Active Bayesian Color Constancy with Non-Uniform Sensors

Sandra Skaff, Tal Arbel, James J. Clark
Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada
E-mail: {sandra,arbel,clark}@cim.mcgill.ca

Abstract

The human impression of the color of an object is the same when viewed foveally or peripherally, despite the non-uniformity of the spatial distribution of photoreceptors on the retina. We propose that this non-uniformity can be used to attain color constancy, the perception of a constant surface color under varying illumination. We develop a multi-sensor Bayesian technique that solves for the surface reflectance and lighting parameters of a bilinear model by sequentially acquiring measurements from independent sensors. We present two cases: (i) two sets of sensors, each with different spectral sensitivies, (ii) a continuous variation in the spectral sensitivities across the sensor array.

1. Introduction

A striking characteristics of the human retina is its non-homogeneity. It is well known that the spatial distribution of photoreceptors on the retina is non-uniform. Less well-known is the fact that the retina has a marked non-homogeneity in the spectral sensitivity of its photoreceptors. Part of this non-homogeneity arises from the fact that the macular pigment, which is a yellowish jelly that covers the macula, absorbs up to 50% of the light in the short wavelength range [1], causing a significant shift in the color sensitivity of foveal receptors. In spite of these non-homogeneities, the human perception of color is remarkably invariant to eye position. We have the subjective impression that the color of an item in the world is the same when viewed peripherally as when viewed foveally. Clark and O'Regan [3] proposed that the human visual system uses a stable-world assumption in developing position invariant color perception. In their approach, it is precisely the changing retinal signal resulting from moving the eye when viewing a stable colored surface patch that determiness the perceived color.

We propose that *color constancy*, the perception of a constant color across variations in illumination, can be explained by a process similar to the Clark-O'Regan explanation of *color stability* across eye movements. The approach that we take follows the Bayesian approach of Brainard and Freeman [2]. Our method extends their technique to the use of a large set of independent sensors (corresponding to the photoreceptors at different retinal locations) from which measurements are obtained sequentially (corresponding to moving the gaze across a surface patch). We present two cases, one in which there are only two sets of sensors, each having a different spectral sensitivity characteristic, and one in which there is a smooth variation in the spectral sensitivity across the sensor array.

2. The Bayesian Approach

Our color constancy algorithm builds on the Bayesian technique of Brainard and Freeman [2]. Their algorithm uses the bilinear model of Maloney and Wandell [7] to provide a parametrization of the surface and illuminant spectra. The bilinear model can be summarized as follows. Describe the light arriving at a location x on an array of sensors by the function $E(\lambda)\ S^x(\lambda)$, where $E(\lambda)$ is the spectral power distribution of the ambient light in the scene, while $S^x(\lambda)$ is the surface spectral reflectance. Let us assume that we have p classes of sensors at each location x, each with a different relative wavelength sensitivity, $R_k(\lambda)$. Then the sensor responses are given by

$$\rho_k^x = \int E(\lambda)\ S^x(\lambda)\ R_k(\lambda)\ d\lambda, \quad k = 1, 2, \ldots, p. \quad (1)$$

S and E can each be represented as linear models:

$$S^x = \sum_{j=1}^{n} \sigma_j^x\ S_j(\lambda), \qquad E(\lambda) = \sum_{i=1}^{m} \epsilon_i\ E_i(\lambda). \quad (2)$$

1051-4651/02 $17.00 © 2002 IEEE

In both cases, the basis functions are fixed. Therefore, finding the surface reflectance and ambient light involves recovering the basis function weights σ_j^x and ϵ_i. The sensor responses are seen to be bilinear functions of the unknown basis function weights. This bilinearity implies that the problem of finding the weights is ill-posed, as different choices of ϵ and σ_x can produce same the sensor measurements.

Brainard and Freeman used a Bayesian technique to regularize the problem of computing the values of the illuminant and surface parameters as follows. Let the vector of surface and illumination model weights be denoted as x, and the sensor responses by y. We can obtain a statistical model for x by the conditional posterior density function, $p(x|y)$, of x given the measurement:

$$p(x|y) = \frac{p(y|x)\ p(x)}{p(y)} \approx p(y|x)\ p(x). \qquad (3)$$

$p(y|x)$ is the *likelihood* which models the relation between the illuminant spectrum model, the surface spectra models and the sensor responses. $p(x)$ represents the prior information on the model parameters. In the Brainard-Freeman formulation, the prior is estimated through principal component analysis (PCA) techniques where the principal components are built from the surface reflectances of a fixed set of Munsell color patches [4]. The likelihood $p(y|x)$ is also represented by a normal distribution.

