Neural Network Colour Constancy and Specularly Reflecting Surfaces

Brian Funt
Vlad Cardei
Kobus Barnard
Simon Fraser University

To appear: Proceedings AIC Colour 97 Kyoto 8th Congress of the International Colour Association, May 1997

We previously developed a neural network which estimates the chromaticity of the illumination under which an given image was taken[1]. This provides colour constancy since, given the chromaticity estimate, the image pixel chromaticities can be converted via a diagonal transformation to what they would be under a canonical illuminant. In tests on synthetically generated and real scene images, the accuracy of the illumination-chromaticity estimate generally surpassed that of most existing color constancy algorithms; however, the errors obtained with real images were significantly larger than those for the synthetic ones. After experiments with adding noise to the synthetic data, we concluded that there was a more fundamental problem than simply the influence of noise which remained to be explained. We hypothesized that specular reflection was causing the problem, so we modeled the specular reflection in the training set. The errors dropped by more than 20 percent.

1. INTRODUCTION

Many colour constancy algorithms[2,3,4,5,6] assume matte surface reflection properties for the objects appearing in images. Lee's[7] algorithm and its descendants[8] provide a notable exception in that they explicitly exploit specularities in calculating the illuminant's chromaticity and will fail if there are no specularities. Those algorithms use the dichromatic model of specular reflection and depend on the fact that the spectrum of the specularly reflected component—that which is reflected directly from the surface of the object rather than entering the object—has approximately the same spectrum as the incident illumination.

We had designed and trained a neural network which on average out performs the matte-surface algorithms applied to real images that happened to include some moderately specular surfaces. While the results were good, they were not as good as we expected based on our tests on synthetic data. At first we could not understand why our tests on synthetic data were so much better than our tests on real images, but eventually we guessed that specularities were causing the problem.

2. NEURAL NETWORK DESCRIPTION

The neural network approach to color constancy involves a multi-layer Perceptron. The network's input is a discritized and binarized chromaticity histogram of a colour image (artificial or real) mapped into a one-dimensional space. The output of the network represents the chromaticity of the incident illumination.

All calculations are done relative to the 3-channel data obtained from, or synthesized with respect to, a calibrated SONY DXC930 3-CCD camera. We will use RGB to refer to this camera output and chromaticity to refer to the chromaticity space (R/(R+G+B), G/(R+G+B)).

The network is trained using the Back-propagation[9] algorithm. The network's input nodes are presented with the binarized histograms of the images of synthesized scenes and simultaneously its output nodes are presented with the chromaticities of the synthesized illuminations. The scenes were synthesized by combining randomly selected reflectance and illumination spectra from a set of 260 real reflectances and 89 real illuminants.

Based on the dichromatic model of reflection[10], the training set was modified to include random amounts of specularity. The dichromatic reflection model states that the reflected light is an additive mixture of a specular component—light reflected directly from the interface layer of the surface, thereby retaining the spectrum of the incident illumination—and a body component. The body component describes the light that enters the object's surface before being re-emitted. Therefore specularity was added to the training set simply by adding random amounts of the scene illumination's RGB to the matte component of the synthesized surface RGB's.

Two different neural network architectures were compared. Both were multi-layer Perceptrons, with two hidden layers. The first neural network contained 2500 nodes in the input layer, 400 nodes in the first hidden layer, 30 nodes in the second hidden layer and 2 in the output layer. All nodes were fully connected to the previous layer, except in the case of the first hidden layer in which each node makes only 200 connections to the input layer. The second neural network contained 3600 nodes in the input layer, 200 nodes in the first hidden layer, 50 nodes in the second hidden layer, and 2 nodes in the output layer. All nodes were fully connected to the previous layer, except for the first hidden layer where again each node had only 400 connections to the input layer.

Because of the large size of the input layer and the fact that the gamut of all the possible RGB's of physically realizable surface reflectances occupies only a portion of the input representation space, we used an adaptive technique which shortened the training time by almost an order of magnitude without affecting the resulting network's performance[11]. This technique consists of deleting the links to those nodes in the input layer which remained dormant throughout an entire training epoch. The deleted links were then replaced with new links to randomly selected input-layer nodes. This process stabilizes after only three or four epochs, at which time all input links point to active areas in the input space.

To train the networks, we used Back-propagation without momentum. By using different training rates for each layer, we further improved the training speed and stability. The learning rates (of 20 for the first hidden layer, 10 for the second hidden layer and 0.25 for the output layer) were kept constant during training. The error measure used to provide feedback to the network during training was the Euclidean distance in the chromaticity space between the target output (illuminant) and the estimated one.

