Color Constancy Using KL-Divergence

Charles Rosenberg, Martial Hebert, Sebastian Thrun
Department of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
chuck@cs.cmu.edu, hebert@cs.cmu.edu, thrun@cs.cmu.edu

Abstract

Color is a useful feature for machine vision tasks. However, its effectiveness is often limited by the fact that the measured pixel values in a scene are influenced by both object surface reflectance properties and incident illumination. Color constancy algorithms attempt to compute color features which are invariant of the incident illumination by estimating the parameters of the global scene illumination and factoring out its effect. A number of recently developed algorithms utilize statistical methods to estimate the maximum likelihood values of the illumination parameters. This paper details the use of KL-divergence as a means of selecting estimated illumination parameter values. We provide experimental results demonstrating the usefulness of the KL-divergence technique for accurately estimating the global illumination parameters of real world images.

1 Introduction

In machine vision and image database applications color can be used as a simple means of segmenting or identifying a specific object as described in [14] or as a means of quickly identifying likely candidate regions for object recognition as in [18]. However, problems arise when images are captured under varying illumination conditions and with cameras with differing sensor characteristics. Color constancy methods try to overcome these problems by estimating the surface reflectance properties of objects in a scene regardless of scene illumination and camera characteristics. In general, if we assume diffuse surface reflections, the measured pixels values for camera sensor channel k at image location i, denoted as $\rho_k(i)$, are the product of the incident illumination $E(\lambda)$, the surface reflectance properties $S(i,\lambda)$ and camera sensor channel spectral sensitivity $C_k(\lambda)$ as a function of the wavelength of the incident light λ, integrated over the visible spectrum, ω:

$$\rho_k(i) = \int_\omega E(\lambda)S(i,\lambda)C_k(\lambda)d\lambda$$

The goal of color constancy is to recover the original surface reflectance properties, $S(i,\lambda)$, regardless of the incident illumination $E(\lambda)$. However, since three camera sensor channels are often employed, it is only typically possible to determine the sensor response $\rho_k(i)$ under some canonical camera model, unless specific constraints are placed on the illuminants and the surface characteristics [17].

2 Model of Image Formation

2.1 Illumination Model

In this work we assume a camera with three sensor channels and utilize a simplified model of how global scene illumination affects the measured pixel responses:

$$R(i) = \alpha S_r(i);\ G(i) = \beta S_g(i);\ B(i) = \gamma S_b(i)$$

In this formulation $R(i)$, $G(i)$, and $B(i)$ are the measured camera sensor responses at each each pixel, $S_r(i)$, $S_g(i)$, and $S_b(i)$ are the responses of the camera sensors under some canonical illuminant, which we will call the "true" or canonical surface color. In this diagonal illumination model, the parameters α, β, and γ capture the effects of the global scene illumination and are constant for all of the pixels in the image. In the domain of our diagonal illumination model the accurate estimation of these parameters permits the computation of a color feature value which is *invariant* of incident global scene illumination. We can further reduce the number of parameters of interest by focusing on recovering only the "color" of the pixels and the illuminant and not their absolute intensity. This allows us to restrict ourselves to cases where $\gamma = 1$. Hence, only the values of α and β need to be estimated.

We adopt a color space in which the intensity of a pixel is inherently factored out. We chose the following log space transformation of the original measured pixel values:

$$C_1(i) = \log\left(\frac{R(i)}{B(i)}\right) = \log\left(\frac{\alpha S_r(i)}{S_b(i)}\right)$$
$$C_2(i) = \log\left(\frac{G(i)}{B(i)}\right) = \log\left(\frac{\beta S_g(i)}{S_b(i)}\right)$$

To simplify our notation, we define the following two quantities:

$$S_1(i) = \log\left(\frac{S_r(i)}{S_b(i)}\right); S_2(i) = \log\left(\frac{S_g(i)}{S_b(i)}\right)$$

We chose to work in a log space because, given our diagonal scene illumination model, an illumination change amounts to a simple shift:

$$C_1(i) = \log(\alpha) + S_1(i); C_2(i) = \log(\beta) + S_2(i)$$

Figure 1a is a graphic representation of this color space, C_1 is on the vertical axis and C_2 is on the horizontal axis, the minimum value for both axes is in the upper left corner.

