Evolving Color Constancy for an Artificial Retina

Marc Ebner

Universität Würzburg, Lehrstuhl für Informatik II,
Am Hubland, 97074 Würzburg, Germany
ebner@informatik.uni-wuerzburg.de
http://www2.informatik.uni-wuerzburg.de/staff/ebner/welcome.html

Abstract Objects retain their color in spite of changes in the wavelength and energy composition of the light they reflect. This phenomenon is called color constancy and plays an important role in computer vision research. We have used genetic programming to automatically search the space of programs to solve the problem of color constancy for an artificial retina. This retina consists of a two dimensional array of elements each capable of exchanging information with its adjacent neighbors. The task of the program is to compute the intensities of the light illuminating the scene. These intensities are then used to calculate the reflectances of the object. Randomly generated color Mondrians were used as fitness cases. The evolved program was tested on artificial Mondrians and natural images.

1 Introduction

The human visual system is able to correctly perceive the color of objects irrespective of the light which illuminates the scene. That is, the leaves of a tree still look green to a human observer even if the tree is illuminated with red light and the leaves actually reflect more red than green light. The task of computing color constant descriptors for an image is known as the problem of color constancy. One is able to somehow discount the illuminant and extract a measure of the object's reflectance properties [26]. The same ability would also be useful for a robot which has to work under different lighting conditions. To date, the lighting still has to be carefully controlled such that the algorithms continue to work. The problem of color constancy is also of particular importance for the task of object recognition [7,10].

Numerous solutions to the problem of color constancy have been proposed, i.e. Land's retinex theory [21], variants of the retinex theory [2,12,18], gamut-constraint methods [1,8,9] recovery of basis function coefficients [13,14,17,23], mechanisms of light adaptation coupled with eye movements [6], neural networks [4,5,11,16,24], minimization of an energy function [25], comprehensive color normalization [7] or committee-based methods which combine the output of several different color constancy algorithms [3]. We now summarize some background on color image formation and discuss the two most widely known algorithms for color constancy: white-patch and the gray world assumption.

The response of a sensor at position $\mathbf{x}_s$ measuring the light reflected from a Lambertian surface at position $\mathbf{x}_o$ is given by

$$\mathbf{I}(\mathbf{x}_s) = \mathbf{n}_l \cdot \mathbf{n}_o \int_\omega R(\lambda, \mathbf{x}_o) L(\lambda) \mathbf{S}(\lambda) d\lambda$$

where $\mathbf{I}(\mathbf{x}_s)$ is a vector of sensor responses, $\mathbf{n}_l$ is the unit vector pointing in the direction of the light source, $\mathbf{n}_o$ is the unit vector corresponding to the surface normal, $R(\lambda, \mathbf{x}_o)$ specifies the percentage of light with wavelength λ reflected by the surface at position $\mathbf{x}_o$, $L(\lambda)$ is the intensity of light hitting the surface and $\mathbf{S}(\lambda)$ specifies the sensor's response functions [7]. The integration is over all wavelengths to which the sensors respond. Assuming ideal sensors for red, green and blue light ($S_i = \delta(\lambda - \lambda_i)$) with $i \in \{r, g, b\}$ and a light source which illuminates the surface at a right angle the equation simplifies to

$$I_i(\mathbf{x}_s) = R(\lambda_i, \mathbf{x}_o) L(\lambda_i)$$

where $I_i(\mathbf{x}_s)$ denotes the i-th component of the vector $\mathbf{I}(\mathbf{x}_s)$.

In this case the light illuminating the scene simply scales the reflectances. If there exists at least one pixel for each band which reflects all light for this particular band, one could simply rescale all color bands to the range $[0, 1]$.

$$R(\lambda_i, \mathbf{x}_o) = \frac{I_i(\mathbf{x}_s)}{L_{\max}(\lambda_i)}$$

with $L_{\max}(\lambda_i) = \max_{\mathbf{x}}\{I_i(\mathbf{x})\}$. This algorithm is called the white-patch retinex algorithm [10].

