Committee-Based Color Constancy

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

Abstract

We show how to achieve better illumination estimates for color constancy by combining the results of several existing algorithms. We consider committee methods based on both linear and non–linear ways of combining the illumination estimates from the original set of color constancy algorithms. Committees of grayworld, white patch and neural net methods are tested. The committee results are always more accurate than the estimates of any of the other algorithms taken in isolation.

Introduction

For our purposes, we consider the goal of a color constancy algorithm to be the precise estimation of the chromaticity of the scene illumination from image data alone. Given such an estimate, the image can be then color corrected to make it look as if it was taken under a standard, canonical illuminant.[1]

Our hypothesis was that by combining several color constancy algorithms, we could obtain a more accurate estimate of the illuminant than any of the algorithms provides individually. A similar approach is known in the neural network literature[2] as using committees of neural networks. Committees of neural networks are based on averaging the outputs of multiple neural networks, trained on the same data, in order to obtain smaller estimation errors. When the estimation errors are uncorrelated with zero-mean, it has been shown [2] that by using n neural networks, the average sum-of-squares estimation error is reduced by a factor of n, relative to the MSE (mean squared error) of individual networks. In practice, the reduction is much smaller because of systematic estimation errors and because the estimation errors of the neural networks are correlated. In any case, the average error given by the committee was found to be smaller than the average of the errors of the individual networks.

In this paper, we explore whether or not a committee of color constancy algorithms leads to a better color constancy. As 'members' of the committee, we used a version of the white patch algorithm,[3] the grayworld algorithm[4] and a neural network algorithm.[5,6]

The grayworld algorithm (GW) determines the chromaticity of the illuminant from an average of all the pixels in an image. The algorithm assumes that the average color of the scene is gray and that any deviation from gray of the image average is caused by the color of the illuminant. To compensate for the possible deviation from gray of the distribution of surface colors, the average is computed relative to the gray world average of the colors in the image database used for training.

The white patch (WP) algorithm independently scales each channel of the image (R,G,B) by the maximum pixel value found in each channel. This is equivalent to estimating the color of the illuminant as being the color given by the maximum pixel value on the R, G and B channels. WP derives from retinex,[3] but is only equivalent to it under special circumstances.

We also used a neural network[5,6] for estimating the chromaticity of the illuminant. This algorithm is more accurate than the other two methods, described above. The neural network was trained to estimate the chromaticity of the illuminant, based on the rg-chromaticity histogram of an image.

Experiments & Results

For our experiments, we used two similar data sets, each composed of 19,800 illuminant estimates. One, the training set, was used for optimizing the committees and the other one was used as a test set for validation. The results reported below are those obtained on the test set.

In a first set of experiments, we compared the individual performance of the NN, WP and GW algorithms to that of three types of committees. It should be noted that the NN algorithm has twice the accuracy of the GW and WP algorithms.

The first type of committee simply averages the outputs of the three color constancy algorithms. The individual r and g chromaticity estimates are averaged, as shown in Equation 1, and the resulting values r_c and g_c are compared to the actual illuminant chromaticities.

$$[r_{NN}\ \ g_{NN}\ \ r_{GW}\ \ g_{GW}\ \ r_{WP}\ \ g_{WP}] \cdot \begin{bmatrix} 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 1/3 & 0 & 1/3 & 0 \end{bmatrix}^T = [r_C\ \ g_C] \quad (1)$$

The second type of committee is a weighted average of the outputs of the individual algorithms. The weights were optimized in the least mean square (LMS) sense, and were computed from the data available in the training set. The actual values of the weights are shown in Equation 2. It is interesting to notice the cross-talk between the red and green channels (i.e. the influence of the green estimates on those of the red).

$$[r_{NN} \quad g_{NN} \quad r_{GW} \quad g_{GW} \quad r_{WP} \quad g_{WP}] \cdot \begin{bmatrix} 0.002 & 0.807 & -.018 & 0.040 & 0.015 & 0.150 \\ 0.675 & 0.113 & 0.041 & -.045 & 0.260 & -.060 \end{bmatrix}^T = [r_C \quad g_C] \quad (2)$$

The first two types of committees are linear. It is possible that there could be some higher-order correlation involved between the different estimates that are not captured by the linear models. Neural networks are good at modeling such non-linear statistical properties, so we experimented with a third type of committee—a neural network (a multi-layer Perceptron) trained to estimate the illuminant, based on estimates provided by the other three color constancy algorithms.

