Tuning Retinex Parameters

Brian Funt[*a], Florian Ciurea[**a] and John McCann[***b]

[a]School of Computing Science, Simon Fraser University, British Columbia, Canada, V5A 1S6
[b]McCann Imaging, Belmont, Massachusetts, USA, 02478

ABSTRACT

Our goal is to specify the retinex model as precisely as possible. The core retinex computation is clearly specified in the recent MATLAB implementation[2]; however, there remain several free parameters which introduce significant variability into the model's predictions. In this paper, we extend previous work[1] on specifying these parameters. In particular, instead of looking for fixed values for the parameters, we establish methods which automatically determine values for them based on the input image. These methods are tested on the McCann-McKee-Taylor asymmetric matching data[5] along with some previously unpublished data that include simultaneous contrast targets.

Keywords: Retinex, color vision, human vision, color constancy.

INTRODUCTION

The MATLAB implementation[2] of the retinex model has three important input parameters: the number of iterations the algorithm performs at each level of its multi-level computation, the "post-lut" output lookup table function, and the input image size. The model's final output depends strongly on the values chosen for the first two of these parameters but not the third[1].

The retinex model aims to predict the sensory response of lightness. In previous work by two of us (Funt and Ciurea)[1], we established values for the parameters based upon fitting the model's predictions to the data originally described over 35 years ago by McCann, McKee and Taylor[5]. This fit led to the conclusion that 33 iterations had the lowest global average of the differences between observer data and computed values, assuming that the number of iterations was constant for all levels of the multi-resolution computation. However, McCann felt that 33 was too high a number, and would not lead to a good model of simultaneous contrast. Hence, together we began the current series of experiments by including previously unpublished data from lightness matching experiments with simultaneous contrast targets. We also added other unpublished data for two other target types: one contained a fixed set of patches of various shades of gray appear on a background which varied from black to gray to white; the other contained a staircase of shades of gray under illumination with rising, falling or constant intensity.

For the simultaneous contrast data, we indeed did find that a much smaller value is required for the iteration parameter in order to make a good fit. However, we could no longer find a single value for the number of iterations that simultaneously would minimized the error for the combined data from the MMT (McCann-McKee-Taylor), GB (fixed scale of grays on different backgrounds), SC (simultaneous contrast) and IG (illumination gradient on staircase) experiments. This led us to consider a method of automatically calculating how many iterations to use based on how the computation was proceeding. In addition, the post-lut processing needs to change as a function of the number of iterations, so this led us to a method of automatically calculating the appropriate post-lut.

NUMBER OF ITERATIONS

The two MATLAB implementations of retinex in Funt et. al.[2] are referred to as McCann99 retinex and Frankle-McCann retinex. For brevity, we concentrate here only on McCann99 retinex, but the results are similar for both versions. McCann99 retinex creates a multi-resolution pyramid from the input by averaging image data. It begins the pixel comparisons at the most highly averaged, or top level of the pyramid. After computing so called New Products (precursors to the final lightness estimates) on the image at a reduced resolution, the resulting New Product values are propagated down, by pixel replication, to the pyramid's next level as initial estimates at that level. Further pixel comparisons refine the estimates at the higher resolution level and then those new estimates are again propagated down a level in the pyramid. This process continues until values have been computed for the pyramid's bottom level.

At each level, the basic step is the comparison of each pixel to each of its immediate neighbors. The number of iterations refers to the number of times all the immediate neighbors are cycled through before moving down to the next level in the pyramid. Since pixels are only directly compared to immediate neighbors, comparisons to more distant pixels at the current pyramid level are only made implicitly by propagation of information from pixel to pixel during these iterations. Hence, increasing the number of iterations increases the spatial distance across which pixels are related during the computation. McCann99 retinex uses the same number of iterations at all levels and so there is only a single iteration parameter to specify and have limited this paper to considering a single value for all levels.

POST-LUT PROCESSING

Postlut processing refers to applying a function f uniformly to every image pixel, $I(x,y)=f(I(x,y))$, for all image locations (x,y) immediately after the main retinex computation. The term "postlut" derives from historical use of image processing hardware using a lookup table (lut) as a final post-processing step. Post-lut processing is important in bringing the final result into the appropriate dynamic range, compensating for differences in overall illumination intensity between test targets, and in converting to the coordinates of Munsell Value scale used in recording the experimental data. Although all these factors can be thought of separately, they are all eventually combined into a single post-lut function.