Another possibility would be to calculate space average color of the image and use this information to estimate the intensities of the light illuminating the scene. If one assumes that the reflectances of the surface are uniformly distributed over the interval $[0, 1]$, one gets

$$\frac{1}{N}\sum_{\mathbf{x}}^{N} I_i(\mathbf{x}) = \frac{1}{N}\sum_{\mathbf{x}}^{N} R(\lambda_i, \mathbf{x}) L(\lambda_i) = L(\lambda_i)\frac{1}{N}\sum_{\mathbf{x}}^{N} R(\lambda_i, \mathbf{x}) = L(\lambda_i)\frac{1}{2}$$

This is the so called gray world assumption. Thus, for a sufficiently complex image one can estimate the intensities of the light illuminating the scene as twice the space average color.

$$L(\lambda_i) = \frac{2}{N}\sum_{\mathbf{x}}^{N} I_i(\mathbf{x})$$

Both cues, space-average scene color as well as the color of the highest luminance patch are used by the human visual system to estimate the color of the light illuminating the scene [22].

Most work on the problem of color constancy has focused on finding an analytical solution. Some research has tried to learn the problem of color constancy using a neural network, e.g. [11,24]. We try to *evolve* the ability of color constancy for an artificial retina. In particular, we want to address the following questions. Is it possible to evolve a function for an artificial retina which estimates the intensities of the light illuminating the scene using a training set which only consists of randomly generated color Mondrians? The function is constrained to obtain and exchange information only locally but not globally. This property is very important in order for the retina to be scalable to arbitrary sizes. Will the results generalize to natural images? We now describe the architecture of the artificial retina.

2 An artificial retina

Our artificial retina consists of a two-dimensional array of elements. Each element is able to exchange information with its neighbors to the left and right as well as its neighbors above and below. The elements receive an image as input and their task is to compute the reflectances of the objects shown in the image. Each element may exchange information only locally in order to calculate the intensities of the light illuminating the scene. Thus, the artificial retina is scalable to any size.

For our experiment we assume that the viewed image is generated by multiplying the red, green and blue components of the pixel values with the intensities of the light illuminating the scene, that is,

$$\mathbf{p}_{\mathrm{v}}(x,y) = (p_{\mathrm{r},r}(x,y) \cdot l_r, p_{\mathrm{r},g}(x,y) \cdot l_g, p_{\mathrm{r},b}(x,y) \cdot l_b)$$

where $\mathbf{p}_{\mathrm{v}}(x,y)$ is the vector with color components of the pixel as it is perceived by the artificial retina, $\mathbf{p}_{\mathrm{r}}(x,y) = (p_{\mathrm{r},r}(x,y), p_{\mathrm{r},g}(x,y), p_{\mathrm{r},b}(x,y))$ is the vector with reflectances, $\mathbf{l} = (l_r, l_g, l_b)$ is the vector which describes the intensities of red, green and blue light illuminating the scene.

The output $\mathbf{p}_{\mathrm{o}}(x,y)$ of the element at position (x,y) is defined as

$$p_{\mathrm{o},i}(x,y) = \begin{cases} \frac{p_{\mathrm{v},i}(x,y)}{l_{\mathrm{e},i}(x,y)} & \text{if } l_{\mathrm{e},i}(x,y) > 0.001 \\ 1 & \text{otherwise} \end{cases}$$

where $i \in \{r,g,b\}$. Thus each element consists of three sub-elements, one for each color band. The output of each sub-element as well as the estimate of the intensities of the ambient light are restricted to the range $[0,1]$. The vector $\mathbf{p}_{\mathrm{v}}(x,y)$ is available to the element at position (x,y). The intensities of red, green and blue light $\mathbf{l}_{\mathrm{e}}(x,y)$ are calculated by a program which is evolved using genetic programming [19,20]. All elements are running the same program. The calculated intensities are stored locally as well as distributed to the neighboring elements.