We tried various network architectures and trained each network a number of times starting from different random initial weights. The network with the smallest average error over the training set has 6 inputs to the neural network, 6 nodes in a hidden layer and two outputs nodes. The 6 input nodes encode the illuminant estimates from the 3 algorithms, while the output nodes encode the new chromaticity estimate. The network was trained on the training set for 50,000 epochs.

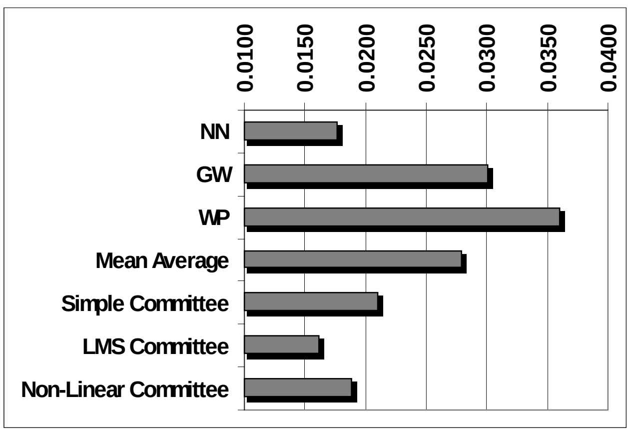

Figure 1. The average RMS error of the 3 raw algorithms and the various committees. NN, GW and WP denote the raw neural network, grayworld and white-patch algorithms' mean RMS errors. The 'Mean Average' is the average of the NN, GW and WP means. 'Simple Committee' refers to the linear committee based on the simple unweighted average of the raw estimates, 'LMS Committee' to the optimized weighted average of raw estimates, and 'Non-linear Committee' to the neural network method of combining results.

The average RMS error for each of the original algorithms as well as the three committees is plotted in Fig. 1 where it can be seen that all three committees result in smaller average errors than the mean error of the raw color constancy algorithms (NN, GW and WP) working alone. The LMS committee provides an 8% improvement over the raw neural network. Despite the generality of the neural network's architecture, this shows that the GW and WP methods still have something additional to offer when their results are combined with the neural network's in an appropriate way.

It is interesting to note that the non-linear committee does not perform as well as the linear LMS committee. This leads to the hypothesis that there are no higher-order statistical relationships between the estimates of the raw color constancy algorithms. Of course, our failure to find a non-linear network architecture with better performance does not prove this hypothesis.

The committee of three algorithms worked well. Would a committee of only WP and GW work well also? Since the non-linear committee method did not work as well as the linear committees, we restrict our attention to the two linear ones based on a simple averaging and LMS optimized weights. Equation 3 shows the simple averaging method, while Equation 4 shows the actual weights, obtained from the training set through the LMS method.

$$[r_{GW} \quad g_{GW} \quad r_{WP} \quad g_{WP}] \cdot \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}^T = [r_C \quad g_C] \quad (3)$$

$$[r_{GW} \; g_{GW} \; r_{WP} \; g_{WP}] \cdot \begin{bmatrix} 0.012 & 0.479 & -0.003 & 0.501 \\ 0.471 & 0.012 & 0.474 & 0.009 \end{bmatrix}^T = \quad (4)$$
$$= [r_C \; g_C]$$

In Figure 2 we compare the results obtained by the WP and GW color constancy algorithms, as well as the two linear committee methods. LMS committee performance improves by 12% over the GW algorithm and 26% over the WP algorithm.

