Visual Evaluation and Evolution of the RLAB Color Space

Mark D. Fairchild
Munsell Color Science Laboratory, Center for Imaging Science
Rochester Institute of Technology, Rochester, New York

Abstract

The RLAB color space has been tested for a variety of viewing conditions and stimulus types. These tests have shown that RLAB performs well for complex stimuli and not-so-well for simple stimuli. This paper reviews the various psychophysical results, interprets their differences, and describes evolutionary enhancements to the RLAB model that simplify it while improving its performance.

Introduction

The accurate reproduction of color images in different media has a number of requirements[1]. One of the most notable is the need to specify and reproduce color appearance across a range of media and viewing conditions. This cannot be accomplished using traditional colorimetry which is only capable of predicting color matches under identical viewing conditions for the original and reproduction. When viewing conditions such as the luminance level, white-point chromaticity, surround relative luminance, and cognitive interpretation of the medium vary, a color-appearance model is necessary to predict the appropriate image transformation required to produce an image that closely resembles the color appearances of the original.

The RLAB color-appearance space was developed by Fairchild and Berns for cross-media color reproduction applications in which images are reproduced with differing white points, luminance levels, or surrounds[2]. Since its development, the RLAB space has been subjected to an extensive series of psychophysical comparisons with other color-appearance models. This paper reviews the RLAB space, briefly describes the results of some visual evaluations of its performance, and outlines the derivation of a revised version of RLAB. The revisions result in a simpler formulation of RLAB with performance equal to or better than the original in all applications evaluated to date.

Overview of RLAB

For a detailed derivation of the original RLAB equations, the reader is referred to reference 2. A descriptive summary of the philosophy and implementation of the RLAB color-appearance space is given below.

RLAB was derived to have color-appearance predictors similar to those of the CIELAB color space[3]. RLAB includes predictors of lightness, L^R, redness-greenness, a^R, yellowness-blueness, b^R, chroma, C^R, and hue angle, h^R. These appearance-predictors are calculated using equations virtually identical to the CIELAB equations after the stimulus tristimulus values are transformed to the corresponding tristimulus values for a reference viewing condition (D65, 318 cd/m^2, hard copy). The transformation is accomplished using a modified von Kries-type chromatic adaptation transformation previously formulated by Fairchild[4]. The end result is that the RLAB color space is identical to (and takes advantage of the excellent performance of) the CIELAB color space for the reference viewing conditions. However, for other viewing conditions, the more accurate chromatic-adaptation transform replaces the normalization of tristimulus values inherent in the CIELAB equations.

The chromatic-adaptation transform utilized in RLAB has several unique features. The first is the capability to predict incomplete levels of chromatic adaptation that result in highly chromatic "white-points" retaining some of their chromatic appearance. In addition, the incomplete-chromatic-adaptation feature can be turned on and off depending on whether or not cognitive "discounting-the-illuminant" mechanisms are active or not. These mechanisms are active when viewing hard-copy images in an illuminated environment and inactive when viewing soft-copy images. A final unique feature is a matrix in the transformation that modeled interaction between the cone-types allowing the prediction of luminance-dependent appearance effects such as the Hunt effect.

Another feature of the RLAB model is that the power-function nonlinearities in the CIELAB equations (cube root) are allowed to vary depending on the image-surround conditions. This is to model the change in image contrast caused by changes in the relative luminance of the image surround. For example, the dark surround in which projected slides are typically viewed causes the perceived contrast to be lower than if the same image luminances were presented in an average surround as is typical of a printed image.

Visual Evaluation of RLAB

A series of experiments have been undertaken to visually evaluate the performance of various color-appearance models under a variety of viewing conditions using both complex stimuli (images) and simple color patches. This section reviews and summarizes the results of five such studies. The performance of the RLAB color-appearance model relative to the other models in these experiments has provided a greater understanding of its relative strengths and weaknesses.