Each element has the structure shown in Figure 1. It has access to the intensities estimated by the neighboring elements (left, right, up, and down) as

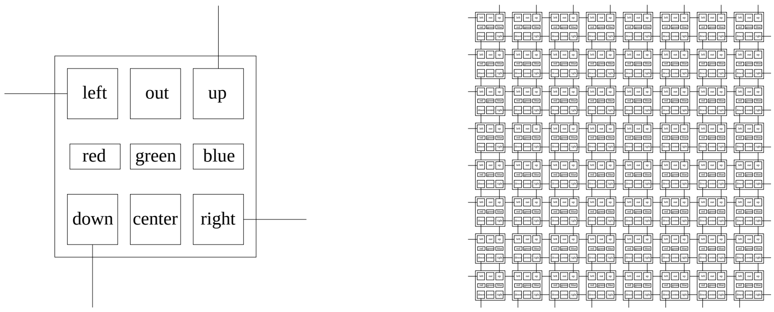

Figure 1. A single element of the artificial retina is shown on the left. Current knowledge of light illuminating the scene is stored inside the element (center) and is also distributed to the left, right, up and down. This knowledge is continually updated. Each element has access to the red, green and blue intensities of the viewed image. Using its knowledge about the light illuminating the scene, each element calculates the reflectances of its pixel (out).Each element only exchanges information locally, thus the individual elements may be combined easily to form a large $n \times n$ array (shown on the right).

well as to its own estimate of the light illuminating the scene. In addition to this information, it has access to the red, green and blue intensities of the viewed image. The intensities of red, green and blue light are calculated by iterating the following update equations:

$$\begin{aligned}
\mathrm{center}_i(x,y,t) &= l_{\mathrm{e},i}(x,y,t-1) \\
\mathrm{left}_i(x,y,t) &= l_{\mathrm{e},i}(x-1,y,t-1) \\
\mathrm{right}_i(x,y,t) &= l_{\mathrm{e},i}(x+1,y,t-1) \\
\mathrm{up}_i(x,y,t) &= l_{\mathrm{e},i}(x,y+1,t-1) \\
\mathrm{down}_i(x,y,t) &= l_{\mathrm{e},i}(x,y-1,t-1) \\
l_{\mathrm{e},i}(x,y,t) &= \mathrm{program}(\mathrm{center}_i(x,y,t), \mathrm{left}_i(x,y,t), \mathrm{right}_i(x,y,t), \\
&\qquad \mathrm{up}_i(x,y,t), \mathrm{down}_i(x,y,t), \mathbf{p}_v(x,y), p_{\mathrm{v},i}(x,y))
\end{aligned}$$

The intensities of the viewed image were used as an initial estimate of the light illuminating the scene.

$$l_{\mathrm{e},i}(x,y,0) = p_{\mathrm{v},i}(x,y)$$

Our task is to find a program which calculates the red, green and blue intensities of the light illuminating the scene. To search the space of possible programs we have used genetic programming [19,20]. The function set consists of the binary arithmetic functions addition (`+`), subtraction (`-`), multiplication (`*`) and protected division (`/`), the unary functions multiply by 2 (`mul2`) and divide by 2 (`div2`). The set of terminal symbols consists of the constant one (`1`), red (`red`), green (`green`), and blue (`blue`) color channel, the sub-element's color channel (`band`), the element's current estimate of the illuminant (`center`) as well as the estimates of the illuminant calculated by the elements to the left (`left`), right (`right`), above (`up`) and below (`down`).

3 Experiments

A population of 1000 individuals was evolved for 50 generations. We did ten runs with different initial seeds for the random number generator. 1% of the next generation was filled with the best individual, 9% of the next generation was filled by applying the reproduction operator, 70% of the next generation was filled by applying the crossover operator and the remainder was filled by applying the mutation operator. We used tournament selection with size 7 to select individuals.

For each generation we randomly generated three Mondrians and evaluated the performance of the individuals on the different Mondrians. Each Mondrian was illuminated with random intensities for the red, green and blue components. The size of the Mondrians was 64x64 pixels with circular boundary conditions. The Mondrians were created by filling the background with a random color and then placing 64 filled rectangles with random colors on top of each other. The size of the rectangles was selected randomly in the range [8,24]. The position of the rectangles was also selected randomly.

