

White Point Estimation for Uncalibrated Images

Vlad C. Cardei, Brian Funt and Kobus Barnard
School of Computing Science, Simon Fraser University
Vancouver, Canada V5A 1S6

Abstract

Color images often must be color balanced to remove unwanted color casts. We extend previous work on using a neural network for illumination, or white-point, estimation from the case of calibrated images to that of uncalibrated images of unknown origin. The results show that the chromaticity of the ambient illumination can be estimated with an average CIE Lab error of 5ΔE. Comparisons are made to the grayworld and white patch methods.

Introduction

To remove a color cast from an image so that it is properly color balanced, the color of the scene illumination must be estimated. Illumination estimation in this sense is also commonly referred to as white point estimation. Even when the imaging device's characteristics are fully known, accurate illumination for color cast removal has proven difficult, but there has been progress.[1–6] In this paper, we test a neural-network-based method for estimating the illuminant in the general case in which the imaging parameters are unknown.

In the case of digital photography images, the camera can be calibrated so that the sensor sensitivities as a function of wavelength are known as well as their response as a function of intensity. Many existing color constancy algorithms[1–6] depend on having calibrated sensors. In many situations—images downloaded over the Internet or scanned from film—the imaging characteristics are either unknown or else, as in the case of film, very difficult to control. In previous work,[7] we proposed a computational framework for color correction of uncalibrated images given an estimate of the illumination chromaticity. In this paper, we test how well various illumination-estimation algorithms (greyworld, white patch, neural net, and bootstrapping) work when presented with uncalibrated image data.

Unknown Imaging Parameters

Calibrating specifies three important aspects of a 3-channel color imaging system: (1) the response of each channel as a function of intensity; (2) the white balance, which is an input spectrum which creates equal, and usually maximal, output across all 3 channels; and (3) for each channel, the relative sensor sensitivity as a function of wavelength. We will assume there is no spatial variation in these characteristics.

In terms of the channel response as a function of intensity, it is possible for this function to be an arbitrary function but generally it will be monotonic. However, we will restrict our attention to "gamma" functions of the form $I=SD^{\gamma}$, where I is the resulting luminance, S is the camera gain and D is a pixel value in the 0..1 range.

Such functions are typical in imaging systems. Poynton[8] discusses gamma in detail. In what follows, we will assume that any non-linearity in the sensor response has been created by gamma, but that the value of gamma is not known since it generally differs between imaging systems. We will term images for which gamma does not equal unity to be 'gamma-on' images. Linear images are 'gamma-off' and have gamma equal one.

For a gamma-on image, we have shown[7] that it is possible to color correct it by a diagonal transformation without first linearizing the image. The off-diagonal terms of the general image transformation (i.e. the best linear fit) are larger for gamma-on images than gamma-off ones, so the average error of a diagonal transformation (which ignores the off-diagonal terms) is greater. However, the perceptual error induced by such a transformation is still small. Also, although gamma introduces a color shift, the shift is independent of pixel intensity. Finally, diagonal color correction[9] and application of gamma are commutative. As a result, we can color correct gamma-on images in the same way we as linear, gamma-off images.

The second problem in handling uncalibrated images concerns the imaging device's color balance. We assume that the device was balanced so that it produced equal RGB values for a white patch under some chosen illuminant, but that we do not know which illuminant it was. In theory, color correcting an image taken with an unknown balance does not pose a problem, since the calibrating coefficients (used to scale the sensors) can be absorbed into the diagonal transformation required for color correction. In effect, the color balance problem folds in with the illumination. However, finding the resulting diagonal transformation could prove difficult for the algorithms[4,5] whose results depend on an expected set of illuminants. The combined balance-illumination transformation will likely fall outside the range of expected illuminants thereby causing inaccurate results.

The third aspect of working with uncalibrated images is not knowing the sensor sensitivity functions. Sensors differ significantly, even for the simple case of digital cameras balanced for the same illuminant. Two camera models balanced for the same illuminant will by definition have the same response to white, but can have different sensors responses to other colors. Figure 1 shows the difference between responses of a SONY DXC-930 and a Kodak DCS460. To eliminate the effects of noise or other artifacts from the comparison, RGB values were synthesized using the sensor sensitivity curves of the two cameras along with the surface reflectances of the 24 Macbeth Colorchecker patches. Both cameras were balanced for the same illuminant. Figure 1 plots the sensor responses in rg-chromaticity space:

$$r=R/(R+G+B) \text{ and} \tag{1}$$

$$g=G/(R+G+B) \tag{2}$$

It can be seen in the figure that the responses vary significantly. This variation in sensor response can adversely affect color constancy algorithms that rely on prior distributions of sensor responses.