The first post-lut step adjusts the dynamic range. Retinex output from the pyramidal spatial comparison stage, falls in the [0,1] range. Because the value 1 represents 'white' and retinex assumes there is at least one white pixel in every image, the value 1 necessarily arises in the output. However, the lowest output value depends on the image content and varies with the number of iterations used. The fewer the iterations, the more local the spatial comparisons will be, and therefore, the less the likelihood of big intensity differences being found. As a result, the fewer the iterations, the higher the minimum retinex output value (Figure 1 in Funt et. al.[2] illustrates this effect). The first purpose of the post-lut is to stretch the retinex output to a reasonable range. Since the amount of stretching needed depends on the number of iterations, and we vary the number of iterations in our experiments, we decided to always linearly scale the retinex output to the full [0,1] range. This stretch does not correct for the fact that the number of iterations performs a non-linear compression of the image. The post-lut is not fixed, but rather depends on the input image and number of iterations used. This decision effectively means that we are assuming that there is at least on black location in the test target. While this assumption need not be true for scenes in general and could lead to errors in retinex predictions, it is true for all the test targets subjects viewed.

After scaling to the [0,1] range, the post-lut then converts the retinex output values, r, to the lightness scale used for recording subject's matches. For the MMT data set, the conversion is to Munsell Value scale V using[3]:

$$V = 2.539\, r^{1/3} - 1.838 \text{ for } r > 0.384$$

For the SC, GB and IG data sets, the conversion is to a lightness scale described by Stiehl et. al.[6]. Based on a fit to the raw data, we use the following function to convert the log luminance to the lightness scale values, L:

$$L = 129.6\, r^{1/100} - 132.45$$

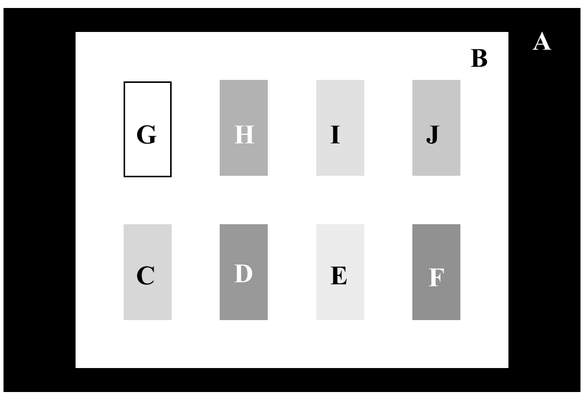
The final post-lut component compensates for differences in overall illumination intensity between the test and match conditions. Only the MMT experiments involved such intensity differences. The compensation is based on data from Figure 8 of McCann, Land and Tatnall[4]. Generally, the effect of this correction is slight. Details can be found in Funt and Ciurea[1].

LIGHTNESS MATCHING DATA

The data for the MMT matching experiments was reported a long time ago[4]. The 'new' data we report here is based on experiments by McCann which were also conducted earlier, but not previously reported in the literature. These experiments involve transparent grayscale targets lit from behind with uniform illumination. Subjects were asked to report the lightness of each patch in the target display using a standard lightness transparency display as a reference. The standard lightness display consists of 25 squares of different lightnesses against a white surround. The squares are arranged in a serpentine path such that the change in lightness from any of the 25 squares to the next is constant[6]. In the resulting lightness scale, 1.0 corresponds to an opaque area and 9.0 to the brightest area. The experiments were based on 4 to 7 subjects, which each subject repeating the matches on 3 different occasions.