Each evaluation consisted of iterating the update equations 100 times. For each Mondrian m we calculated the difference d_m between the output of the artificial retina and the known reflectances.

$$d_m = \frac{1}{3 \cdot \text{width} \cdot \text{height}} \sum_{x=1}^{\text{width}} \sum_{y=1}^{\text{height}} \sqrt{(\mathbf{p}_{\mathrm{o}}(x,y) - \mathbf{p}_{\mathrm{r}}(x,y))^2}$$

The largest difference over all Mondrians is used to calculate the fitness of the individual. Fitness is defined as

$$\text{fitness} = \frac{1}{1 + max_m\{d_m\}}.$$

The worst performance of an individual was used to calculate its fitness in order to reward generalists. Note that fitness is to be maximized and the error is to be minimized. Harvey et al. [15] have chosen a similar approach to evolve robust control algorithms for a simulated robot.

We also tried averaging the errors of several fitness cases instead of using the maximum error to determine fitness. However, if the number of fitness cases is large, then on average the red, green and blue intensities are 0.5 and a program which simply outputs $2\mathbf{p}_{\mathrm{v}}$ might seem like a good solution. Obviously, this is not a solution we are looking for. Therefore we chose to use the maximum error over several fitness cases to calculate the fitness of an individual. In this case, an individual has to produce very good results for all fitness cases in order to survive more than one generation. If an individual performs badly on even a single fitness case it is likely that this individual will be eliminated from the population. The individual with the highest fitness value at the last generation is our result.

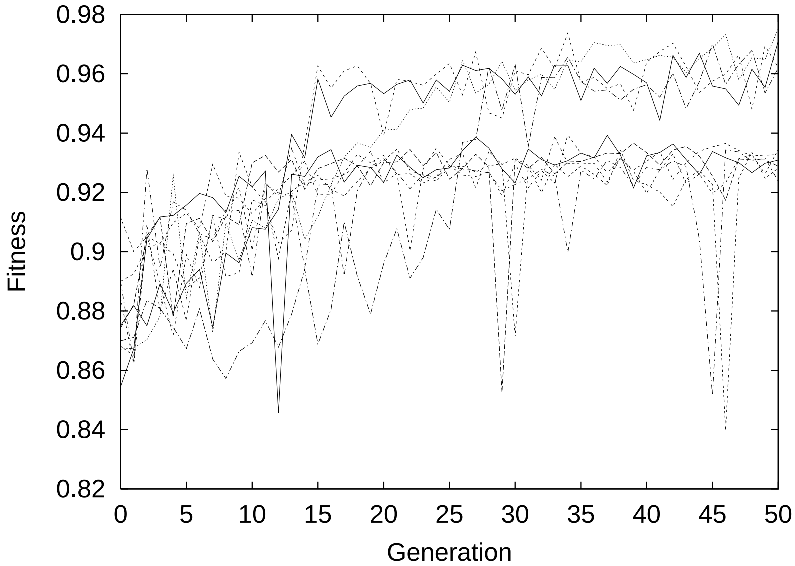

Figure 2. Ten runs were executed with different seeds for the random number generator. At the end of 50 generations, 6 of the runs were stuck in a local optimum. The best individual of all ten runs was analyzed in detail.

We performed 10 runs with different seeds for the random number generator. Figure 2 shows the maximum fitness for each of the runs. 6 of the runs were stuck in a local optimum. We now analyze the best individual of those 10 runs in detail.

4 Results

Run number 4 had the highest fitness value at generation 50. Figure 3 shows the output of the best individual found after 50 generations for three different fitness cases[1]. Each row shows the results for one fitness case. The first image in each row shows the original image. The second image shows the light illuminating

[1] The images can be viewed in color on the authors web page http://www2.informatik.uni-wuerzburg.de/staff/ebner/research/color/color.html

the scene. The third image shows the image viewed by the camera. The fourth image shows the estimated intensities and the fifth image shows the estimated reflectances. The evolved individual uses mostly addition and division by 2 to solve the task of color constancy.