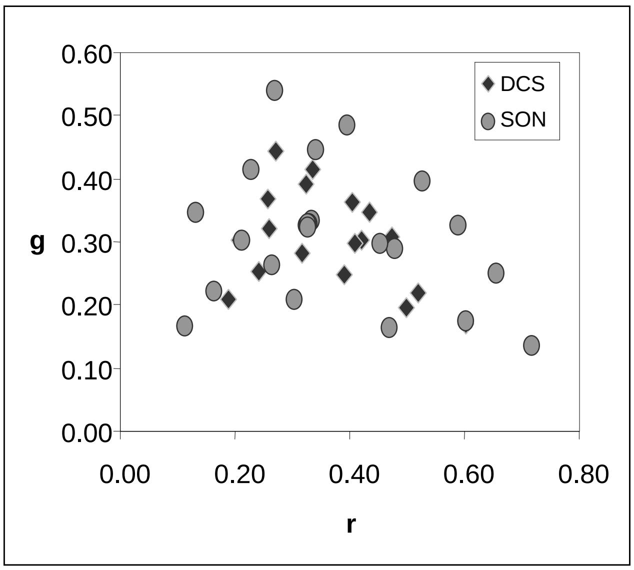

Figure 1. Comparison of rg-chromaticities obtained by two cameras balanced for the same illuminant.

Illumination-Estimation Algorithms

We test several different illumination-estimation algorithms on a database of 'uncalibrated' images. The images are uncalibrated in the sense that the imaging characteristics are not provided to the algorithms, even though we have the calibration parameters available so that we can evaluate the results. In particular, we test the white patch algorithm (WP), a version of the grayworld algorithm (GW) and two neural-network-based methods. The gamut-constraint methods[3,4] were not tested because they require information about the expected gamuts of reflectances or illuminants.

The image database contains 116 images taken with a Kodak DCS-460 camera and 67 images scanned with a Polaroid Sprintscan 35+ slide scanner from various film types: Kodak Gold, Kodak Royal, Agfa Optima, Polaroid HiDef and Fuji Superia. The slides were scanned using a 'generic' pre-defined scanner setting. This setting is consistent with the assumption of unknown pre-processing. Using the manufacturer's setting for each film type would have allowed the scanner driver to accomdate for the different film characteristics.

We divide the image database into two sets, the first for training and the second for testing. The training set contains 102 images and is used for training the neural network and computing the average color for use in the database grayworld algorithm. The test set contains the other 81 images (57 DCS images and 24 slides).

For the GW algorithm, the chromaticity of the illuminant is determined from the average of all the pixels in an image. GW assumes that the average color of the scene is gray and that any departure from this average in the image is caused by the color of the illuminant. The average is computed relative to the average chromaticity computed using all pixels in the training database. Using the database average as the definition of gray compensates for the fact that gray may not have exactly equal r and g chromaticities. Nonetheless, GW's performance will be poor when the test images have different average distributions than the ones used for computing the database average.

The WP algorithm determines white, and hence the illuminant color, as the maximum R, maximum G and maximum B found in the image. The WP algorithm has roots in the family of retinex algorithms,[1] but it is only equivalent to it under restricted circumstances.

Two differently trained neural networks were used for illumination estimation. The network architecture was the same in both cases; namely, a Perceptron with two hidden layers as we have previously described.[10,11] The networks are trained to estimate the chromaticity of the illuminant based on the binarized rg-chromaticity histogram of an input image. The 3600-node input layer is fed binary values representing the presence or absence of chromaticities falling within a particular chromaticity bin. The first hidden layer contains 50 neurons and the second layer 20 neurons. The output layer consists of only two neurons representing the chromaticity of the illuminant. All neurons have a sigmoid activation function.

Both neural networks were trained using the back propagation algorithm. The error function for training and testing is the Euclidean distance in rg-chromaticity space between the actual illuminant and its estimate.

The difference between the training of the two networks concerns the method of determining the actual illuminant. For the first network, the illuminant chromaticity is simply measured from the reference white standard that was contained within each image. This provides an accurate value for the illuminant's chromaticity. For the second network, a less accurate method is used, which we have called the bootstrapping method.[11] The bootstrapped network uses the GW algorithm to "measure" the chromaticity of the illuminant for training. Clearly, the illuminant value determined by GW will only be approximately correct; nonetheless, previous experiments with calibrated image data have shown that the network "learned" to make a better estimate than the simple GW algorithm used to train it. Our new experiments described below show that bootstrapping works even for the more general case of non-linear images acquired from various sources. This approach allows us to train a neural network for a range of uncalibrated cameras and scanners, without explicitly having to measure white patches in the set of training images.

Experimental Results

The algorithms presented above were tested on an image database containing 81 images. Figures 2 and 3 show the relative performance of the color constancy algorithms. The figures show the average errors over the whole test set as well as for each type of input (i.e. for DCS images and slides).

In Figure 2, the average errors are computed in the rg-chromaticity space, the same space in which the neural network was trained. "Nothing" refers to assuming that the illuminant is the one for which the device is calibrated and reflects the variation in the chromaticity of the illuminant across the test set of images, relative to white (located at r=g=1/3 in rg-chromaticity space). "NN" refers to the neural network trained with accurately measured illumination data, while "Bootstrapped NN" refers to the same network trained using GW illumination estimates.