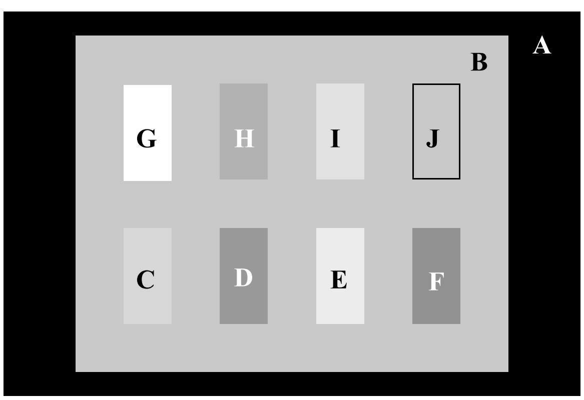
The matching procedure was set up such that in the normal viewing position, the subject saw the test display as the only thing in the field of view. By turning 90 degrees to the right, the subject would see instead the standard lightness display as the only thing in the field of view. Subjects were allowed to look back and forth between viewing the test display and the standard display as many times as desired without a time constraint[4]. The test display and the standard lightness display had the same level of luminance.

Figures 1-9 illustrate the targets along with the corresponding luminance, pixel value for each patch as input to the retinex algorithm, and average observed lightness reported for each patch. With the exception of the IG targets, all the patches have uniform luminance. For the IG targets, the five patches present a gradient of luminance specified by the left and right edge. It should be noted that the figures are intended only to illustrate the corresponding targets. They are not accurate reproductions of the targets. Their printed appearance is not the same as under the controlled experimental conditions.



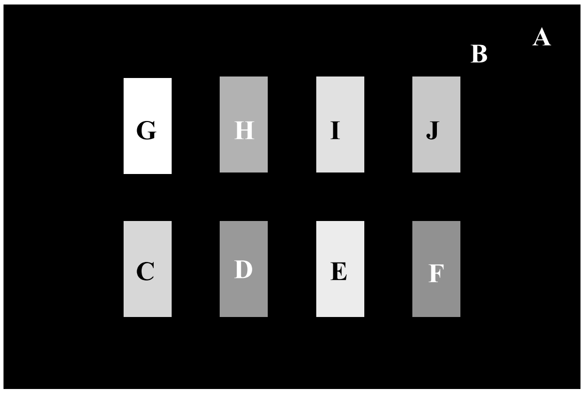
Patch	Luminance	Pixel value	Observed lightness	Standard deviation
G	1001	255	8.75	0.15
E	595	236	7.55	0.20
I	439	225	6.25	0.25
C	336	215	5.94	0.31
J	228	200	5.19	0.19
H	125	178	4.36	0.26
D	63	153	3.37	0.50
F	50	145	2.80	0.30
B	1001	255	8.80	0.20
A	1	0	1.0	0

Figure 1: Grays on White



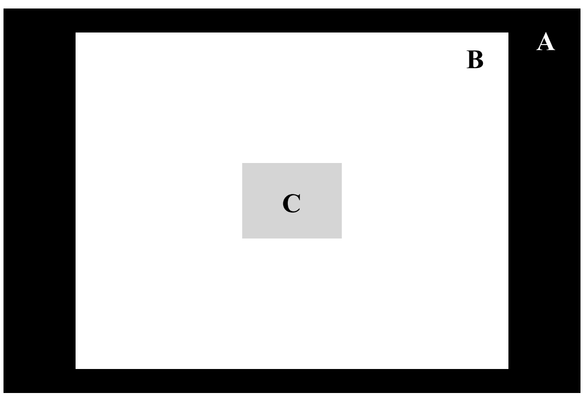
Patch	Luminance	Pixel value	Observed lightness	Standard deviation
G	1001	255	9.00	0.20
E	595	236	8.50	0.50
I	439	225	7.31	0.31
C	336	215	7.06	0.31
J	228	200	5.88	0.38
H	125	178	4.98	0.28
D	63	153	4.08	0.48
F	50	145	3.05	0.55
B	228	200	5.75	0.25
A	1	0	1.0	0