3. The Sequential Multiple Sensor Approach

In this paper, we introduce a strategy that is an extension to the Brainard-Freeman Bayesian approach, in which only one sensor was modeled. In our approach, measurements are acquired from a large set of independent sensors, each with its own spectral sensitivities. The inherent ill-posedness of the problem is therefore addressed through the introduction of more sources of information. Measurements are acquired sequentially from each sensor, much as when a person's gaze moves across a surface in the scene. *Color stability* is achieved through the accumulation of evidence from the various sensors through a sequential Bayesian estimation process.

As in the Brainard and Freeman appraoch, we represent evidence for the lighting and surface color parameters of various surface patches in a scene by a conditional probability density function, given the sensor measurements. This probabilistic evidence is then accumulated sequentially over sensors with different spectral sensitivities through a Bayesian chaining approach. We model two cases: (i) one in which there are only two different types of sensors, and (ii) one in which there is a smooth variation in spectral sensitivity curves across the sensor array. In practice, this multi-sensor formulation can be modeled through the placement of a filter with the appropriate absorption characteristics onto a single camera lens. In this case, evidence can be accumulated spatially over the pixels of the image of the scene (see Section 4).

3.1. The Binary Formulation

We first consider the case where measurements of a scene are acquired by two sensor types, each with a different spectral sensitivity. We will assume that each individual sensor actually produces three independant measurements, which we will refer to as R, G, and B. RGB measurements of a particular surface obtained from sensors of type 1 will be denoted RGB_{1n}, and those from sensors of type 2 as RGB_{2n}, where n refers to the surface patch being viewed. Let us denote the surface spectral model weight vector by a_n, and the illuminant spectral model weights vector by b. Consider the case of three surface patches in a scene. Suppose that surfaces 1 and 3 are visible to sensor type 1 alone, while surface 2 is visible to both sensor types. Let $\{RGB\}$ denote the total set of measurements of the scene:

$$\{RGB\} \triangleq RGB_{11}, RGB_{13}, RGB_{21}, RGB_{22}, RGB_{23}.$$

The conditional posterior density function for the parameters, a_1, a_2, a_3 and b, given the set of measurements of the scene,$\{RGB\}$, is given by:

$$\begin{aligned}
&p(a_1, a_2, a_3, b|\{RGB\}) \\
&\approx p(\{RGB\}|a_1, a_2, a_3, b)\ p(a_1, a_2, a_3, b) \\
&\approx p(RGB_{11}, RGB_{13}|a_1, a_3, b) \\
&\quad p(RGB_{21}, RGB_{22}, RGB_{23}|a_1, a_2, a_3, b) \\
&\quad p(a_1, a_3, b)\ p(a_2) \\
&\approx p(a_1, a_3, b|RGB_{11}, RGB_{13})\ p(RGB_{11}|a_1, b) \\
&\quad p(RGB_{22}|a_2, b)\ p(RGB_{23}|a_3, b)\ p(a_2) \qquad (4)
\end{aligned}$$

In obtaining this formulation, various simplifying assumptions are made. First, it is assumed that the prior probabilities for each surface reflectance weight a_n are statistically independent of each other and of the spectral function weights of the illuminant, b. Next, the measurements are assumed to be conditionally statistically independent. Finally, it is assumed that there is no interreflection between surfaces, and therefore the spectral reflectance for a surface and the lighting vectors are sufficient statistics for the measurement of that

surface. This leads to a simplification:

$$p(RGB_{22}|a_1, a_2, a_3, b) = p(RGB_{22}|a_2, b).$$

From equation 4, we can conclude that the posterior for the entire scene is a function of the posterior for one sensor (sensor A), the likelihood of the RGB measurement of the other surface (sensor B) and the prior of its spectral function weights. These manipulations indicate that the active Bayesian formulation is, in fact, *recursive*, where the posterior for each sensor acts as a prior for the next sensor. Given the assumptions above, the posterior for sensor A can be obtained by:

$$\begin{aligned} & p(a_1, a_3, b|RGB_{11}, RGB_{13}) \\ \approx\ & p(RGB_{11}, RGB_{13}|a_1, a_3, b)\ p(a_1, a_3, b) \\ \approx\ & p(RGB_{11}|a_1, b)\ p(RGB_{13}|a_3, b)\ p(a_1)\ p(a_3)\ p(b) \\ \approx\ & p(a_1, b|RGB_{11})\ p(RGB_{13}|a_3, b)\ p(a_3) \end{aligned}$$

Note that this posterior is a function of the likelihood and prior for surface 3, and the posterior for surface 1. The recursive nature of the strategy can be seen once again as, for each sensor, the posterior for each surface patch acts as a prior for the next surface patch.