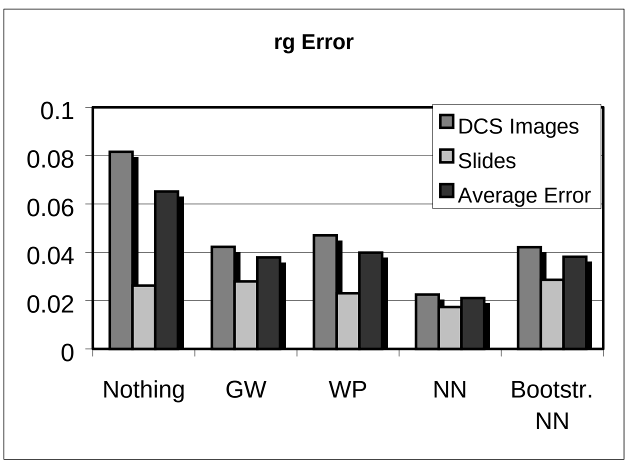

Figure 2. Average errors measured in rg-chromaticity space.

Figure 3 presents similar results, but with the error measured in CIE Lab space. The conversion from the RGB space to CIE Lab assumes the images are to be viewed on a sRGB-compliant[9] monitor.

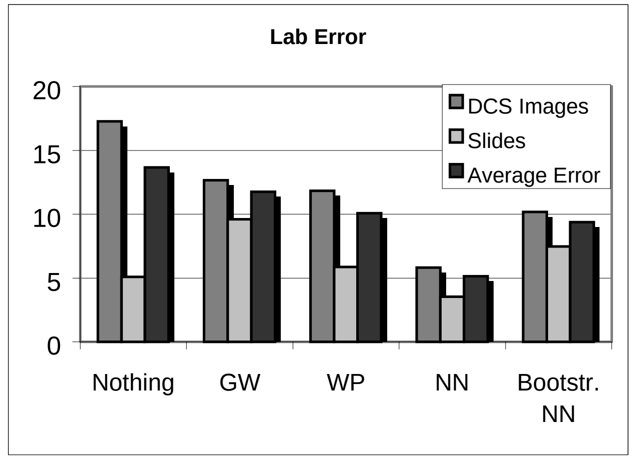

Figure 3. Average CIE Lab ΔE space between actual and estimated illuminant fixed to the same L value.*

Discussion

Previous studies[10,11] based on calibrated, linear image data have shown that a neural network can accurately estimate the illumination chromaticity. Often we must work with uncalibrated image data, so we trained and tested several algorithms on uncalibrated data, but in a controlled manner. On this test data, the neural net average error is 5.14ΔE. We believe this to be useful for removing color casts from images of unknown origin. In the tests with the bootstrapping method of training the neural network, the ΔE error increased to 9.38. Nonetheless, this is better than either the GW or WP methods and the bootstrapping method can be applied in situations where accurate measurements of the illuminant chromaticity are unavailable for training.

Acknowledgements

The authors gratefully acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada and Hewlett Packard Incorporated, and Lindsay Martin and Fred Kyba for their help acquiring much of the image data used in the experiments.

References

1. E. H. Land, "The Retinex Theory of Color Vision," Scientific American, pg. 108-129. (1977).
2. G. Buchsbaum, "A Spatial Processor Model for Object Colour Perception," *J. Franklin Institute,* 310 (1), pg. 1-26. (1980).
3. D. A. Forsyth, "A Novel Algorithm for Color Constancy," *Int. J. of Computer Vision,* 5:1, 5-36. (1990).
4. G. Finlayson, "Color in Perspective," *IEEE Trans. PAMI*, 18 (10), pg. 1034-1038. (1996).

5. B. Funt, V. Cardei, and K. Barnard, Learning Color Constancy, *Proc Fourth Color Imaging Conf.,* pg. 58-60. (1996).
6. G. Finlayson, P. Hubel and S. Hordley, Color by Correlation, *Proc Fifth Color Imaging Conf.*, pg. 6-9. (1997).
7. V. Cardei, B. Funt and M. Brockington, Issues in Color Correction of Images from Unknown Origins, Proc. *CSCS'12. (*1999).
8. C. Poynton, "The Rehabilitation of Gamma," B.E. Rogowitz and T.N. Pappas (eds.), *Proc. of SPIE 3299*, pg. 232-24. (1998).
9. G. Finlayson, M. Drew, and B. Funt, "Color Constancy: Generalized Diagonal Transforms Suffice," *J. Opt. Soc. Am. A*, 11(11), pg. 3011-3020. (1994).
10. V. Cardei, B. Funt, and K. Barnard, Adaptive Illuminant Estimation Using Neural Networks, *Proc. of the 8th Int. Conf. on Artificial Neural Networks*, pg. 749-754. (1998).
11. B. Funt and V. Cardei, Bootstrapping Color Constancy, *SPIE Electronic Imaging'99. (*1999).
12. M. Anderson, R. Motta, S. Chandrasekar and M. Stokes, Proposal for a Standard Default Color Space for the Internet – sRGB, *Proc. Fourth Color Imaging Conf.*, pg. 238-246. (1996).