Figure 2: Grays on Gray



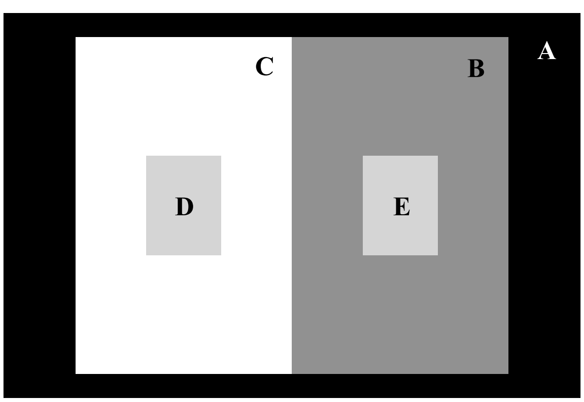
Patch	Luminance	Pixel value	Observed lightness	Standard deviation
G	1001	255	9.00	0.05
E	595	236	8.55	0.45
I	439	225	7.53	0.78
C	336	215	7.29	0.54
J	228	200	6.80	0.50
H	125	178	5.65	0.35
D	63	153	5.20	0.50
F	50	145	4.68	0.38
B	1	0	0.90	0.10
A	1	0	0.83	0.08

Figure 3: Grays on Black



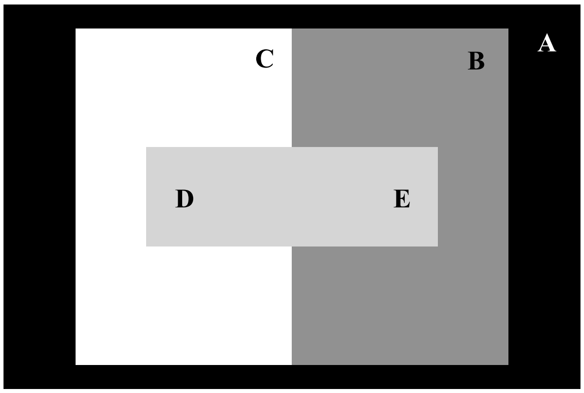
Patch	Luminance	Pixel value	Observed lightness	Standard deviation
B	1001	255	9.03	0.23
C	321	213	6.15	0.52
A	1	1	1.13	0.13

Figure 4: Simultaneous Contrast ‘Single’



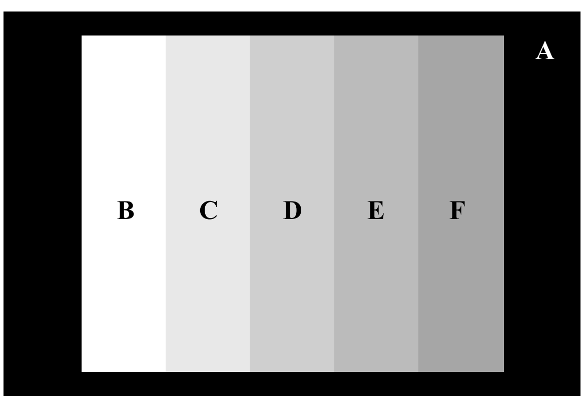
Patch	Luminance	Pixel value	Observed lightness	Standard deviation
C	1001	255	8.85	0.18
D	321	213	6.16	0.4
E	321	213	6.95	0.45
B	50	145	3.95	0.45
A	1	1	1.08	0.27

Figure 5: Simultaneous Contrast ‘Double’



Patch	Luminance	Pixel value	Observed lightness	Standard deviation
C	1001	255	9.09	0.29
D	321	213	6.2	0.5
E	321	213	7.69	0.58
B	50	145	4.04	0.54
A	1	1	1	0.25

Figure 6: Simultaneous Contrast 'Strip'



Patch	Luminance	Pixel value	Observed lightness
A	1	1	1.02
B	1001	255	9.1
			9
			90
			8.9
C	840.8	249	7.8
			7.9
			7.9
			8
D	706.3	242	7.2
			7.4
			7.4
			7.6
E	593.3	236	6.8
			7
			7
			7.2
F	498.4	229	6.5
			6.6
			6.6
			6.8

Figure 7: Staircase under Constant Illumination. The subjects were asked visually to divide each of the five patches B, C, D, E and F in four equal-width strips and report the lightnesses in each of these regions. The 4 observed lightness values from top to bottom associated with each patch are those from the strips in left to right order across each patch.

The other two staircase targets are of the same form and use the same patch labelling. The Figure 7 target was of uniform luminance across each patch. Figures 8 and 9 list the corresponding luminances and observed lightness for staircase targest in which the luminance varies across each patch as if it were under an illumination gradient. In one case it increases, in the other it decreases.