The evolved individual uses the following code
(div2 (+ (+ (div2 (+ (div2 (+ (div2 down) (div2 (+ (div2 (+ (div2 (+ band (div2 down))) (div2 (+ band (div2 (+ band (div2 center))))))) (div2 down))))) (* (div2 down) (/ right (div2 (+ down (+ (div2 down) (div2 (+ (div2 (+ band (div2 down))) (div2 (+ band (div2 (+ band (div2 center))))))))))))) (div2 (+ (div2 (+ (div2 right) (div2 (+ band (div2 down))))) (* (div2 down) (/ right (div2 (+ down (+ (div2 down) (div2 (+ (div2 (+ band (div2 center))) (div2 (+ band band)))))))))))) down))
to estimate the intensities of the light illuminating the scene and thereby estimate the reflectances of the viewed object ($d_1 = 0.0242$, $d_2 = 0.0343$, $d_3 = 0.0334$, after 100 updates).

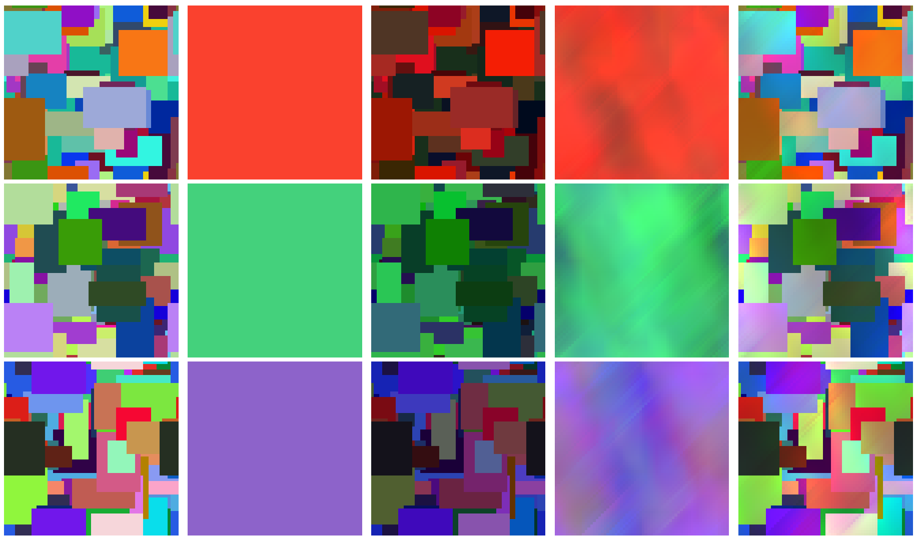

Figure 3. Results of one individual for three different fitness cases. The first image of each row shows the original Mondrian. The second image shows the light illuminating the scene. The third image shows the viewed image. The fourth image shows the estimated light intensities. The fifth image shows the reflectances which were extracted from the viewed image by the artificial retina.

The same individual was also tested on natural images. The results of these tests are shown in Figure 4. The images on the left show the results for an outdoor photograph transformed in the same way as the artificially created Mondrians. That is, we simply multiplied the pixel values with a randomly chosen color.

The first image shows the original photograph. The second image shows the color illuminating the scene. The third image shows the colored photograph. The fourth image shows color estimated by the artificial retina. The fifth image shows the reflectances estimated by the artificial retina. Again the individual is able to restore the original colors (d=0.0246, after 300 updates, 190×128 image). Thus, we have shown that although the individual was only trained on color Mondrians it was able to estimate the reflectances of a color adjusted photograph.

The images on the right show the results for a test image created by Funt et al. [10]. The original image as well as the transformed image are part of a larger set for the problem of color constancy. The original image was taken with a camera under a uniform light model, the third image shows the same object viewed with colored light. We calculated the second image by dividing the pixel values of the third image by the pixel values of the first image to extract the color illuminating the scene. Note that in this case the assumption of a flat image does not hold. The fourth image shows the color estimated by the artificial retina and the fifth image is the output of our artificial retina (d=0.098, after 300 updates, 168×160 image). In this case the assumptions used during evolution, e.g. that the viewed object is flat, do not hold and the results are not as good as for the photograph. However the results show, that even in this case, the evolved individual is able to improve the appearance of the input image.

In order to compare the results we also applied the white-patch retinex algorithm and the algorithm using the gray world assumption to both images. The results for the white-patch retinex algorithm are shown in Figure 5 (d=0.032 for the photograph and d=0.063 for the ball) and the results for the algorithm using the gray world assumption are shown in Figure 6 (d=0.038 for the photograph and d=0.162 for the ball). The results for the photograph are comparable for all three algorithms. The images produced by white-patch retinex and gray world are a bit brighter than the output image of the evolved algorithm. In terms of the calculated error, the white-patch algorithm produced the best results for the ball image. The results for the ball seem to suggest that the evolved algorithm applies a mixed strategy.