2.2 Sensor Noise Model

The color space we have chosen factors out individual pixel intensities, however we do not want to completely discard this information. In terms of accurately measuring color component values, not all pixels are created equal. Sensor noise affects our ability to accurately determine the color of a pixel. In general, the measured color component values of darker pixels have higher uncertainty than lighter pixels. We introduce a sensor model which captures this effect. We model sensor measurements as Gaussians centered on the true value with constant variance noise parameters σ_r^2, σ_g^2, σ_b^2, as per [3]. Using our original illumination model the measured pixel values are modeled as follows:

$$R(i) \sim N\left(\alpha S_r(i), \sigma_r^2\right); G(i) \sim N\left(\beta S_g(i), \sigma_g^2\right);$$
$$B(i) \sim N\left(\gamma S_b(i), \sigma_b^2\right)$$

If we assume that the three color measurements are independent, the joint probability density function for a single measurement at location i becomes, $f(r,g,b) =$

$$\frac{1}{\sqrt{2\pi}^3\sigma_r\sigma_g\sigma_b}\exp\left[-\frac{\left(\frac{r-\alpha s_r}{\sigma_r}\right)^2+\left(\frac{g-\beta s_g}{\sigma_g}\right)^2+\left(\frac{b-\gamma s_b}{\sigma_b}\right)^2}{2}\right]$$

Since we will be using a transformed version of the original color space we are interested in how the noise affects measurements in that transformed space. The color space transformation we introduced in the previous section is as follows:

$$C_1(i) = \log R(i) - \log B(i)$$
$$C_2(i) = \log G(i) - \log B(i)$$
$$C_3(i) = \log B(i)$$

$C_3(i)$ is introduced in order to facilitate the transformation of the probability density function for a measurement in the original color space to this new color space. This is achieved by a standard transformation of random variables involving the Jacobian of the transformation, as per [5]. After this procedure we obtain the following expression for the joint probability density function for a single measurement in the transformed color space, $g(c_1,c_2,c_3) =$

$$\frac{rgb/(\alpha\beta\gamma)}{\sqrt{2\pi}^3\sigma_r\sigma_g\sigma_b}\exp\left[-\frac{\left(\frac{r-\alpha s_r}{\sigma_r}\right)^2+\left(\frac{g-\beta s_g}{\sigma_g}\right)^2+\left(\frac{b-\gamma s_b}{\sigma_b}\right)^2}{2}\right]$$

The effect is that for a given value of α, β, and γ, the relative probability of a pixel measurement at its mean is related to the product of r, g, and b. In this work, instead of modeling the full noise distribution, we take a simplified approach and scale the contribution of each pixel to our model probabilities by the product of its r, g, and b values.

3 Parameter Estimation

In this work we evaluate two statistical methods of estimating the global scene illumination parameters: the maximum likelihood method, similar to color by correlation described in [10], and a new method which uses KL-divergence.

3.1 Statistical Model

For both methods of parameter estimation that will be evaluated it is important to be able to calculate the probability of observing a particular color measurement given a set of illumination parameters:

$$P(C_1(i), C_2(i) \mid \alpha, \beta)$$

The practical difficulty of performing such a computation is that it would seem to be necessary to have a different probability distribution for each value of α and β. However, this can be avoided if we make the assumption that the parameters of the global scene illumination are independent of the canonical ("true") colors of the image pixels. This independence assumption seems reasonable because one would not generally expect the "true" color of an object in the scene to be affected by the color of the scene illumination. Given this assumption, it is possible to utilize a single distribution over canonical colors to compute the probability values we need. The single distribution over canonical colors which we utilize for our computation takes the following form:

$$P(S_1(i), S_2(i))$$

Note that since this distribution is strictly over the canonical colors, it is not a function of the illumination parameters α and β. For a given, fixed set of values for the illumination parameters, for example $\alpha = \widetilde{\alpha}$ and $\beta = \widetilde{\beta}$, and because of our independence assumption, we can substitute into our illumination model and obtain:

$$P(S_1(i), S_2(i)) = P(C_1(i) - \log(\widetilde{\alpha}), C_2(i) - \log(\widetilde{\beta}))$$
$$= P(C_1(i), C_2(i) \mid \alpha = \widetilde{\alpha}, \beta = \widetilde{\beta})$$