Print-to-Print Image Reproduction

Experiment 1 examined the reproduction of printed images viewed under different light sources at different luminance levels. Details of this experiment were described by Kim, *et al*[5]. Four pictorial images were used in this experiment. The originals were viewed under a CIE illuminant A simulator at a luminance level (white) of 214 cd/m^2. Reproductions were viewed under fluorescent CIE illuminant D65 simulators at one of three different luminance levels (71, 214, and 642 cd/m^2). The reproductions were produced by applying color-appearance transformations as described by each of eight models. The reproductions were viewed pairwise in every possible combination and 30 observers were asked to choose which image in each pair was a better reproduction of the original. The data were then analyzed using Thurstone's Law of Comparative Judgements to derive interval scales of model performance. Confidence limits were also calculated about each of the scale values. The images were viewed using a successive-*Ganzfeld* haploscopic viewing technique.[6] The rank order of each model's performance (averaged over all images and conditions) is given in Table I. Models that did not perform significantly differently than one another are given identical ranks. Only the data for the five appearance models common to all of the experiments is given in Table I. In experiment 1, the RLAB, CIELAB, Hunt, and von Kries models all performed similarly, while the Nayatani model performed significantly worse. The other three models performed worse than each of these five, which is why they were not included in further experiments.

Table I. Rank order of model performance in each of the 4 visual experiments.

Model	Exp. 1	Exp. 2	Exp. 3	Exp. 4
RLAB[2]	1	5	1	1
CIELAB[3]	1	2	2	2
von Kries[7]	1	2	3	2
Hunt[8]	1	1	4	4
Nayatani[9]	5	2	5	(5)

Simple Object-Color Reproduction

Experiment 2 was virtually identical to experiment 1 with the exception that simple color patches on gray backgrounds were used as stimuli rather than pictorial images. Details of this experiment were described by Pirrotta.[10] Ten different original colors, chosen to maximize differences between the appearance model predictions, were used. The originals were viewed under a CIE illuminant A simulator at a luminance level (white) of 73 cd/m^2. Reproductions were viewed under fluorescent CIE illuminant D65 simulators at a luminance of 763 cd/m^2. Nine different color-appearance transformations were evaluated using the same experimental procedure and analysis as experiment 1 and viewed via the successive-*Ganzfeld* haploscopic technique[6] by 26 observers. The rank order of each model's performance (averaged over all colors) is given in Table I. In experiment 2, the Hunt model performed significantly better than the others followed by the Nayatani, von Kries, and CIELAB models with similar performance. The RLAB model performed significantly worse than all of the other models in this experiment. It is of interest that the RLAB model performed best for pictorial images and worst for simple color patches under similar experimental conditions.

Further analysis of the RLAB model showed that it introduced an unwanted shift in the lightness of the color samples upon changes in luminance level. This resulted in the poor performance of RLAB for the simple color patches. This problem was not apparent in the experiments using pictorial images since the lightness shift occurred for all of the image colors and the image contrast was properly reproduced. This deficiency in the RLAB model was traced to the **C** matrix, which models interactions between the cone types. The problem is corrected by removal of the **C** matrix in the revised formulation of RLAB given below.

Print-to-CRT Image Reproduction

Experiment 3 examined the performance of five color-appearance transformations for reproductions of printed original images as CRT-displayed images. The experiment was carried out using five different viewing techniques to determine which was most appropriate for such comparisons[11]. A memory-matching technique was determined to be the best. Thus, the memory-matching results are summarized below. Five different pictorial images were used as originals. In one session, the originals were viewed under a fluorescent CIE illuminant D50 simulator. In the second session, the originals were viewed under a CIE illuminant A simulator. The reproductions were viewed on a CRT monitor with CIE illuminant D65 white-point chromaticities. The luminance of white for all conditions was 75 cd/m^2. Both the originals and reproductions were viewed with white borders, gray backgrounds, and dark surrounds. Fifteen observers took part in this experiment. A paired-comparison experiment with data analysis similar to the first two experiments was used. The model-performance rank order (averaged over images and print white points) is given in Table I. This experiment proved to be the most sensitive test of model performance with each model performing significantly differently than the others. The order of performance from best to worst was RLAB, CIELAB, von Kries, Hunt, Nayatani. The problems exhibited by RLAB in experiment 2 were not apparent in this experiment due to the use of equal luminance levels and complex images.