Patch	Luminance	Pixel value	Observed lightness
A	1	1	1.2
B	840.8 (left edge)	249 (left edge)	8.6
			8.6
	1001 (right edge)	255 (right edge)	8.6
			8.6
C	840.8 (left edge)	249 (left edge)	7.8
			7.9
	1001 (right edge)	255 (right edge)	7.9
			8
D	840.8 (left edge)	249 (left edge)	7.5
			7.7
	1001 (right edge)	255 (right edge)	7.7
			7.9
E	840.8 (left edge)	249 (left edge)	7.4
			7.6
	1001 (right edge)	255 (right edge)	7.6
			7.8
F	840.8 (left edge)	249 (left edge)	7.3
			7.5
	1001 (right edge)	255 (right edge)	7.5
			7.8

Figure 8: Staircase under Rising Illumination Intensity

Patch	Luminance	Pixel value	Observed lightness
A	1	1	1.0
B	1001 (left edge)	255 (left edge)	9
			8.9
	841 (right edge)	249 (right edge)	8.9
			8.8
C	706.3 (left edge)	242 (left edge)	8.9
			8.8
	498.4 (right edge)	229 (right edge)	8.8
			8.8
D	418.6 (left edge)	223 (left edge)	7
			7
	351.6 (right edge)	216 (right edge)	7
			7.1
E	295.4 (left edge)	210 (left edge)	6.6
			6.6
	248.1 (right edge)	204 (right edge)	6.7
			6.7
F	208.4 (left edge)	197 (left edge)	6.2
			6.2
	175.1 (right edge)	191 (right edge)	6.4
			6.5

Figure 9: Staircase under Decreasing Illumination Intensity

AUTOMATIC SELECTION OF THE NUMBER OF ITERATIONS

To determine the optimal number of iterations (i.e., cycles of comparing a pixel to its neighbors at each pyramid level), we plotted the RMS (root mean square) error in retinex predictions as a function of the number of iterations. The variation in error is shown for the case of the SC data in Figure 1.

Since subjects reported a single lightness value for each patch, we calculate retinex's lightness estimate based on the mean of the lightness estimates over all pixels within a patch. The prediction error for a patch therefore measures the difference between retinex's lightness estimate and the mean across all subjects of the lightnesses of the matches made for that patch. The overall error for a target is the simply the RMS of the errors for the individual patches it contains.

For the simultaneous contrast targets, the minimum error occurs at a small number of iterations as can be seen from Figure 10. The "Single" line shows the average RMS error of retinex predictions in lightness units for the case of a target (Figure 4) in which there are three areas: the gray center, the white surround and the black background. At one iteration, with a linear scaling of max and min, the RMS value is 0.9. That is much larger than the standard deviation of observer results of 0.52, 0.23 and 0.13. Increasing the number of iterations to 10 causes a drop in RMS values to 0.2 units. From 10 to 50 units the values drop from 0.2 to 0.1. For this target any value above 5 iterations does a reasonable job of matching the observer data.

The thin line "Double" describes the data from Figure 5. Here the average data show a minimum value around 6 or 7 iterations. This is because the dark gray surround and the gray area in the dark gray are very sensitive to the number of iterations. They have such a large effect that it makes the average of five areas exhibit the minima. The luminances from the two central grays is the same. This target is of interest because the two grays do not look the same. With too few iterations the calculated value for the gray in black is too high. At the point of minimum error, the calculation renders the gray in black one lightness unit higher than the gray in white. When increasing the number of iterations above the minimum, the calculations report that the two grays are identical, losing the ability to predict simultaneous contrast and increasing the RMS error.

Figure 11 shows the average error for the 24 targets in the combined MMT, SC, GB and IG data sets versus the number of iterations. The minimum error now occurs when the number of iterations is quite large; although, the curve is quite flat so the minimum is also not very distinct. Since it is clear from Figure 11 that a single choice for the number of iterations cannot be determined from the RMS error measurement, we considered what test might serve as an automatic stopping condition so that the number of iterations could be adjusted automatically on a case-by-case basis. The stopping condition cannot be based on minimizing the RMS error directly, since the subjects' matches are not available to the retinex—they are after all what retinex is supposed to be predicting.

