From [R,G,B] to Surface Reflectance: Computing Color Constant Descriptors in Images

Ron Gershon, Allan D. Jepson*, and John K. Tsotsos*

Department of Computer Science
University of Toronto
Toronto, CANADA M5S 1A4

Abstract

This paper addresses the issue of color constancy, which is the the perceptual ability of the human visual system to assign the same colors to objects under different lighting conditions. We propose a method, based on finite-dimensional linear models of reflectance and illumination, which allows the transformation [R,G,B] images into color constant images. In contrast to previous work, we show that good results can be obtained using a 3-receptor system and some knowledge about the spectral properties of natural materials and illuminants. In the method developed, an estimate of illuminant in the scene is computed, which allows the computation of color constant descriptors of the pixel values in the image. In addition, we show a method of computing the actual reflectances of the materials in the scene out of the computed color descriptors.

1 Introduction

One of the most interesting perceptual abilities of the human visual system is to assign the same colors to objects under different lighting conditions. In other words, a human observer looking at a certain scene perceives the colors of surfaces in a consistent way, although the spectral distribution of the illuminant may vary considerably. This degree of independence of perceived object color on the illuminating spectrum is called *color constancy.*

In this paper, we describe an algorithm for generating color constant descriptors of surfaces. The idea is to transform an input image representing the intensities reflected from objects into a color constant image which is a representation of reflectances, therefore independent of illumination effects. When put formally, the problem is to find a conversion matrix A that transforms the original [R,G,B] values of the image into color constant descriptors *6.* If the original [R,G,B] values of the image are defined as *PR, PG,* and *PB,* denoted as $\vec{P}$, then the following relation can be formulated:

$$A \cdot \vec{C} = \vec{P} \qquad (i)$$

The main goal of the work described in this paper is to find the values of the vector $\vec{C}$, which for each pixel, is a set of (three) numbers representing the color descriptors of the material, regardless of the illumination.

The approach we take is similar to the one forwarded by Buchsbaum [1]. Thus, it is a two-step process: estimating the illuminant, and using the estimated illuminant to obtain the descriptors. The method makes use of a finite-dimensional linear model which represents light sources and reflectances. Using this model, we show how an estimate to the illuminant is obtained, and consequently, color constant descriptors. In contrast with Buchsbaum's work, we base our choice of basis functions on statistical measurements of naturally occurring reflectances and illuminants. In addition, our approach shows that a 3-sensor input (red, green, blue) is sufficient for the estimation of the illuminant and reflectances in the scene, in contrast with the method suggested by Maloney [6].

* Also, Canadian Institute for Advanced Research.

2 Finite-dimensional Linear Models

The model which will be used for the description of surface spectral reflectances and illumination is a finite-dimensional linear model [1,6]. The idea behind this model is to describe the surface reflectances and light sources through a weighted sum of a fixed set of basis functions. The reflectances and illuminants generated by the model must be physically realizable, otherwise they will not represent real-life phenomena. The basis functions need not be physically realizable, and their only constraint is that they be linearly independent. Since the quantity being measured is the reflected intensities, this implies that the physical interaction of light and surface reflectance has to be captured by sensor(s) with some spectral sensitivity curves. Thus, a third component to be included in the model is the response functions of the sensor in use, which introduces a third set of linearly independent functions.

We now turn to the choice of the appropriate basis functions. The easiest one is the sensitivity functions. As mentioned above, if we are to deal with the human visual system, then such curves have already been determined in many different ways. If the sensor is other than the eye, such as optical scanners for digitized images, then one can easily find out the specifications of the filters used in the process of capturing the scene onto a device. Whatever sensor is being used, these sensitivity curves will be denoted as $S_i(\lambda)$ where $i = 1,2,3$ (corresponding to red, green, and blue sensitivity curves) and $\lambda \in [380, 770]$.

For the purpose of finding a finite-dimensional linear model of surface reflectance, one has to measure a large number of spectral reflectances of materials and derive the appropriate set of functions. One such study was conducted by Cohen [2], who computed the characteristic vectors of 150 Munsell chips randomly selected from a total of 433 chips. In the analysis it turned out the the first three vectors accounted for 99% of the variance in the fit to the data. We will therefore use these three vectors, which will be denoted as $R_j(\lambda)$ where $j = 1,2,3$ and $\lambda \in [380, 770]$.