3. THE TRAINING SETS

The networks were trained with large training sets containing synthesized scenes. Each training set consisted of 8900 artificially generated scenes (100 scenes for each of 89 illuminants). Each scene was generated by randomly selecting *n* surfaces (ranging from 10 to 100) from the surface reflectance database and then multiplying by an illuminant spectrum picked at random from the illuminant database. The camera's sensitivity functions were then applied to the resulting spectrum to produce a set of *n* RGB values. To these values, we added a random amount *r* of the scene illumination. The value of *r* for a scene *i* was computed as the product between the maximum value of the specular component *S* (usually in the range of 0%-100%) and a random, sub-unitary coefficient *p*: $r_i = S*p$

Since surface specularity is not uniformly distributed in a real image, we created a non-uniform distribution by squaring a uniformly distributed random function: $p=rnd()^2$. This model has a mean value of 25% of the maximum specularity and assures that generally only a few surfaces in the scene will be highly specular. A random amount of white noise to a maximum ±5% of the RGB values was then also included.

We generated training sets with different amounts of maximum specularity (ranging from 0% to 100%) and trained the networks for 10 epochs on each training set. All networks of the same architecture were trained starting from identical untrained networks. This assures that the training depends only on the training sets and not on the initial random weights of the network. In the end, we obtained a separate neural network for each training set.

After training, the average error in estimating the illumination chromaticity for the training set data ranged from 0.83% to 1.1%. When tested on scenes that were not part of the training set, the average error then ranged from 1.2% to 2.2%.

RESULTS

The neural networks obtained after training on synthetic data with varying amounts of specularity were then tested using a set of 48 **real** images. These real images were obtained using a variety of different light sources. The results are presented in the tables 1 and 2 below and compared with existing methods in table 3:

Specularity (%)	Mean Error	Std. Dev.	Improvement (%)
0% (no spec.)	.059	.043	-
5%	.051	.035	13.5%
10%	.044	.026	25.4%
25%	.044	.030	25.4%
>50%	≈.044	≈.035	25.4%

Table 1. Results for the 3600-200-50-2 network trained for different amounts of specularity and then tested on images of real scenes

Specularity (%)	Mean Error	Std. Dev.	Improvement (%)
0% (no spec.)	.058	.047	-
5%	.051	.037	11.3%
10%	.056	.038	3.4%
25%	.045	.036	13.9%
50%	.047	.032	19.1%

Table 2. Results for the 2500-400-30-2 network trained for different amounts of specularity and then tested on images of real scenes.

Method	**Mean**	**Std. Dev.**
Illumination Chromaticity Variation	.090	.062
Grey world using average R, G, and B	.071	.051
Retinex using maximum R, G, and B	.075	.049
2D gamut-constraint method using surface constraints only [3]	.054	.047
2D gamut-constraint method using surface and illumination constraints	.047	.039
Neural network with 25% specularity model	.044	.032

Table 3. Comparison of the 3600-200-50-2 to other colour constancy methods tested on real images**.** The top row is the average chromaticity difference in the light sources used.

CONCLUSION

The results above show that there is a significant improvement in the network performance when trained on data that models specular reflectance. The neural network also obtains more accurate estimates of the illumination chromaticity than any of the existing methods tested.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada and by a gift from Hewlett-Packard Incorporated.

REFERENCES

[1] FUNT, B.V., CARDEI, V., and BARNARD, K., "Learning Color Constancy," *Proc. IS&T/SID Fourth Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, Arizona, November, pp. 58-60, 1996.

[2] FORSYTH, D., "A novel algorithm for color constancy," *Intl. J. of Comp. Vis.* 5:5-36, 1990.

[3] FINLAYSON, G., "Color Constancy in Diagonal Chromaticity Space," *IEEE Proc. Fifth Intl. Conf. on Comp. Vision*, June 20-23,1995.

[4] MALONEY, L. and WANDELL, B., "Color constancy: a method for recovering surface spectral reflectance," *J. Opt. Soc. Am. A*, 3:29-33, 1986.

[5] LAND, E.H., and MCCANN, J.J., "Lightness and retinex theory," *J. Opt. Soc. Am*, 61:1-11, 1971.

[6] GERSHON, R. JEPSON, A.. TSOTSOS, J., "From RGB to Surface Reflectance: Computing Color Constant Descriptors in Images," *Perception* 755-758, 1988

[7] LEE,H., "Method for computing the scene-illuminant chromaticity from specular highlights," *J. Opt. Soc. of Am. A*, Vol 3, pp1694-1699, Oct.1986

[8] RICHARD, W.M., *Automatic Detection of Effective Scene Illuminant Chromaticity from Specular Highlights in Digital Images*, M.Sc. Thesis, Rochester Inst. of Tech., April 1995.

[9] HERTZ, J., KROGH, A., and PALMER, R.G. *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company, 1991.

[10] SHAFER, S.A., "Using color to separate reflection components," *Color Res. Appl.*,Vol.10, pp. 210-218, 1985.

[11] CARDEI V., FUNT B., BARNARD K., "Modeling Color Constancy with Neural Networks", *Proc. Int. Conf. on Vision, Recognition, and Action: Neural Models of Mind and Machine*, Boston, May 29-31, 1997

AUTHORS' ADDRESSES

Brian FUNT
Vlad CARDEI
Kobus BARNARD
School of Computing Science, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6.
E-mail: (funt, cardei or kobus)@cs.sfu.ca