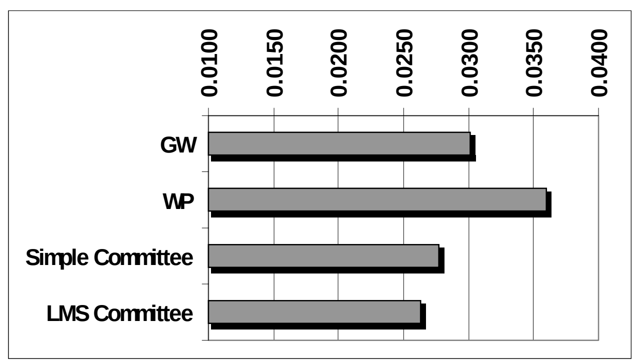

Figure 2. RMS average error of individual algorithms and committees

Systematic errors in the raw algorithms could adversely affect committee performance. In particular, GW is prone to systematic errors if the colors in the test images do not average to the database average used to compensate for the deviation from gray. To test the effect of systematic error on the committees, we introduced a systematic shift into the data set by assuming that the red component of the RGB values of the surfaces in the test set is 10% higher than its actual value.

This systematically biases the illuminant estimates to be too red. The actual amount by which the red chromaticity is increased is a function of pixel brightness and is not necessarily 10%. The new r chromaticity is given by:

$$r = 1.1 \cdot R / (1.1 \cdot R + G + B) \qquad (5)$$

Since the purpose of this test is to test if committees can eliminate systematic errors, we assumed that WP algorithm is not affected by this color shift.

Figure 3 shows the performance of two committees, one employing a simple average and one using a LMS weighted average.

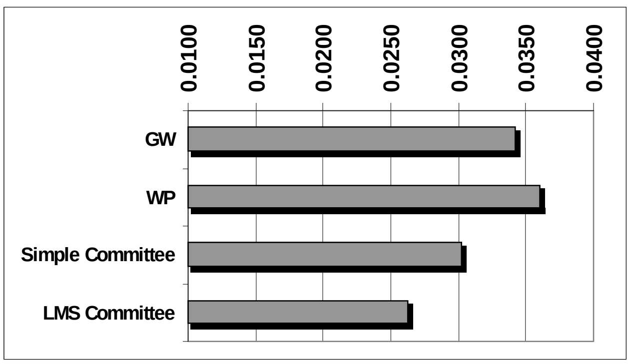

Figure 3. RMS error of individual algorithms and committees. The systematic errors induced in the GW algorithm do not affect the performance of the LMS committee

The estimation errors of the GW algorithm are larger due to the systematic estimation error induced by the color shift described above. However, the LMS model compensates for the systematic error and yields the same performance as the model shown in Figure 2.

Conclusion

We have shown that committee models, which combine the results of two or more color constancy methods, can significantly improve overall color constancy performance. The implementation of these models is simple and the computational overhead is very small. Thus, committees provide a useful tool for improved color constancy.

Acknowledgements

The authors gratefully acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada and Hewlett Packard Incorporated.

References

1. G. Finlayson, M. Drew, and B. Funt, “Color Constancy: Generalized Diagonal Transforms Suffice,” *J. Opt. Soc. Am. A*, 11(11), pg. 3011-3020. (1994).
2. C.M. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford. (1995).
3. E.H. Land, “The Retinex Theory of Color Vision,” Scientific American, pg. 108-129. (1977).
4. G. Buchsbaum, “A Spatial Processor Model for Object Colour Perception,” J. Franklin Institute, 310 (1), pg. 1-26. (1980).
5. B. Funt, V. Cardei, and K. Barnard, Learning Color Constancy, *Proc IS&T/SID 4th Color Imaging Conf*., pg. 58-60. (1996).
6. V. Cardei, B. Funt, and K. Barnard, Adaptive Illuminant Estimation Using Neural Networks, *Proc. 8th Int. Conf. on Artificial Neural Networks*, pg. 749-754. (1998).