Similarly, it is possible to determine a basis set which will describe typical daylight conditions. In a study carried out by Judd, Mac Adam, and Wyszecki [4], spectral distributions of 622 samples of daylight have been subjected to characteristic vector analysis. The study showed that the mean and two characteristic vectors accounted for most of the variance in the data leading them to suggest a linear model of light with three degrees of freedom. Therefore our choice of three basis functions describing the luminance will be the mean and first two characteristic vectors, which will be denoted as $I_l(\lambda)$ where $l = 1,2,3$ and $\lambda \in [380, 770]$.

To combine the basis functions into the linear model, we take the product of the three functions. This is due to the fact that reflection is a product of illumination and reflectance, and to measure the reflection one needs sensitivity curves to filter the reflected intensities. Therefore a three-dimensional tensor $T_{i,l,j}$ is defined as:

$$T_{i,l,j} = \sum_{\lambda=380}^{770} S_i(\lambda) I_l(\lambda) R_j(\lambda) \qquad i,l,j = 1,2,3 \qquad (2)$$

It is important to note that the tensor is a fixed system of constants, and does not depend at all on the materials being viewed.

3 Estimating the Illuminant

Buchsbaum [1] proposed a partial solution for the recovery of the illuminant: he assumed that the average reflected intensities (a $3 \times 3 \times 3$ system along the lines of $T_{i,l,j}$) corresponds to the one obtained from the actual illuminance acting on a standard homogeneous field, internally embedded in the model. This standard internal reflectance was assumed to approximate the entire field average reflectance. There are two assumptions Buchsbaum made which weaken the solution considerably: (1) the tensor $T_{i,l,j}$ was picked in an *ad hoc* manner, without any relevance to naturally-occurring reflectances and illuminants; (2) a fixed internal reflectance vector for the overall actual field average was built into the system. This vector was chosen to be equal reflectance at all wavelengths, implying that the overall mean pixel values of all images are grey. Moreover, there is no indication to whether this vector should be updated if the scene is known to be very different from normal rich scenes (e.g., a forest full of green colors).

The way we suggest to estimate the illuminant is to match a system representing a *general* model of material reflectances with an average of reflected intensities exhibited by the materials within a given image. This poses two problems which we address in the following sections:

1. What can be considered as a "general" set of materials which includes, with high probability, most of the materials found in scenes?
2. How should the image average reflectance be generated such that it will exhibit dependency on the illuminant and represent all materials equally?

3.1 "Ideal" Material Space

In order to be able to treat most of the naturally occurring materials, one has to find a large ensemble of materials and record their spectral properties. One such study was conducted by Krinov [5], who measured and documented the spectral sensitivity curves of 370 materials of various types, such as soil, vegetation, and water. This sample is "ideal" since the more materials it contains, the closer its average will be to the average of the sample being examined (the image). As mentioned in Section 2, the characteristic vectors suggested by Cohen [2] represent a good approximation of reflectances. Therefore we suggest to represent this ideal material space with the three characteristic vectors, and to compute the average of this ideal space in terms of the characteristic vectors. Thus, if $K_k(\lambda)$ is the k-th material in Krinov's sample set, then its representation using Cohen's basis functions is:

$$c_j[K_k(\lambda)] = Q^{-1} \sum_{\lambda=380}^{770} R_j(\lambda) K(\lambda) \qquad j = 1,2,3 \qquad (3)$$

where Q is a 3×3 matrix defined as $Q_{a,b} = \sum_{\lambda=380}^{770} R_a(\lambda) R_b(\lambda)$. Thus the average of all materials 0 is a vector computed as:

$$U_j = \frac{1}{370} \sum_{k=1}^{370} c_j[K_k(\lambda)] \qquad j = 1,2,3 \qquad (4)$$

yielding a triplet of numbers which are the average of the ideal material space as represented by three characteristic vectors.

3.2 Average Image Reflectance

There are two main requirements being imposed on the average image reflectance. The first requirement is that it should depend on the illumination and its effects on the scene. The motivation behind this requirement is that different light sources will affect the scene differently, resulting in different reflected intensities. It is therefore desirable to generate an average which will be sensitive to these changes and represent each light source differently. The second requirement is that all materials should be represented equally, no matter how large a region they occupy in the image. This is due to the fact that in the ideal material space presented in the previous section, each material is represented only once. Since our technique tries to find correspondence between the ideal space and the image space, it is desirable that the image space be sampled in the same manner.