We have chosen instead to use the change in retinex output as the number of iterations is increased from n to $n+1$. While this is analogous to the situation of numerical solution of a typical optimization problem where the minimization process is iterated until the change becomes small enough, it is not precisely the same. The difference is in the meaning of the term 'iteration'. In the optimization case, the entire process is repeated until convergence; whereas, in the retinex case, retinex in its entirety is not being repeated. Instead, it is the number of times the process of cycling through the neighbors is repeated at each level.

Let R_x^n be the retinex output at location x when retinex's iterations parameter has been set to n. The proposed retinex stopping condition for image size N and threshold ε can be expressed as:

$$\sqrt{\frac{\sum_x (R_x^{n+1} - R_x^n)^2}{N}} \leq \varepsilon$$

Using this stopping condition, the number of retinex iterations will vary with the input target. What is the optimal value of ε? We determine an optimal value for it by brute force search. In other words, we chose an initial high value for ε, ran retinex on all the test targets and calculated the RMS prediction error, decreased ε by a small amount and repeated the process. A minimum occurs at $\varepsilon = 0.015$. The average prediction error drops to 0.62. In comparison, the minimum average error for any fixed choice of the number of iterations (as shown in Figure 11) remained at 1.71.

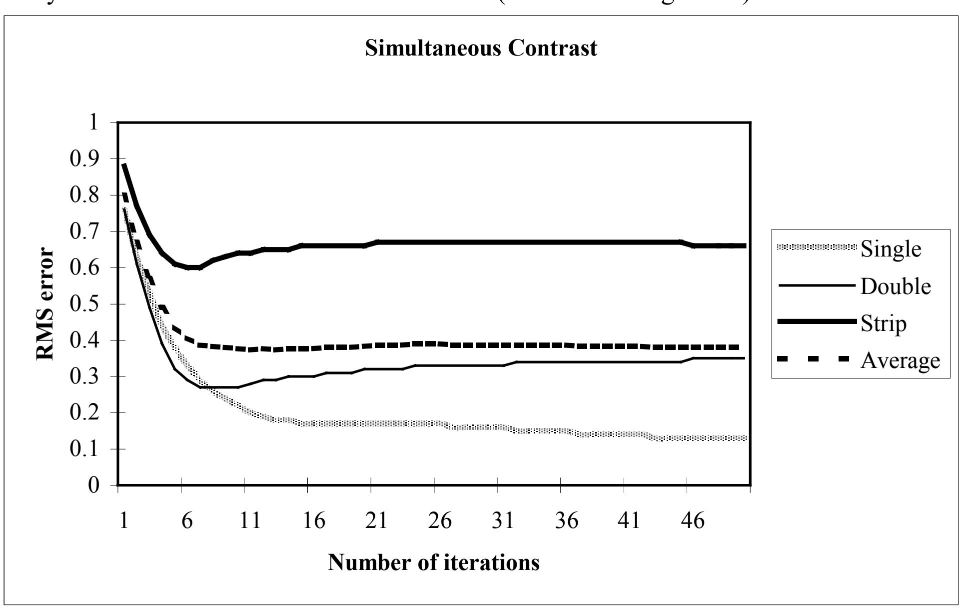


Figure 10: Simultaneous Contrast Targets: RMS error measuring the difference between retinex lightness predictions and subjects' reported matching lightness as a function of the number of iterations.