5 Conclusion

We have evolved an algorithm to solve the problem of color constancy for an artificial retina using genetic programming. The retina was designed such that information is only exchanged locally, not globally. The artificial retina is thus scalable to arbitrary sizes. The individuals were trained on artificially created Mondrian images. The best program was tested on additional color Mondrians as well as natural images. Although the individual was only trained on Mondrian images it was also able to estimate the intensities of the light illuminating the scene for natural images and thereby estimate the reflectances of the viewed object.

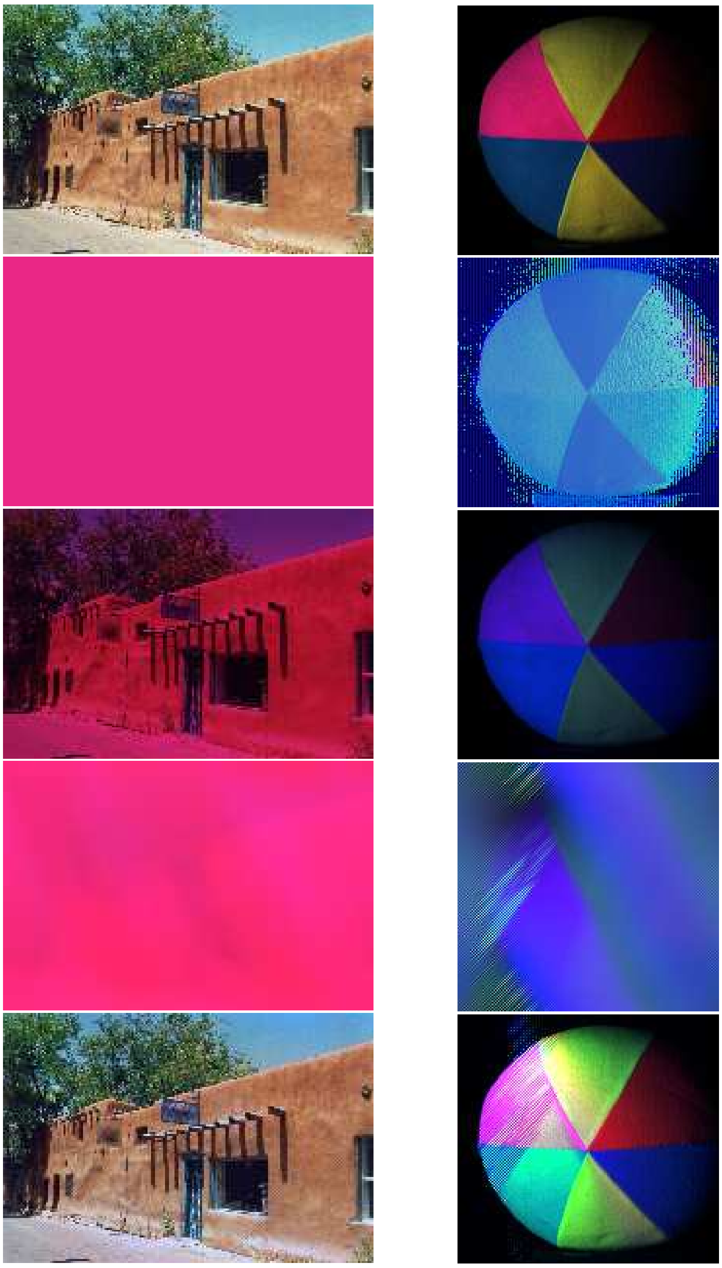

Figure 4. Results of the evolved program for the artificial retina on two natural images. The images on the left show the results for an outdoor photograph. The images on the right show the results for an image which was created by Funt et al. [10] to test color constancy algorithms.