3.2. The Non-Uniform Formulation

Next, we consider the case where there is a smooth variation in the spectral sensitivities across the sensor array. The formulation is once again described by considering three surfaces in the scene. Surfaces 1, 2, and 3 have N, M and K sensor response functions, respectively. Thus, the sensor responses can be denoted as $\{RGB\}_{n1\ldots np} \triangleq RGB_{n1}, RGB_{n2}, \ldots, RGB_{np}$ for each surface n, where p is the number of sensor responses for that surface. Based on the assumptions made in the previous section, the posterior for the entire scene is again recursive and can be formulated as:

$$\begin{aligned} & p(a_1, a_2, a_3, b|\{RGB\}_{11\ldots1N}, \{RGB\}_{21\ldots2M}, \{RGB\}_{31\ldots3K}) \\ \approx\ & p(a_1, a_2, b|\{RGB\}_{11\ldots1N}, \{RGB\}_{21\ldots2M}) \\ & p(\{RGB\}_{31\ldots3K}|a_3, b)\ p(a_3) \end{aligned}$$

This implies that the posterior is computed by multiplying the prior and likelihood for the third surface by the posterior from the first two surfaces, i.e. these posteriors become priors for the third surface. Similarly, the posterior for the first two surfaces is simply the product of the likelihood and prior for the second surface and the posterior for the first, taking care at the transition between surfaces:

$$\begin{aligned} & p(a_1, a_2, b|\{RGB\}_{11\ldots1N}, RGB_{21}) \\ \approx\ & p(\{RGB\}_{11\ldots1N}, RGB_{21}|a_1, a_2, b)\ p(a_1, a_2, b) \\ \approx\ & p(\{RGB\}_{11\ldots1N}|a_1, a_2, b)\ p(RGB_{21}|a_1, a_2, b) \\ & p(a_1)\ p(a_2)\ p(b) \\ \approx\ & p(a_1, b|\{RGB\}_{11\ldots1N})\ p(RGB_{11}|a_2, b)\ p(a_2). \end{aligned}$$

Note that the subjective prior term (e.g. $p(a_2)$) only comes in at the start of the recursion for each surface. The posterior for each surface is simply a recursive, sequential update of the posterior given each of the sensor responses for that surface:

$$\begin{aligned} & p(a_1, b|\{RGB\}_{11\ldots1N}) \\ \approx\ & p(a_1, b|\{RGB\}_{11\ldots1N-1})\ p(RGB_{1N}|a_1, b) \end{aligned}$$

For both of the cases described here we obtain a sequential update of the posteriors given each of the sensors in turn. The result is that, for each surface patch, each new sensor response leads to an increase in information regarding the color of the surface patch and the illumination. The addition of more surface patches in the scene provides even more information regarding the illumination and thus, in turn, about the surface colors themselves.

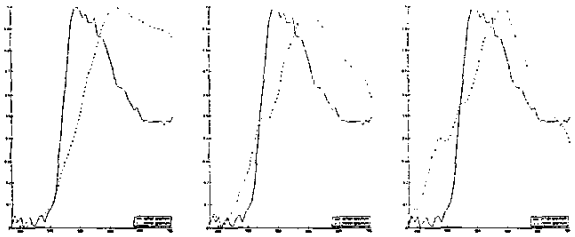
4. Simulations

To demonstrate our approach, a series of simulations were performed. We constructed "Mondrian" scenes consisting of several surface patches illuminated by a single light source. The multi-sensor environment is modeled through the simulated placement of an optical filter onto the (simulated) camera lens, which non-uniformly modifies the spectrum of the light falling on the sensors. We simulated the binary case, in which half of the filter is tinted yellow with the other half left transparent, and also simulated the non-uniform case, in which a filter with a Gaussian absorption profile is used; the center is tinted deep yellow, and the color fades smoothly towards the periphery.