This means that we only need to to maintain a single distribution over canonical colors and can "shift" that distribution to compute the probability of observing a particular pixel measurement. To facilitate the discussion in the following sections, we adopt the following simplified notation to represent the transformed canonical color distribution:

$$P_{\alpha,\beta}(C_1(i), C_2(i)) \equiv P(C_1(i) - \log(\widetilde{\alpha}), C_2(i) - \log(\widetilde{\beta}))$$

3.2 Parameter Estimation Using Maximum Likelihood

When using the maximum likelihood technique to estimate the values of α and β we would like to find the values of the parameters which are most likely given the observed data, which in our case are the observed pixel values. To do this it is necessary to assume a generative model of image formation. In this work we will assume that the pixel values are independently and identically distributed as in [10]. This results in the following expression for the joint probability density function, where the n pixels in the image are indexed from $1 \ldots n$:

$$P(\alpha, \beta \mid \{C_1(i), C_2(i)\}\, for\, i = 1 \ldots n) = \frac{\prod_{i=1}^{n} P(C_1(i), C_2(i) | \alpha, \beta)\, P(\alpha, \beta)}{\prod_{i=1}^{n} P(C_1(i), C_2(i))}$$

As is standard procedure in a maximum likelihood formulation, if we assume a uniform prior over the illumination parameters and because we are only interested in the maximum likelihood values of these parameters and not their actual posterior probabilities, this expression can be simplified as follows:

$$P(\alpha, \beta \mid \{C_1(i), C_2(i)\}\, for\, i = 1 \ldots n) \propto \prod_{i=1}^{n} P(C_1(i), C_2(i) \mid \alpha, \beta)$$

It is customary to reformulate this likelihood function in terms of the log likelihood which is monotonic in the likelihood:

$$L(\alpha, \beta \mid \{C_1(i), C_2(i)\}\, for\, i = 1 \ldots n) = \sum_{i=1}^{n} \log P(C_1(i), C_2(i) \mid \alpha, \beta)$$

We can now substitute our transformed canonical color distribution:

$$L(\alpha, \beta \mid \{C_1(i), C_2(i)\}\, for\, i = 1 \ldots n) = \sum_{i=1}^{n} \log P_{\alpha,\beta}(C_1(i), C_2(i))$$

To simplify the practical computation of this quantity and facilitate the comparison to the KL-divergence method in the following section, we discretize our two dimensional color space into $j = 1 \ldots m_1$ and $k = 1 \ldots m_2$ bins and define $p_{\alpha,\beta}(j,k)$ as the probability that observed color j, k would appear in an image from that class given a set of illumination parameters α and β, and $\widetilde{p}(j,k)$ as the empirical probability that observed color j, k actually appears in the specific image we are working with. Given these definitions, the log likelihood takes the following form:

$$L(\alpha, \beta \mid \widetilde{P}) = n \sum_{j=1}^{m_1} \sum_{k=1}^{m_2} \widetilde{p}(j,k) \log p_{\alpha,\beta}(j,k)$$

Note that this new expression is now parameterized over the colors in the color space instead of the pixels in the image. The discretization has allowed us to collect up and count all of the pixels which have the same color and capture those counts in the empirical distribution, $\widetilde{p}(j,k)$.

3.3 Parameter Estimation Using KL-Divergence

In this work we explore the use of the Kullback-Leibler (KL) divergence, a measure of the closeness of two distributions, as a means of selecting a set of illumination parameters. Typically KL-divergence is used to measure the closeness between a true distribution, P and some approximate distribution, $\widetilde{P}$. The KL-divergence for a discrete distribution is defined as follows, as per [2], where m is the number of classes in the discrete distribution:

$$KL(P \| \widetilde{P}) = -\sum_{x=1}^{m} p(x) \log \frac{\widetilde{p}(x)}{p(x)}$$

