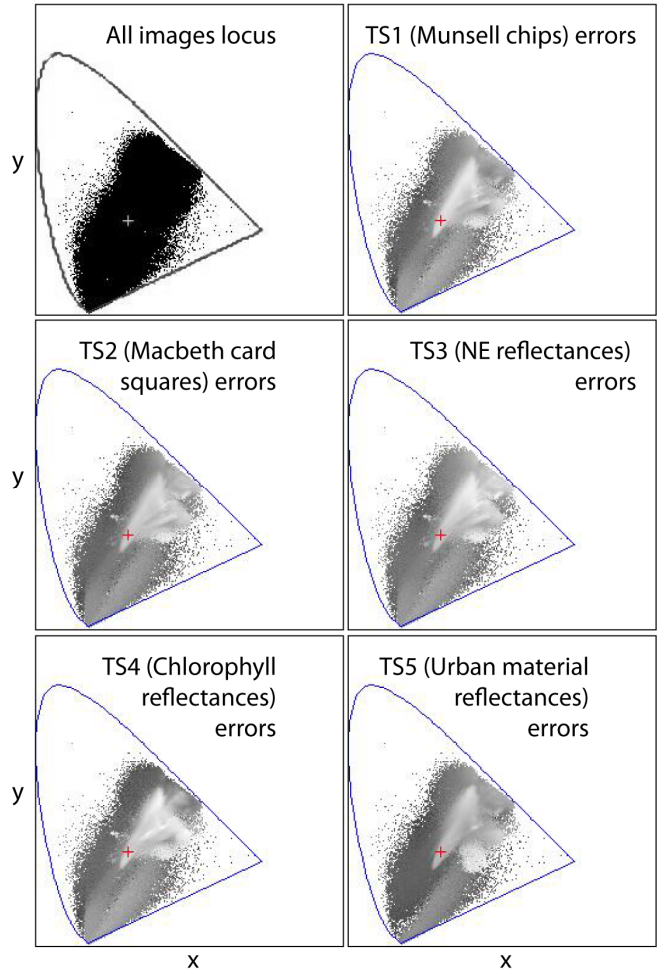

Figure 5: histograms representing the frequency of errors in terms of the numbers of pixels per error interval per L, M or S sensor. The plot also shows for each training set (TS1 to TS5) the mean error, and the absolute maximum error.

Training set	L plane	M plane	S plane
TS1	-0.063	-0.065	-0.047
TS2	-0.077	-0.065	-0.086
TS3	-0.067	-0.057	-0.123
TS4	-0.034	-0.029	0.024
TS5	-0.031	-0.041	-0.097

Table 3: mean relative errors for each of the three (LMS) planes considered. The errors were calculated by dividing each pixel difference (err) in the largest pixel value of the pair, for each (x,y) position in each of the 3 planes. The predominance of negative values (systematic error) may reflect a slight bias in the camera sensors' dependency with radiant power (gamma-correction).

The values obtained in Table 3 are larger than those of Table **2** and this is a natural consequence of computing values that contain image noise in our calculations. Noise present in the images becomes more relevant for lower intensity values such as those of shaded areas, and will increase the estimated averages. However, it is interesting to compare this measure for the different training sets considered. Again the best polynomial is that determined by TS4 (Chlorophyll-rich samples) supporting the idea that a significant improvement can be obtained by "customizing" the dataset to the expected content of the target scenes.

Figure 6: The first plot corresponds to the locus of all colors contained in the 8 images of Figure 4. All other plots show the mean errors (gray-levels) as a function of chromaticity (in the CIE 1931 diagram) for each of the 5 Training sets considered. For ease of representation given the disparity between the largest values and the smallest ones, all plots show the logarithm of the absolute value of the error.

Discussion

Our tests have shown the feasibility of a "customized" characterization of our trichromatic camera, taking on board the predicted features of the scenery to be photographed. There are many datasets of reflectances that could serve particularly well as

"training sets" for this characterization. These include the dataset tested here and other hyperspectral scenery both urban and "natural" available from the Foster-Nascimento databases[32,34]. The decision on whether to choose a training dataset or another (e.g. "generalistic" or "foliage-based") should be based on the percentage of pixels that are likely to be present in each category and its details will be analyzed in the future. However, it is indeed surprising the robustness of the measures obtained by applying the Munsell-based polynomial (TS1) to various imagery. This has been the default settings for our LMS cone sensitivity images to date. The results of these transformations are available online[13]. It might be interesting to explore several other ways to improve the overall accuracy of the characterization in the future, e.g. by trying a different polynomial mapping algorithm along with ways of converting images generated by one optimization choice into another. These issues will be explored in the future.

References

1.   W. S. Geisler, "Visual perception and the statistical properties of natural scenes", Annu Rev Psychol, **59**: 167, (2008).

2.   F. Martínez-Verdú, J. Pujol, & P. Capilla, "Calculation of the color matching functions of digital cameras from their complete spectral sensitivities", Journal of Imaging Science and Technology **46**: 15, (2002).

3.   S. Tominaga, "Multichannel vision system for estimating surface and illumination functions", J. Opt. Soc. Am. A, **13**: 2163, (1996).