The RGB components of each sensor's response were simulated by multiplying the spectrum of the light passing through the optical filter by the sensor spectral sensitivity curves denoted $R_k(\lambda)$. For the purposes of this paper, we used the Stockman and Sharpe estimates for the sensitivities of the cones in the human retina [6]. We used randomly selected Munsell color patches and random combinations of the Parkkinen and Silftsen daylight and skylight spectra [5] to generate the spectrum of the light falling on the sensor array. The likelihoods in our Bayesian formulations were computed using the model predictions of

the sensor measurement. The basis functions for the surface spectrum model were taken to be the principal components of the spectra of the 1250 Munsell color patches as measured by Parkkinen *et al.* [4]. The basis functions for the light source spectrum model were taken to be the principal components of a set of daylight and skylight spectra as measured by Parkkinen and Silfsten [5]. The likelihoods are assumed to have Gaussian distributions. The prior distributions for the spectral model parameters are assumed to be independent and Gaussian. The means and variances of the priors are computed from the distribution of weights corresponding to the 1250 Munsell spectra and the 37 different skylight and daylight spectra that were used. These weights were obtained by projecting the measured spectra onto the basis function sets.The location of the maximum of the posterior distribution was estimated by a standard MATLAB optimization package, resulting in a set of estimated surface and light spectra weight vectors.

Experiments with several surface patches in the scene (as described in Section 3) were performed, and the resulting estimates for the model and the actual spectra (Munsell, patch 300) for one surface in the scene as well as the spectra for the light source can be found in Figures 1 and 2, respectively. The figures illustrate the spectra for the two different types of optical filters as compared to the case with no filter. The results indicate that there is considerable improvement with the introduction of the multi-sensor method. In fact, the more sensors introduced, the better the estimates. This can also be seen by examining the RMS errors depicted in Table 1.

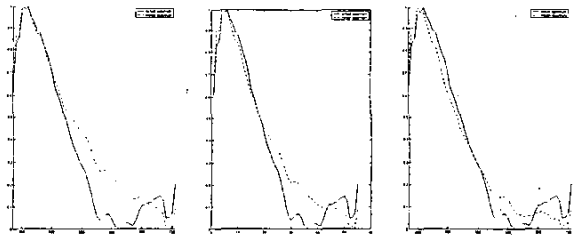


(a) (b) (c)

Figure 1. The estimated and actual (Munsell patch 300) surface spectra in the cases of (a) no filter, (b) binary filter, (c) Gaussian filter are plotted for the relevant wavelength range.

5. Summary

In this paper we have proposed an *active Bayesian* strategy that is an extension to the Brainard-Freeman Bayesian approach to color constancy. Here, evidence for the parameters of the surface and illuminant spectral models, in the form of conditional probability density functions, is sequentially accumulated over a multi-sensor environment. We develop two types of multi-sensor environments, one with a binary and one with smoothly varying spectral absorption curves.



(a) (b) (c)

Figure 2. The estimated and actual illuminant spectra in the cases of (a) no filter, (b) binary filter, (c) Gaussian filter.

Case	No Filter	Binary Filter	Gaussian Filter
Surf	0.315	0.2232	0.1931
Illum	0.1143	0.0867	0.0634

Table 1. RMS error for the surface and illuminant spectra for the 'no filter', binary and Gaussian filter cases.

References

[1] Bone, R. A., Landrum, J. T., and Cains, A., "Optical density spectra of the macular pigment in vivo and in vitro", *Vision Research*, Vol. 32(1), pp. 105-110, 1992.

[2] Brainard, D.H. and Freeman, W.T., "Bayesian color constancy", *Journal of the Optical Society of America*, 14:7, 1393-1411.

[3] Clark, J.J. and O'Regan, J.K., "A Temporal-difference learning model for perceptual stability in color vision", *2000 International Conference on Pattern Recognition*, Barcelona, Spain, pp 503-506, September 2000.

[4] Parkkinen, J. P. S., Hallikainen, J. and Jaaskelainen, T., "Characteristic spectra of Munsell colors", *Journal of the Optical Society of America*, Vol. 6, No. 2, pp. 318-322, 1989.

[5] Database of Spectra, University of Joensuu, Finland. http://cs.joensuu.fi/spectral

[6] Stockman, A., Sharpe, L. T., and Fach, C. C., "The spectral sensitivity of the human short-wavelength cones", *Vision Research*, Vol. 39, pp. 2901-2927, 1999.

[7] Maloney, L. T. and Wandell B. A., "Color constancy: a method for recovering surface spectral reflectance", *Journal of the Optical Society of America*, Vol. 3, pp. 29-33, 1985.