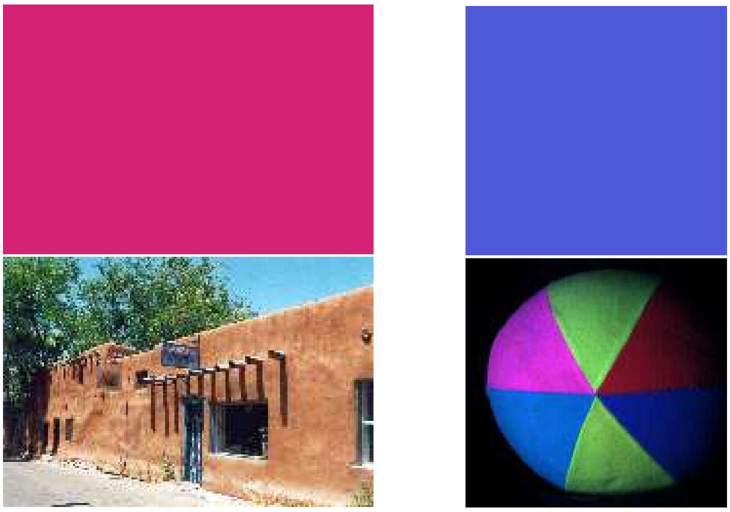

Figure 5. Results for the white-patch retinex algorithm. The first row of images shows the estimated color for the photograph and the image of the ball . The second row shows the output of the white-patch retinex algorithm for both images.

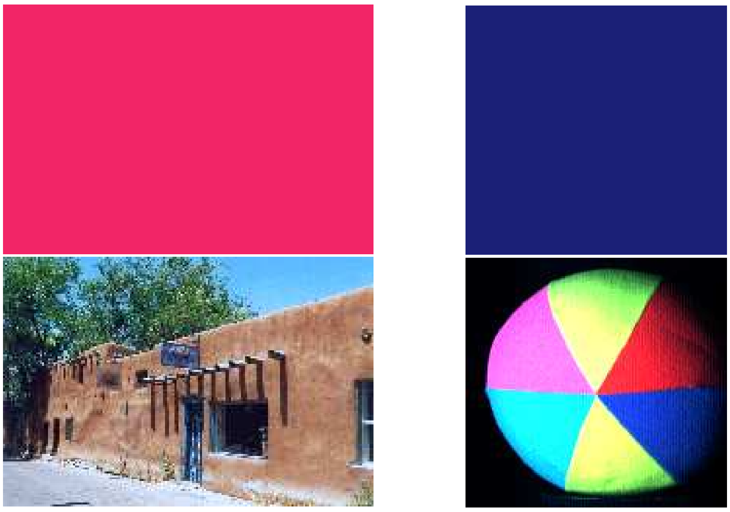

Figure 6. Results for the gray-world assumption. The first row of images shows the estimated color for the photograph and the image of the ball. The second row shows the output of the gray-world algorithm for both images.

Acknowledgements

We have used the lilgp Programming System Vers. 1.1 [27] for our experiments.