CRT-to-Projected Slide Image Reproduction

Experiment 4 was carried out in a manner similar to experiment 3. However, the original images were presented on a CRT display with white-point chromaticities of either CIE illuminant D65 at 53 cd/m^2 or CIE illuminant D93 at 60 cd/m^2 and the reproductions were projected 35mm transparencies with a white-point correlated color temperature of 3863K at a luminance of 109 cd/m^2. The CRT images were viewed in a dim surround of office lighting and the projected transparencies were viewed in a dark surround to test the models' abilities to predict surround effects. Fifteen observers completed the experiment. The data were collected using a memory matching technique and analyzed in a way similar to the first three experiments. The Nayatani model was excluded from the psychophysical experiments since the images produced by it were clearly inferior to those produced by other models. The rank order results (averaged over three pictorial images) are given in Table I. The RLAB model performed best followed by CIELAB and von Kries in a tie, Hunt performed the worst of the models actually evaluated. Details of this experiment are described by Fairchild, *et al.*[12] in these proceedings.

Image and Color Dependence

It should be noted that the results described in this paper are the overall average results for each experiment. There are many details worthy of further investigation in the complete results of each experiment. For example, the performance of the models is typically somewhat image dependent. Usually the rank order of the models remains approximately the same, but occasionally more drastic dependencies can be noted. For example, CIELAB performs poorly for blue hues. Thus, if an experiment were designed using images that all had a preponderance of blue, the performance of CIELAB would likely be much worse than indicated by the results summarized above. The same is also true for experiment 2 in which simple color patches were used. The models' performance differed for the various colors investigated. This color dependency is likely to be a major cause of the observed image dependency.

Evolution of RLAB

The RLAB model performs as well as, or better than, all of the other color-appearance transformations in the experiments dealing with images. Since the original objective in the derivation of RLAB was to develop a simple model that would perform at least as well as more complicated models in color reproduction applications it seems that it has been successful. However, the poor performance of RLAB in experiment 2 highlighted a flaw in the model that could easily be corrected without affecting the good performance in the other experiments. In addition, further simplifications of the equations have been derived that allow easier implementation and inversion of the model (both necessary for imaging applications). This was accomplished by replacing the "if-then" linear/power functions of the CIELAB equations with approximately equivalent simple power functions that do not require the "if-then" implementation and its complex inversion. Further flexibility was added to RLAB by allowing the cognitive "discounting-the-illuminant" mechanisms to be partially active. This is likely the case in situations such as large projected transparencies in a darkened room. Lastly, the capability to express hue as percentage combinations of the unique hues was added to provide a more precise definition of perceived hue. These changes are detailed below in the new RLAB equations.

Summary of RLAB Equations

The following equations describe the forward implementation of the new RLAB equations. Changes are described as they are presented. One begins with a conversion from CIE tristimulus values (Y = 100 for white) to fundamental tristimulus values as illustrated in Eqs. 1 and 2.

$$\begin{vmatrix} L \\ M \\ S \end{vmatrix} = \mathbf{M} \begin{vmatrix} X \\ Y \\ Z \end{vmatrix} \tag{1}$$

$$\mathbf{M} = \begin{vmatrix} 0.4002 & 0.7076 & -0.0808 \\ -0.2263 & 1.1653 & 0.0457 \\ 0.0 & 0.0 & 0.9182 \end{vmatrix} \tag{2}$$

The next step is calculation of the **A** matrix that is used to model the chromatic adaptation transformation.

$$\mathbf{A} = \begin{vmatrix} a_L & 0.0 & 0.0 \\ 0.0 & a_M & 0.0 \\ 0.0 & 0.0 & a_S \end{vmatrix} \tag{3}$$

$$a_L = \frac{p_L + D(1.0 - p_L)}{L_n} \tag{4}$$

$$P_L = \frac{(1.0 + Y_n^{1/3} + l_E)}{(1.0 + Y_n^{1/3} + 1.0 / l_E} \tag{5}$$

$$l_E = \frac{3.0(L_n / 102.70)}{L_n / 102.70 + M_n / 98.47 + S_n / 91.82} \tag{6}$$

The a terms for the short- (S) and middle-wavelength (M) sensitive systems are derived in a similar fashion using analogous functions. Y_n is the absolute adapting luminance in cd/m^2. Terms with n subscripts refer to values for the adapting stimulus. The D factor was added in Eq. 4 to allow various proportions of cognitive "discounting-the-illuminant". D should be set equal to 1.0 for hard-copy images, 0.0 for soft-copy displays, and an intermediate value such as 0.5 for situations such as projected transparencies in completely darkened rooms.