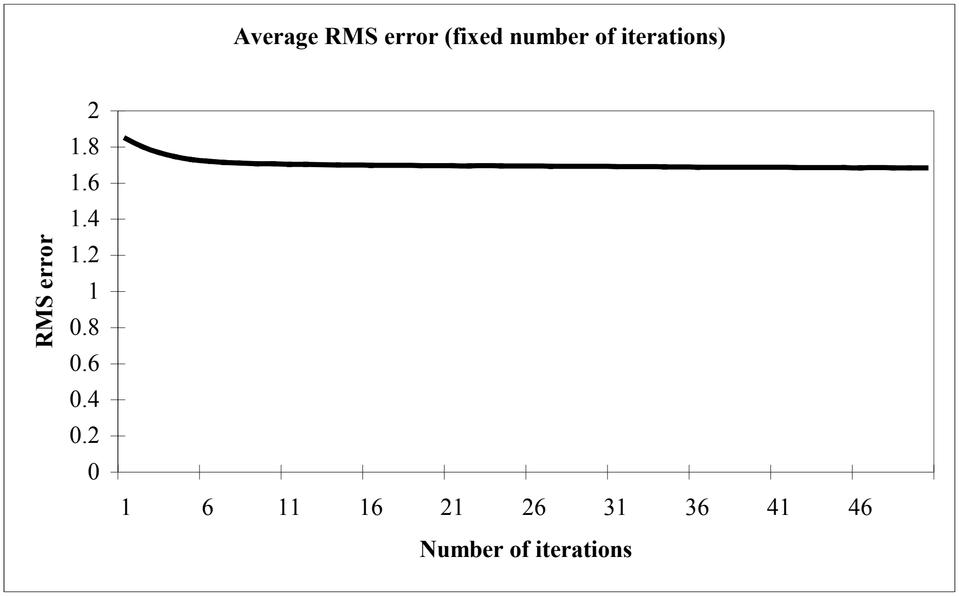


Figure 11: RMS error in retinex lightness prediction averaged across MMT, SC, GB and IG experiments as a function of the number of iterations. For each choice of the number of iterations parameter, the same choice is then used for retinex for all targets.

CONCLUSION

Our goal has been to specify retinex as completely as possible so that it can be tested unambiguously. With the proper selection of parameters, retinex can reduce the average RMS prediction error to 0.62 units on a 1-to-9 lightness scale. This requires the retinex parameters 'post-lut' and 'number of iterations' be set. In this paper, we presented a new method of setting them automatically. Optimizing for a fixed setting for the number of iterations resulted in an overall average RMS error of 1.71, so the new automatic-stopping-condition technique constitutes a significant improvement over a single choice for the number of iterations. Since the method changes only retinex's input parameters, the retinex model itself has not changed. However, the advantage of using the retinex model in conjunction with automatic parameter selection is that it can be applied in a hands-off manner without requiring further intervention. Future work will include modifying retinex to employ different numbers of iterations automatically at each pyramid level.

ACKNOWLEDGEMENTS

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

REFERENCES

1. Funt, B. and Ciurea, F. "Control Parameters for Retinex" *AIC 2001 Proc 9th Congress of the International Color Association*, June 2001, in press.

2. Funt, B.V., Ciurea, F., and McCann, J.J., "Retinex in Matlab" *Proc. IS&T/SID Eighth Color Imaging Conference*, 112-121, Scottsdale 2000.

3. Glasser, L.G., McKinney, A.H., Reilly, C.D., and Schnelle, P.D. "Cube-root color coordinate system". *J. Opt. Soc. Am.* 48, 736-740, 1958.

4. McCann, J.J., Land, E.H., and Tatnall, S.M.V., “A Technique for Comparing Human Visual Responses with a Mathematical Model for Lightness” *Amer. J. Of Optometry and Archives of Amer. Academy of Optometry*, Vol. 47, 845-855, 1970.

5. McCann, J.J., McKee, S.P., and Taylor, T.H., “Quantitative Studies in Retinex Theory: A Comparison Between Theoretical Predictions and Observer Responses to the ‘Color Mondrian’ Experiments”, *J. of Vision Res.* Vol. 16, 445-458, 1976.

6. Stiehl, W.A., McCann, J.J., and Savoy, R.L., “Influence of Intraocular Scattered Light on Lightness-Scaling Experiments”, *J. Opt. Soc. Am.* 73, 1143-1148, 1983.

---

[*] funt@sfu.ca; phone 604-291-3126; fax 604-291-3045; http://css.sfu.ca/members/funt.html; School of Computing Science, Simon Fraser University, 8888 University Drive, Burnaby, B.C. Canada V5A 1S6

[**] fciurea@sfu.ca; phone 604-291-4717; fax 604-291-3045; http://www.cs.sfu.ca/~fciurea/; School of Computing Science, Simon Fraser University, 8888 University Drive, Burnaby, B.C. Canada V5A 1S6

[***] mccanns@tiac.net; phone 617-484-7865; fax 617-484-2490; McCann Imaging, 161 Claflin Street, Belmont, MA, USA, 02478