The closer the two distributions are to one another, the smaller the KL-divergence. In this work we calculate the KL-divergence between the expected distribution of observed colors over a set of images drawn from some class of images, assuming a specific set of illumination parameters, and the actual distribution of observed colors for the specific image we are working with. Using our discretized two dimensional color space introduced in the previous section the KL-divergence takes the following form:

$$KL(P_{\alpha,\beta} \| \widetilde{P}) = -\sum_{j=1}^{m_1} \sum_{k=1}^{m_2} p_{\alpha,\beta}(j,k) \log \frac{\widetilde{p}(j,k)}{p_{\alpha,\beta}(j,k)}$$

In order to more easily compare this to the maximum likelihood formulation, we note that the summation of the terms related to the image class model is approximately a constant because it is related to the entropy of the canonical color distribution which does not change significantly as it shifted. (It is not identically a constant due to minor clipping effects at the edges of the color space.) This results in the following expression:

$$-KL(P_{\alpha,\beta} \| \widetilde{P}) \approx \sum_{j=1}^{m_1} \sum_{k=1}^{m_2} p_{\alpha,\beta}(j,k) \log \widetilde{p}(j,k) + constant$$

We can put the maximum likelihood formulation in a similar form to facilitate the comparison:

$$L(\alpha, \beta \mid \widetilde{P}) \propto \sum_{j=1}^{m_1} \sum_{k=1}^{m_2} \widetilde{p}(j,k) \log p_{\alpha,\beta}(j,k)$$

In this form, both equations appear to be quite similar. One way to understand the difference between the two is to evaluate them as two different functions for scoring possible matches between the canonical color distribution and the observed, empirical color distribution of an image. The likelihood formulation is more sensitive to large peaks in the observed color distribution and greatly penalizes solutions where the peak observed image colors are far from what would be expected, given a particular set of illumination parameters and the canonical color distribution. The KL-divergence formulation, because it contains a term which is a function of the $\log$ of the observed color distribution, is less sensitive to large peaks in the observed image color distribution and tends to be less affected by images which contain large amounts of a single color. Another way to examine these two formulations is to identify the conditions under which each would consider a particular match to be the "best" match. In the case of the likelihood formulation,

the best match occurs when as many of the observed colors as possible are aligned with high likelihood colors in the canonical distribution. In the case of the KL-divergence formulation, the best match occurs when the observed color distribution is, in effect, most similar to the canonical color distribution.

3.4 Estimating the Canonical Color Distribution

In order to implement these parameter estimation algorithms it is necessary to estimate the probability distribution of the canonical colors for a specific image class:

$$P(S_1(i), S_2(i))$$

We decided to take a simple approach and approximate this distribution as a histogram, as a table of counts. To estimate these counts a set of images from the desired image class is needed with ground truth information regarding the associated illumination parameters or, alternately, a set of images captured under a single canonical illuminant and camera.

Unfortunately it is quite difficult to collect a large number of images under controlled conditions. To avoid this issue we use *bootstrapping*, as described in [15], to approximate the ground truth. What we mean by *bootstrapping* in this case is that we utilize the estimates from some other color constancy algorithm as a proxy for the ground truth for a set of images. This would seem to be problematic in that it might limit any algorithm based on these estimates to perform only as well as the original algorithm used to generate the estimates. However, if the errors made by the "base" algorithm are relatively unbiased, then the estimates of the parameters of the resulting approximate distribution will converge to their true mean values, but may or may not be useful depending on their variance.

We chose as our image class natural images as might accompany a news article. As a data set for bootstrapping we used approximately 2300 randomly selected JPEG images from news sites on the web. These images consisted mostly of outdoor scenes, indoor news conferences, and sporting event scenes. (Examples of these images are available at: http://www.cs.cmu.edu/∼chuck/iccv01/.) We used the white patch algorithm (the illuminant is estimated by taking the peak value in each color channel) as our "base" algorithm. We utilized the log based color space described in a previous section and discretized the color space into 128×128 histogram bins, Figure 1a is a graphic representation of that color space. We initialized each histogram bin with a small number of initial counts to implement a smoothing prior and compensate for the fact that certain colors might not appear in our training data. Figure 1b is a plot of the probability distribution collected, where lighter regions represent higher probability values, white in the color space is at the center of the diagram. Note that no prior distribution over illumination parameters was collected because we assumed a uniform prior distribution over the illumination parameters.