After the **A** matrix is calculated, the tristimulus values for a stimulus color are converted to corresponding tristimulus values under the reference viewing conditions using Eqs. 7 and 8.

$$\begin{vmatrix} X_{ref} \\ Y_{ref} \\ Z_{ref} \end{vmatrix} = \mathbf{RAM} \begin{vmatrix} X \\ Y \\ Z \end{vmatrix} \tag{7}$$

$$\mathbf{R} = \begin{vmatrix} 186.01 & -112.95 & 21.98 \\ 36.12 & 63.88 & 0.0 \\ 0.0 & 0.0 & 108.89 \end{vmatrix} \tag{8}$$

The RLAB coordinates are then calculated using Eqs. 9-13.

$$L^R = 100(Y_{ref} / 100.00)^{\sigma} \tag{9}$$

$$a^R = 430[(X_{ref} / 95.05)^{\sigma} - (Y_{ref} / 100.00)^{\sigma}] \tag{10}$$

$$b^R = 170[(Y_{ref} / 100.00)^{\sigma} - (Z_{ref} / 108.88)^{\sigma}] \tag{11}$$

$$C^R = \sqrt{(a^R)^2 + (b^R)^2} \tag{12}$$

$$h^R = \tan^{-1}(b^R / a^R) \tag{13}$$

Equations 9-11 have been simplified as described above to avoid complexities in the implementation and inversion of the CIELAB-style equations. The exponents have changed slightly, but their ratios have remained the same. For an average surround $\sigma = 1/2.3$, for a dim surround $\sigma = 1/2.9$, and for a dark surround $\sigma = 1/3.5$. The hue composition, H^R, can be calculated via linear interpolation of the values in Table II. These were derived based on the notation of the Swedish Natural Color System (NCS). Example values are listed in the table in italics. The inversion of the revised RLAB equations is straightforward.

Table II. Data for conversion from hue angle to hue composition

h^R	R	B	G	Y	H^R
24	100	0	0	0	R
0	*82.6*	*17.4*	*0*	*0*	*R17B*
270	*17.4*	*82.6*	*0*	*0*	*R83B*
246	0	100	0	0	B
180	*0*	*21.4*	*78.6*	*0*	*B79G*
162	0	0	100	0	G
90	0	0	0	100	Y
24	100	0	0	0	R

Comparison of New and Old RLAB Equations

To compare the new and old RLAB equations, a sample of 125 colors was generated (5 levels each of X, Y and Z). The RLAB coordinates of each of these colors were calculated with both the old and new equations and then the RLAB (*i.e.* CIELAB) color differences between the old and new predictions were calculated. For an average surround the mean color difference was 6.18 units with a maximum of 12.17. The mean color differences were 5.49 and 4.94 with maximums of 11.40 and 10.56 for dim and dark surrounds respectively. While these changes might seem large, they are not significant when compared to the inter-observer variability in color-appearance judgments which can often exceed 20 CIELAB units[13] and the differences between similarly-performing color-appearance models which are even larger. It should also be noted that the gamut of the 5X5X5 XYZ sampling (a simulation) far exceeds the gamut of physically realizable colors, thus producing a more rigorous comparison of the equations. The changes are not systematic as illustrated in Figure 1, a vector plot indicating the change from the old to new RLAB equations.