4.   F. Martínez-Verdú, J. Pujol, & P. Capilla, "Characterization of a digital camera as an absolute tristimulus colorimeter", Journal of Imaging Science and Technology **47**: 279, (2003).

5.   M. J. Vhrel, & H. J. Trussell, "Color device calibration: A mathematical formulation", IEEE - Transactions on Image Processing, **8**: 1796, (1999).

6.   H. J. Trussell, "Application of set theoretic methods to color systems", Color Research and Application, **16**: 31, (1991).

7.   W. K. Pratt, & C. E. Mancill, "Spectral estimation techniques for the spectral calibration of a color image scanner", Appl Optics, **15**: 73, (1976).

8.   P. L. Vora, & H. J. Trussell, "Mathematical methods for the analysis of color scanning filters", IEEE - Transactions on Image Processing, **6**: 321, (1997).

9.   V. Cheung, S. Westland, D. Connah, & C. Ripamonti, "A comparative study of the characterisation of colour cameras by means of neural networks and polynomial transforms", Coloration Technology, **120**: 19, (2004).

10.   T. Johnson, "Methods for characterising colour printers", Displays, **16**: 193, (1996).

11.   T. Johnson, "Methods for characterising colour scanners and digital cameras", Displays, **16**, (1996).

12.   L. T. Maloney, & B. A. Wandell, "Color constancy: A method for recovering surface spectral reflectance", J. Opt. Soc. Am. A, **3**: 29, (1986).

13.   C. A. Parraga, P. G. Lovell, T. Troscianko, D. J. Tolhurst, R. Baldrich, & M. Vanrell (2009). The barcelona calibrated images database, from http://www.cvc.uab.es/color_calibration/Index.html

14.   J. Guild, "The colorimetric properties of the spectrum", Philosophical Transactions of the Royal Society of London, **230**: 149, (1931).

15.   W. D. Wright, "A re-determination of the trichromatic coefficients of the spectral colours", Transactions of the Optical Society, **30**: 141, (1928-1929).

16.   D. I. A. MacLeod, & R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance", J. Opt. Soc. Am., **69**: 1183, (1979).

17.   A. M. Derrington, J. Krauskopf, & P. Lennie, "Chromatic mechanisms in lateral geniculate-nucleus of macaque", J Physiol-London, **357**: 241, (1984).

18.   V. C. Smith, & J. Pokorny, "Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm", Vision Research, **15**: 161, (1975).

19.   J. J. Vos, "Colorimetric and photometric properties of a 2-deg fundamental observer", Color Research and Application, **3**: 125, (1978).

20.   A. Stockman, & L. T. Sharpe, "The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype", Vision Research, **40**: 1711, (2000).

21.   R. T. D. Luther, "Aus dem gebiet der farbreizmetrik", Zeitschrift für technische Physik, **8**: 540, (1927).

22.   H. E. Ives, "The transformation of color-mixture equations from one system to another", J. Franklin Inst., **16**: 673, (1915).

23.   P. C. Hung, *Colorimetric calibration for scanners and media*. Camera and Input Scanner Systems, San Jose, CA, USA. (SPIE Publishing 1991) 164

24.   P. C. Hung, "Colorimetric calibration in electronic imaging devices using a look-up-table model and interpolations", Journal of Electronic Imaging **2**, (1993).

25.   G. D. Finlayson, & M. S. Drew, "Constrained least-squares regression in color spaces", Journal of Electronic Imaging **6**: 484, (1997).

26.   G. W. Chang, & Y. C. Chen, "Colorimetric modeling for vision systems", Journal of Electronic Imaging, **9**: 432, (2000).

27.   G. Hong, M. R. Luo, & P. A. Rhodes, "A study of digital camera colorimetric characterization based on polynomial modeling", Color Research and Application, **26**: 76, (2001).

28.   P. Morovic, & G. D. Finlayson, "Metamer-set-based approach to estimating surface reflectance from camera rgb", J Opt Soc Am A Opt Image Sci Vis, **23**: 1814, (2006).

29.   R. S. Berns, & M. J. Shyu, "Colorimetric characterization of a desktop drum scanner using a spectral model", Journal of Electronic Imaging **4**: 360, (1995).

30.   J. Malo, & M. J. Luque (2002). Colorlab: A color processing toolbox for matlab. http://www.uv.es/vista/vistavalencia/software.html

31.   R. F. Lyon, & P. M. Hubel, *Eyeing the camera: Into the next century*. 10th Color Imaging Conference: Color Science, Systems

and Applications., Scottsdale, AZ. (IS&T - The Society for Imaging Science and Technology 2002) 349-355.

32. D. H. Foster, S. M. C. Nascimento, & K. Amano, "Information limits on neural identification of colored surfaces in natural scenes", Vis. Neurosci., **21**: 331, (2004).

33. J. Parkkinen, T. Jaaskelainen, & M. Kuittinen, *Spectral representation of color images*. IEEE 9th International Conference on Pattern Recognition, Rome, Italy. 1988) 933-935.

34. S. M. C. Nascimento, F. P. Ferreira, & D. H. Foster, "Statistics of spatial cone-excitation ratios in natural scenes", J. Opt. Soc. Am. A, **19**: 1484, (2002).