4 Results

4.1 Evaluation Specifics

The data set used for the evaluation was collected by the Computational Vision Lab at Simon Fraser University as detailed in [14]. Two images each of 11 different household objects under five different illuminants on black backgrounds were collected: two different colored balls, a book, a coffee cup, a cereal box, flowers, a bottle of bleach, a box of macaroni, colored rope, a shampoo bottle and a detergent box. The five illuminants that were used were: Halogen, Cool White Fluorescent, Ultralume Fluorescent, the Macbeth Judge II 5000 Fluorescent, and the Macbeth Judge II 5000 Fluorescent with a Roscolux 3202 blue filter. They describe the camera used to capture the images as a Sony DXC-930 3-CCD color video camera balanced for 3200K lighting with gamma correction turned off. Each image also included gray patches which were used to determine the best fit illumination parameters under the diagonal model. These gray patches were cropped before the images were presented to the algorithms under evaluation. Image intensity values were scaled by a factor of 2.5 and values over 255 were clipped to more closely match the exposure settings under which images might typically be captured, as suggested in [14].

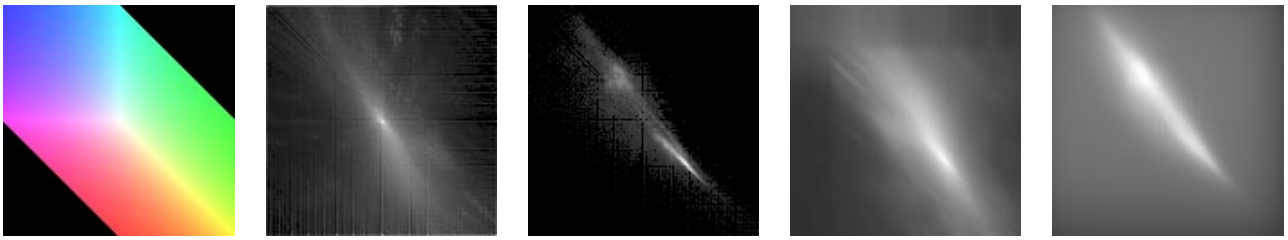
The error in the estimated illuminant values for each algorithm is reported as the Euclidean distance between the chromaticity of the "best fit" illuminant to the estimate produced by the algorithm being evaluated. (The "best fit" illuminant was estimated as described in the previous paragraph.) The distance error (where larger distances represent larger errors) was measured in the following intensity normalized color space:

$$D_1 = \frac{R}{R+G+B}; D_2 = \frac{G}{R+G+B}$$

This color space, which is different than the one used for the parameter estimation process, was used because it is the same color space used by [13] and therefore allows for a more direct comparison of results.

4.2 Algorithm Specifics

In our evaluation we compared the global illumination estimation performance of our KL-divergence based algorithm to the following algorithms: the likelihood based algorithm described previously (similar to color by correlation as described in [10]), gray world, white patch and white



(a) (b) (c) (d) (e)

Figure 1. Plot (a) is a graphic representation of the log color space, $\log\left(\frac{R}{B}\right)$ is on the vertical axis and $\log\left(\frac{G}{B}\right)$ is on the horizontal axis, the minimum value for both axes is in the upper left corner, this image is available in color at: http://www.cs.cmu.edu/$\sim$chuck/iccv01/. Plot (b) is a graphic representation of probability distribution of canonical colors, brighter regions are regions of higher probability, the overall brightness has been boosted to show structure. Plot (c) is the observed color distribution for the book image under the MB-5000 illuminant shown in Figure 2. Plot (d) is the likelihood based illuminant posterior and (e) is the equivalent plot for KL-divergence.

patch 1%. The gray world algorithm estimated the illuminant using the mean values of image color channels. The white patch algorithm estimated the illuminant using the color channel maximums. White patch 1% estimated the illuminant using the average of the brightest 1% of the pixel values of each of the color channels, to reduce noise sensitivity. To find the best estimate for the KL-divergence and likelihood based algorithms, each of the illumination parameters, α and β were each discretized into 128 levels and an exhaustive search was performed over all 16384 possibilities.