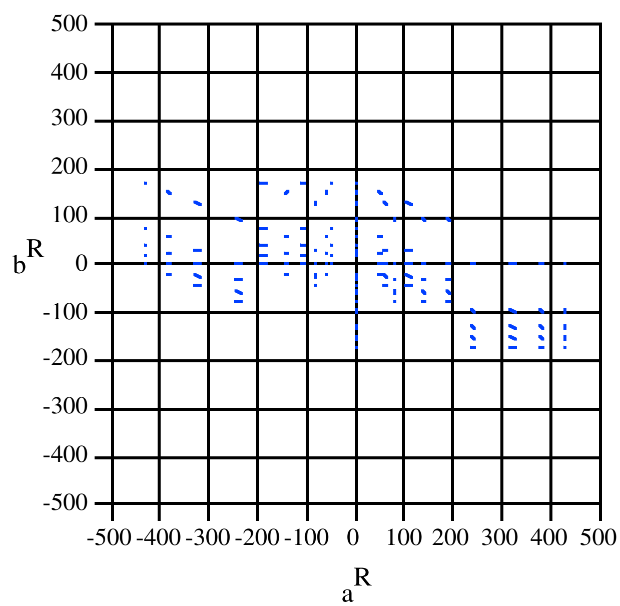

Figure 1. Vectors indicating the magnitude of the change from the old to new RLAB equations for an average surround.

Conclusion

The RLAB color-appearance space performs as well as or better than more complex appearance models in imaging applications. This is likely due to the complex nature of image-color appearance judgments in comparison with judgments of simple color patches. The added complexity in other appearance models might be useful for predicting subtle color-appearance effects. However, these effects are apparently masked in image judgments.

The RLAB equations have been further simplified while at the same time improving their performance for all types of applications.

The Hunt model performed very well in experiment 2 on simple patches. Thus it is surprising that it did not perform equally well in other situations. One reason for this is some ambiguity in deciding the values of the various parameters in the Hunt model for a particular application. The model was implemented exactly as published in the above experiments. However, it is clear that the Hunt model can perform as well as the RLAB model if its various parameters are optimized to the particular viewing conditions.[13] An advantage of the RLAB model is that its simplicity leaves little room for ambiguity in its implementation.

Acknowledgments

This research was supported by the NSF-NYS/IUCRC and NYSSTF-CAT Center for Electronic Imaging Systems.

References

1. M. D. Fairchild, Some hidden requirements for device-independent color imaging, *SID International Symposium,* 865-868 (1994).
2. M. D. Fairchild and R. S. Berns, Image color-appearance specification through extension of CIELAB, *Color Res. Appl.,* **18**: 178-190 (1993).
3. CIE, *Colorimetry*, CIE Publication **15**.2, Vienna (1986).
4. M. D. Fairchild, Formulation and testing of an incomplete-chromatic-adaptation model, *Color Res. Appl.,* **16**: 243-250 (1993).
5. T. G. Kim, R. S. Berns and M. D. Fairchild, Comparing appearance models using pictorial images, *IS&T/SID Color Imaging Conference*, 72-77 (1993).
6. M. D. Fairchild, E. Pirrotta and T. G. Kim, Successive-Ganzfeld Haploscopic Viewing Technique for Color-Appearance Research, *Color Res. Appl.*, **19:** 214-221 (1994).
7. J. von Kries, Chromatic adaptation, Festrchrift der Albrecht-Ludwig Universität, Fribourg, 1902 (Translation: D. L. MacAdam, *Sources of Color Science*, MIT Press, Cambridge, 1970).
8. R. W. G. Hunt, An improved predictor of colourfulness in a model of color vision, *Color Res. Appl.,* **19**: 23-26 (1994).
9. Y. Nayatani, K. Takahama, H. Sobagaki and K. Hashimoto, Color-appearance model and chromatic adaptation transform, *Color Res. Appl.,* **15**: 210-221 (1990).
10. E. Pirrotta, Testing chromatic adaptation models using object colors, M.S. Thesis, RIT, (1994).
11. K. Braun and M. D. Fairchild, Viewing Environments for Cross-Media Image Comparisons, *IS&T's 47th Annual Conference/ICPS*, 391-396 (1994).
12. M. D. Fairchild, R. S. Berns, A. A. Lester and H. K. Shin, Accurate color reproduction of CRT-displayed images as projected 35mm slides, *IS&T/SID Color Imaging Conference,* 69-73 (1994).
13. R. W. G. Hunt and M. R. Luo, Evaluation of a model of color vision by magnitude scalings: Discussion of collected results, *Color Res. Appl.,* **19**: 27-33 (1994).