To satisfy both requirements, we chose a method which segments the image into regions which differ from each other in their chromatic appearance. The averaging of the image reflectance is computed by segmenting the image into regions which differ from each other according to some criterion of homogeneity (set by the segmentation scheme). Each meaningful region is then averaged, and this average is added up to the total average of the image reflectance. The total average is computed for all three dimensions (i.e., [R,G,B]) and will be denoted as $\vec{V}$. Details regarding the segmentation are given in [3}.

4 The Algorithm

We now turn to describe the actual algorithm for computing color descriptors in detail. As mentioned before, the proposed technique consists of two steps: estimating the illuminant vector e, and finding the transformation matrix *A,* which will enable the conversion from [R,G,B]-space to color constant space, which will be referred to as C-space. The two steps will be discussed in detail below.

4.1 Step 1 — Computation of $\vec{\epsilon}$

1. Compute the average image reflectance $\vec{V}$.
   (a) Segment the image into regions using some homogeneity criterion.
   (b) For each region, compute the average [R,G,B] values and add to the average reflectance $\vec{V}$.
2. Compute $\vec{\epsilon}$.
   (a) Define the 3 x 3 matrix $B(\vec{U})$ to be a representation of the reflectances in the image space in terms of

the linear model $T_{i,l,j}$. Thus each element of B is defined as:

$$b_{i,l} = \sum_{j=1}^{3} T_{i,l,j}\, U_j \quad . \qquad (5)$$

(b) Find the correspondence between the average image reflectance ($\vec{V}$) and the average of the ideal material space ($\vec{U}$). The underlying assumption is that there are enough materials in the image to justify such a correspondence. Formally:

$$B(\vec{U}) \cdot \vec{\epsilon} = \vec{V} \quad . \qquad (6)$$

The estimated illuminant is therefore:

$$\vec{\epsilon} = B^{-1}(\vec{U}) \cdot \vec{V} \quad . \qquad (7)$$

4.2 Step 2 — Computation of A

1. Define the 3×3 matrix A to be the fundamental linear model $T_{i,l,j}$ as updated by illumination vector $\vec{\epsilon}$. Thus A has the form:

$$a_{i,j} = \sum_{l=1}^{3} T_{i,l,j}\, \epsilon_l \quad . \qquad (8)$$

2. To obtain the specific C-values of the image, reorganize Equation 1 into:

$$\vec{C} = A^{-1} \cdot \vec{P} \quad . \qquad (9)$$

The output of this step is an "image" of C-values; that is, each pixel value in the image is transformed from [R,G,B] to a new set of three numbers: C_1, C_2, and C_3.

4.3 Discussion

One of the assumptions made, used in Step 1 of the algorithm, is that the image reflectance space has enough materials to allow the transformation to the ideal material space. The motivation behind this assumption was the fact that the ideal material space contains a rich ensemble of materials, and thus its average is likely to be similar to averages of many scenes. The implication of this assumption is that the accuracy of the algorithm depends on the richness in materials of the input image. Clearly, the situation can be simplified if there exists *a priori* knowledge about the domain from which the images are taken, since the ideal material space can be reduced to include materials which are expected to be in the input images. For example, if it is known in advance that forests will be the subject of the images, then only materials which may appear in forests, such as leaves and wood, should be included in the ideal material space, thus increasing the accuracy of the algorithm.

The C-values computed by the algorithm provide a means of recovering the actual reflectance of the material, as represented by the reflectance basis functions. Since the C-values are color descriptors which are constant regardless of the illuminant, and since they were computed using the linear model, it is a straightforward computation that allows the recovery of the reflectances. If we define the recovered reflectance as $RR(\lambda)$, then this computation is simply:

$$RR(\lambda) = \sum_{j=1}^{3} R_j(\lambda)\, C_j \qquad (10)$$

5 Experiments

In order to evaluate the effectiveness of the algorithm, a few experiments were designed and carried out on simulated and real images. Although each experiment was designed with a different goal in mind, the overall goal was to demonstrate the robustness and accuracy of the algorithm.