4.3 Experimental Results

The most basic evaluation of these algorithms is the error of the chromaticity of the estimated illuminant. Table 1 details the average distance results for the algorithms evaluated. The column labeled "SE of Error" is the standard error, a common measure of confidence for the mean error measure. As this table demonstrates, the KL-divergence based algorithm empirically achieved the lowest error, 0.0591.

Another way to evaluate these algorithms is to calculate the error based on compensated illumination estimates, as per [15]. The compensation process attempts to factor out any systematic error that an algorithm might have, for example the gray world algorithm explicitly assumes that white should be mapped to $R = G = B$, but in a specific canonical color space white might map to some other RGB triplet. We computed a compensation for each algorithm by splitting the test images into two equal sized sets and computing the mean difference in the color estimates between the images in the compensation set and the ground truth for those images. The result of the compensation process is a single constant value, for each algorithm, which is added to its estimate on the test set. Both the likelihood algorithm and the KL-divergence based algorithm obtained a significant reduction in error through the compensation process. The likelihood algorithm had a compensated mean error of 0.0692 (standard error of this estimate is 0.0078) and the KL-divergence had a mean compensated error of 0.0543 (standard error of this estimate is 0.0051). Even though both algorithms benefited from compensation, the KL-divergence based algorithm still achieved the lowest average error.

Algorithm	Mean Distance Error	SE of Error
White Patch	0.0999	0.0059
White Patch 1%	0.0837	0.0042
Gray World	0.0833	0.0054
Likelihood	0.0810	0.0064
KL-Divergence	0.0591	0.0044

Table 1. This table contains the error values for uncompensated illumination estimates evaluated on a test set. The column labeled "SE of Error" is the standard error, a common measure of confidence of the mean distance error.

It is also informative to look at examples of specific images. Figure 2 contains images of the "book" object, a predominantly yellow object, where the KL algorithm works particularly well when compared to the other algorithms evaluated here. (Note that all of the images in this figure are color images available

at: http://www.cs.cmu.edu/∼chuck/iccv01/ and were post-processed after being color corrected so that they would have the same average intensity to facilitate a direct comparison.) The KL algorithm performs better because it gives higher weight to the non-yellow colors in the image, whereas the likelihood algorithm tries to shift the predominant color towards gray. This effect can be observed in the plots of the posterior illuminant distribution in Figure 1, subfigures (d) and (e). A similar result was observed over all the "book" images in the data set, not just the single example included in figure 2. The mean uncompensated error for the "book" object over all illuminants was 0.0390 for the KL algorithm versus 0.2421 for the likelihood algorithm. There was also a single image of a specific object in our test set in which the KL algorithm did much worse than the likelihood algorithm, it achieved an uncompensated error of 0.2144 versus 0.1033 for the likelihood based algorithm. In this image of the "ball1" object, one of the colors had a very low measured intensity under halogen illumination and the KL algorithm aligned neutral gray with one of the primary colors in the image. It should be noted that the overall mean uncompensated error over all illuminants for this object was 0.1431 versus 0.1394, both algorithms performing similarly. (Color images of this example are available at: http://www.cs.cmu.edu/∼chuck/iccv01/.)

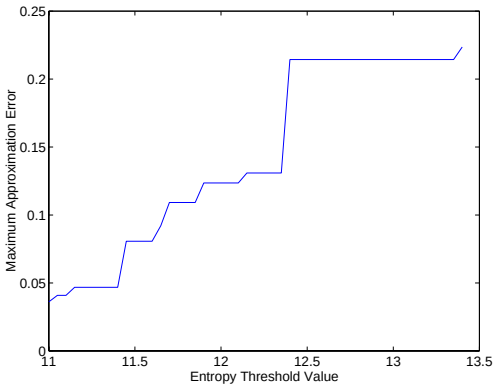


Figure 3. This is a plot of the maximum approximation error versus entropy threshold for the KL-divergence algorithm. (See the text for more details.)