References

1. K. Barnard, G. Finlayson, and B. Funt. Color constancy for scenes with varying illumination. *Computer Vision and Image Understanding*, 65(2):311–321, February 1997.
2. D. H. Brainard and B. A. Wandell. Analysis of the retinex theory of color vision. In G. E. Healey, S. A. Shafer, and L. B. Wolff, editors, *Color*, pages 208–218, Boston, 1992. Jones and Bartlett Publishers.
3. V. C. Cardei and B. Funt. Committee-based color constancy. In *Proceedings of the IS&T/SID Seventh Color Imaging Conference: Color Science, Systems and Applications*, pages 311–313, 1999.
4. S. M. Courtney, L. H. Finkel, and G. Buchsbaum. A multistage neural network for color constancy and color induction. *IEEE Transactions on Neural Networks*, 6(4):972–985, July 1995.
5. P. A. Dufort and C. J. Lumsden. Color categorization and color constancy in a neural network model of V4. *Biological Cybernetics*, 65:293–303, 1991.
6. M. D'Zmura and P. Lennie. Mechanisms of color constancy. In G. E. Healey, S. A. Shafer, and L. B. Wolff, editors, *Color*, pages 224–234, Boston, 1992. Jones and Bartlett Publishers.
7. G. D. Finlayson, B. Schiele, and J. L. Crowley. Comprehensive colour image normalization. In *Fifth European Conference on Computer Vision (ECCV '98)*, 1998.
8. D. A. Forsyth. A novel approach to colour constancy. In *Second International Conference on Computer Vision (Tampa, FL, Dec. 5-8)*, pages 9–18. IEEE Press, 1988.
9. D. A. Forsyth. A novel algorithm for color constancy. In G. E. Healey, S. A. Shafer, and L. B. Wolff, editors, *Color*, pages 241–271, Boston, 1992. Jones and Bartlett Publishers.
10. B. Funt, K. Barnard, and L. Martin. Is colour constancy good enough? In *Fifth European Conference on Computer Vision (ECCV '98)*, pages 445–459, 1998.
11. B. Funt, V. Cardei, and K. Barnard. Learning color constancy. In *Proceedings of the IS&T/SID Fourth Color Imaging Conference*, pages 58–60, Scottsdale, 19-22November 1996.
12. B. V. Funt and M. S. Drew. Color constancy computation in near-mondrian scenes using a finite dimensional linear model. In *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition*, pages 544–549. Computer Society Press, 5-9 June 1988.
13. B. V. Funt, M. S. Drew, and J. Ho. Color constancy from mutual reflection. *International Journal of Computer Vision*, 6(1):5–24, 1991.
14. R. Gershon, A. D. Jepson, and J. K. Tsotsos. From [R,G,B] to surface reflectance: Computing color constant descriptors in images. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, volume 2, pages 755–758, 1987.
15. I. Harvey, P. Husbands, and D. Cliff. Issues in evolutionary robotics. In J.-A. Meyer, H. L. Roitblat, and S. W. Wilson, editors, *From animals to animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior, Honolulu, Hawaii, 1992*, pages 364–373. The MIT Press, 1993.

16. J. Herault. A model of colour processing in the retina of vertebrates: From photoreceptors to colour opposition and colour constancy phenomena. *Neurocomputing*, 12:113–129, 1996.
17. J. Ho, B. V. Funt, and M. S. Drew. Separating a color signal into illumination and surface reflectance components: Theory and applications. In G. E. Healey, S. A. Shafer, and L. B. Wolff, editors, *Color*, pages 272–283, Boston, 1992. Jones and Bartlett Publishers.
18. B. K. P. Horn. *Robot Vision*. The MIT Press, Cambridge, Massachusetts, 1986.
19. J. R. Koza. *Genetic Programming, On the Programming of Computers by Means of Natural Selection*. The MIT Press, Cambridge, Massachusetts, 1992.
20. J. R. Koza. *Genetic Programming II, Automatic Discovery of Reusable Programs*. The MIT Press, Cambridge, Massachusetts, 1994.
21. E. H. Land. The retinex theory of colour vision. *Proc. Royal Inst. Great Britain*, 47:23–58, 1974.
22. K. J. Linnell and D. H. Foster. Space-average scene colour used to extract illuminant information. In C. Dickinson, I. Murray, and D. Carden, editors, *John Dalton's Colour Vision Legacy. Selected Proceedings of the International Conference*, pages 501–509, London, 1997. Taylor & Francis.
23. L. T. Maloney and B. A. Wandell. Color constancy: a method for recovering surface spectral reflectance. *Journal of the Optical Society of America A3*, 3(1):29–33, January 1986.
24. C. L. Novak and S. A. Shafer. Supervised color constancy for machine vision. In G. E. Healey, S. A. Shafer, and L. B. Wolff, editors, *Color*, pages 284–299, Boston, 1992. Jones and Bartlett Publishers.
25. S. Usui and S. Nakauchi. A neurocomputational model for colour constancy. In C. Dickinson, I. Murray, and D. Carden, editors, *John Dalton's Colour Vision Legacy. Selected Proceedings of the International Conference*, pages 475–482, London, 1997. Taylor & Francis.
26. S. Zeki. *A Vision of the Brain*. Blackwell Science, Oxford, 1993.
27. D. Zongker and B. Punch. *lil-gp 1.01 User's Manual (support and enhancements Bill Rand)*. Michigan State University, 1996.