5.1 Accuracy as a Function of Knowledge About Materials

The first goal was to evaluate the accuracy of the algorithm and the claim made earlier that *a priori* knowledge about the materials being viewed increases the accuracy of the results. Therefore the experiment was intended to compare the real C-values of materials to the ones output by the algorithm given different illuminants. The real C-values were computed in the following way. Seven materials were chosen at random from the data collected by Krinov [5] and their real C-values were computed according to Equation 3. In addition, knowing that only the seven materials chosen were going to be used in the experiment, the ideal space was reduced to include only those materials, thus resulting in an updated set of averages of this space.

As pointed out, the real C-values were to be compared to the C-values computed by the algorithm. The input to the algorithm was computed in the following way. The materials were combined into a Mondrian. Since the only information about the materials was their spectral reflectance curves, there was a need to create an [R,G,B] image out of these reflectances, using some random illuminants ϵ'_l and assuming some [R,G,B] filters. Therefore the [R,G,B] values of the patches were computed as:

$$P_i = \sum_{l=1}^{3} \sum_{\lambda=400}^{650} S_i(\lambda)\, K(\lambda)\, I_l(\lambda)\, \epsilon'_l \qquad i = 1,2,3 \qquad (11)$$

where $S_i(\lambda)$ represents the sensitivity curves of the filters, and $I_l(\lambda)$ represents the characteristic vectors describing light sources. The [R,G,B] image was used as input to the algorithm. As a result of Step 1, an estimate of the illuminant was computed. The comparison between the illuminant we used in the experiment ϵ'_l and the estimated illuminant as computed by the algorithm ϵ_l are depicted in Figure 1. In Step 2 of the algorithm, the

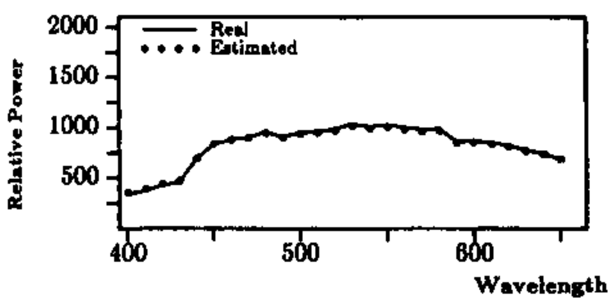

Figure 1: The real and estimated illuminant.

C-values were computed for each material. The absolute error in the C-values compared to the true C-values are given in Table 1. As can be observed from the figure, given the *a priori* knowledge about the materials, the results of the algorithm are accurate.

5.2 Constancy Under Different Illuminants

In this experiment the goal was to verify that the C-values generated by the algorithm were indeed constant under different illumination conditions. Therefore a simple method was used: the same scene was photographed and digitized under two different illuminants; the first was illuminated with white

Absolute error in C-values						
Material #						
1	2	3	4	5	6	7
0.183 0.141 0.010	0.325 0.629 0.260	0.030 0.216 0.314	0.378 0.732 2.247	0.449 0.864 0.757	0.195 0.869 0.306	0.000 0.401 0.055
Real C-values						
7.147 -0.664 -0.368	23.488 2.910 8.831	9.583 -0.884 2.197	25.392 -2.094 10.716	55.477 12.184 2.499	91.055 -3.985 8.057	25.947 3.473 4.512

Table 1: Comparison between the true C-values and the error of the estimated ones.

The numbers in each box, from top to bottom, represent C_1, C_2, and C3 respectively. All numbers were multiplied by a factor of 100.

incandescent light (termed as Image #1) and the second with the same light source with a yellow filter in front of it (Image #2).[1] For each image, the algorithm was run and C-values of the same region were recorded. The results appear in the upper portion of Figure 2. We also plotted the recovered illuminants as computed by the algorithm (Figure 2 — center portion), and the estimated reflectances in each area in both images (Figure 2 — bottom portion). It is clear from the table and the plots that the C-values generated are very close, as are the estimated reflectances. In addition, the estimated illuminants do exhibit curves which are close to approximating white and yellow illuminants, respectively. Note that the estimates of the reflectances, as well as the illuminants, are not accurate due to the fact that the sample of "materials" is not rich enough (there is only one material!). Nevertheless, the color descriptors are fairly constant, indicating the robustness of this technique.