Another important aspect of color constancy which has often been overlooked, but was recently explored in [9] with another technique, is the ability of an algorithm to generate some measure indicative of the quality of its prediction. Although we have not completed an in depth evaluation, we have empirically examined this issue. The measure we computed was an entropy-like measure: we treated the negative of the KL-divergence of each illuminant estimate as if it were a log probability value and calculated the entropy of that distribution. This measure has the property that if there tends to be a single dominant peak, the entropy will be low. And if there is no clearly superior solution, the entropy will be high. Figure 3 plots maximum error on the vertical axis and the entropy threshold value on the horizontal axis for the images in our test set. The entropy threshold value can be interpreted as follows: all of the images whose estimates had an entropy value less than the threshold value are included in that data point on the graph. So at the left hand side of the graph only the images with the lowest entropy estimates are included and on the right hand side all of the images in the test set are included. As this graph demonstrates, there is a strong correlation between this entropy measure and the maximum observed error. This suggests that this might provide a useful means of estimating the worst-case error of a particular estimate generated by the algorithm.

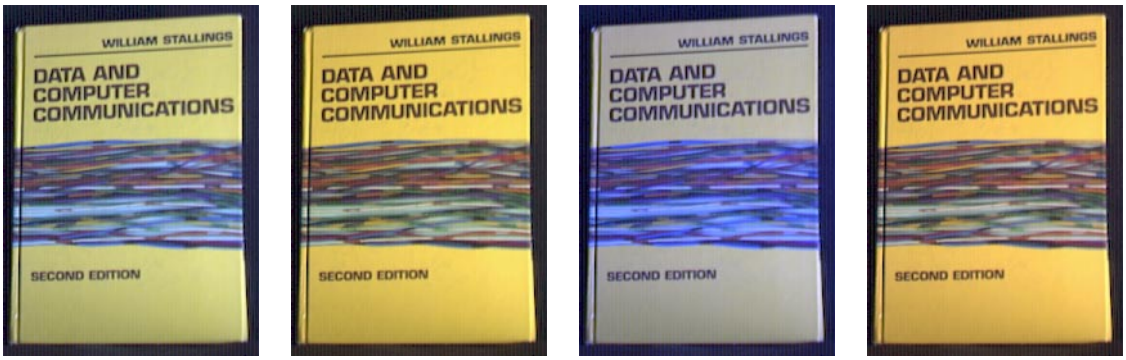
5 Related Work

Many color constancy methods have been described in the literature: [17], [16], [11], [7], [12], [13], [4], [8], [1], [19], [15], [10]. One class of algorithms described by Forsyth [11] and Finlayson [7] utilizes constraints about the distribution of the extreme colors in an image and a search procedure to determine the transformation which best satisfies the given constraints. Funt [13] uses a neural network which takes as input the image color histogram to estimate chromaticities. Recent work by Funt [15] has made use of bootstrapping techniques, as was used here, as a source of training data. Freeman and Brainard, in [12] and [4], also describe a probabilistic model and attempt to solve the more ambitious problem of recovering full surface reflectance spectra.

The work presented here is most closely related to the work described by Finlayson [8], [10] and Sapiro [19]. Both of the algorithms described in these papers are the basis of the maximum likelihood approach that we use as a benchmark for evaluating our algorithm. Both, in effect, assume a generative model of the data and find a set of parameters which have the maximum likelihood given the model and the observed data. Our technique is different in that it utilizes a distance measure between distributions, the KL divergence, to select a set of estimated parameters.

6 Conclusions and Future Work

We have demonstrated empirically that KL-divergence appears to be a robust means of estimating a good set of global illumination parameters, outperforming an algorithm we evaluated which used maximum likelihood as its basis. One possible explanation for this is that the probabilistic



Uncorrected Best Fit Likelihood KL-Divergence

Figure 2. An example of an image where the KL algorithm works particularly well is the "book" image, captured here under the MB-5000 illuminant. The "Uncorrected" image is the raw uncorrected image. The other images are renderings of how the image would look if corrected given the illumination parameters estimated by the specific algorithm. The "Best Fit" image is the best possible rendition, given the ground truth and a diagonal illumination model. Here the likelihood algorithm achieved an uncompensated error of 0.2756 and the KL algorithm achieved an error of 0.0312. (Note that these are color images and are available at: http://www.cs.cmu.edu/∼chuck/iccv01/.)

models typically used in these maximum likelihood settings do not accurately model the generative process of image formation. It is possible that with a better generative model (possibly one which explicitly models the dependencies between measured pixel colors), the maximum likelihood approach would fare better. However, building such a model is no simple task. Our experiments also empirically demonstrated that an entropy-like measure can be used to predict a bound on the maximum error of the illuminant estimate.

There are a number of directions for future work. The first is to improve the speed of our algorithm. Currently we exhaustively search for the best estimate. We believe that a gradient descent like approach or a multi-resolution approach in color space would be faster and yield similar quality results. We would also like to better understand the role of the canonical color distribution and how tightly coupled it is to obtaining good results for a specific class of images. We would also like to explore the use of a parametric model of the canonical color distribution to see if it would work as well as the histogram based approach we have taken. A parametric model has the advantage of having less free parameters and requiring less data to train.

Acknowledgements

We would like to thank the Computational Vision Lab at Simon Fraser University for collecting the data set used in this work and making it publicly available. This research is sponsored in part by National Science Foundation LIS grant number REC-9720374, which is gratefully acknowledged.

References

[1] K. Barnard, *Practical colour constancy*, PhD thesis, Simon Fraser University, School of Computing, 1999.

[2] C. M. Bishop, *Neural Networks for pattern recognition*, New York: Oxford University Press, 1995, pp. 58-59.

[3] R. A. Boie and I. J. Cox, "An analysis of camera noise," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, June 1992.

[4] D. H. Brainard and W. T. Freeman, "Bayesian color constancy," *Journal of the Optical Society of America A*, vol. 14, no. 7, pp. 1393-1411, 1997.

[5] M. H. DeGroot, *Probability and statistics (second edition)*, Menlo Park: Addison-Wesley Publishing Company, 1989, pp. 158-169.

[6] G. D. Finlayson and M. S. Drew and B. V. Funt, "Diagonal transforms suffice for color constancy," *Proceedings of the International Conference on Computer Vision*, pp. 164-171, 1993.

[7] G. D. Finlayson and B. V. Funt and K. Barnard, "Color constancy under a varying illumination," *Proceedings of the Fifth International Conference on Computer Vision*, pp. 720-725, 1995.

[8] G. D. Finlayson and S. D. Hordley and P. M. Hubel, "Color by correlation," *Proceedings of the IS&T / SID Fifth Colour Imaging Conference*, pp. 6-11, 1997.

[9] G. D. Finlayson and S. D. Hordley, "Colour constancy with error bars," *Proceedings of the Seventh International Conference on Image Processing and Its Applications*, vol. 1, pp. 368-371, 1999.

[10] G. D. Finlayson and S. D. Hordley and P. M. Hubel, "Colour by correlation: a simple, unifying approach to colour constancy," *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, pp. 835-842, 1999.

[11] D. A. Forsyth, "A novel algorithm for color constancy," *International Journal of Computer Vision*, vol. 5, no. 1, pp. 5-46, 1990.

[12] W. T. Freeman and D. H. Brainard, "Bayesian decision theory, the maximum local mass estimate, and color constancy," *Proceedings of the International Conference on Computer Vision*, 1995.

[13] B. Funt and V. Cardei and K. Barnard, "Learning color constancy," *Proceedings of Imaging Science and Technology / Society for Information Display Fourth Color Imaging Conference*. pp. 58-60, 1996.

[14] B. Funt and K. Barnard and L. Martin, "Is colour constancy good enough?," *Proceedings of the Fifth European Conference on Computer Vision*, pp. 445-459, 1998.

[15] B. Funt and V. Cardei. "Bootstrapping color constancy," *Proceedings of SPIE: Electronic Imaging* IV, 3644, 1999.

[16] A. Gershon and A. D. Jepson and J. K. Tsotsos, "From [R,G,B] to surface reflectance: computing color constant descriptors in images," *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 755-758, 1987.

[17] L. T. Maloney, and B. A. Wandell, "Color constancy: a method for recovering surface spectral reflectance," *Journal of the Optical Society of America A*, vol. 3, no. 1, pp. 29-33, 1986.

[18] H. A. Rowley and S. Baluja and T. Kanade. "Neural network-based faced detection,", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp, 23-28, 1998.

[19] G. Sapiro, "Color and illuminant voting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 11, pp. 1210-1215, 1999.