6 Summary

This paper dealt with the problem of color constancy, or the fact that the perceived color of surfaces tends to remain constant despite changes in illumination that alter the intensities reflected off the surfaces. We have presented an algorithm which generates color constant descriptors for different surfaces. There are two steps in the algorithm: the first estimates the illuminant present in the scene, and the second step makes use of the estimated illuminant to generate the color descriptors. We have shown that by computing an image average reflectance, the algorithm can estimate the illuminant present in the scene and with it, produce color descriptors for each of the surfaces in the image. Although the same materials may exhibit different reflected intensities under different illumination conditions, the algorithm produces the same descriptors in each case. The main result of the algorithm is that it allows the treatment of surfaces independent of the spectral power of the light source affecting them.

References

[1] Buchsbaum, G. (1980) "A spatial processor model for object colour perception", *J. Franklin Institute,* 310(1), pp. 1-26.

[1] These images were provided by Professor Steve Shafer of the Calibrated Imaging Laboratory at Carnegie Mellon University, supported by the National Science Foundation and the Defense Advanced Research Projects Agency.

Coordinates		Image # 1	Image # 2
row	column	(white illuminant)	(yellow illuminant)
180	115	(3.855, 1.398, 1.074)	(4.150, 1.303, 0.792)
180	135	(3.736, 1.444, 1.119)	(3.961, 1.524, 0.856)
200	115	(3.872, 1.369, 1.124)	(3.996, 1.151, 0.831)
200	135	(3.724, 1.473, 1.070)	(3.760, 1.531, 0.916)
195	125	(3.831, 1.373, 1.080)	(4.085, 1.130, 0.949)
185	120	(3.831, 1.383, 0.850)	(3.819, 1.332, 0.907)
190	130	(3.902, 1.343, 1.080)	(4.192, 1.196, 0.702)

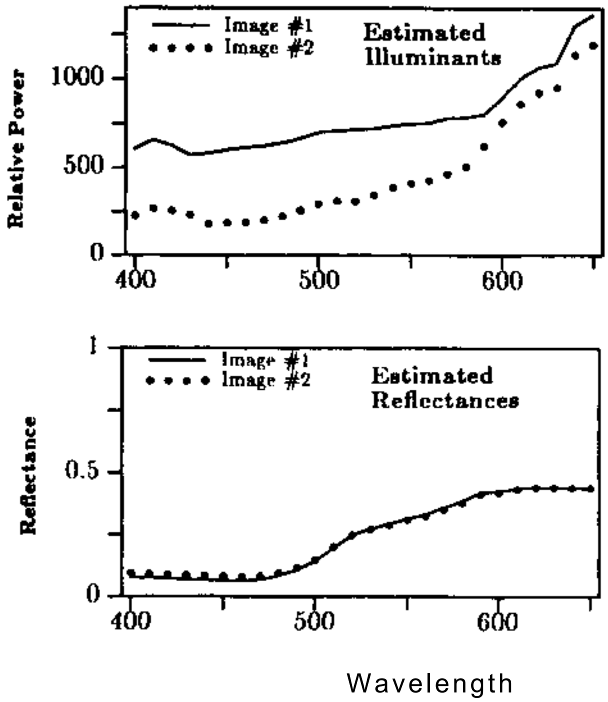

Figure 2: Results of the algorithm run on two different images.

[2] Cohen, J. (1964) "Dependency of the spectral reflectance curves of the Munsell color chips", *Psychonomic Sci.,* 1, pp. 369-370.

[3] Gershon, R. (1987) *The Use of Color in Computational Vision,* Ph.D. Thesis, Dept. of Computer Science — University of Toronto, forthcoming.

[4] Judd, D.B., MacAdam, D.C., and Wyszecki, G. (1964) "Spectral distribution of typical daylight as a function of correlated color temperature", *J. Opt. Soc. Am.,* 54, pp. 1031-1040.

[5] Krinov, E.L. (1947) "Spectral reflectance properties of natural formations", Technical Translation TT-439, National Research Council of Canada.

[6] Maloney, L.T. (1985) "Computational approaches to color constancy", Technical Report 1985-01, Dept. of Psychology — Stanford University.