Accurate Color Reproduction of CRT Displayed Images as Projected 35mm Slides

Mark D. Fairchild, Roy S. Berns, Audrey A. Lester and Hae Kyung Shin
Munsell Color Science Laboratory, Center for Imaging Science
Rochester Institute of Technology, Rochester, New York

Abstract

Accurate color reproduction of images presented on a computer-controlled CRT display as projected 35mm transparencies is a complicated procedure requiring the characterization and control of several imaging processes and application of appropriate color-appearance modeling to account for the changes in viewing conditions. This paper reviews a process for image-recorder characterization, projection-system characterization, and testing of color-appearance models for this application.

Introduction

With the recent and rapid technological growth of affordable, open, color-image input and output devices has come a strong need for color management systems to facilitate device-independent color imaging. One application is the development of presentations and other imagery on desktop computer systems with the final intention being the production of 35mm slides for projection. Ideally users should be able to select colors and view images on the computer-driven CRT display that are accurate representations of the colors that will ultimately be seen on the projected slides. There are numerous requirements for the successful implementation of such a device-independent, open, color imaging system.[1] This paper reviews the development and evaluation of a system for reproducing CRT-displayed images as projected 35mm slides. Particular attention is given to the issues of device characterization and color-appearance transformation. The equally critical issue of gamut mapping was not addressed in this work by limiting the images evaluated to colors common to the gamuts of both image displays.

CRT-Display Characterization

The colorimetric characterization of CRT displays has been the subject of considerable study. No new techniques were developed for this project. The following section briefly reviews the techniques that were used.

Colorimetric Characterization

The CRT characterization was carried out using the techniques detailed by Berns *et al.*[2, 3] The tristimulus values of the display phosphors are carefully measured with a high-accuracy colorimeter in order to derive the 3×3 linear transformation from linearized RGB phosphor excitations to CIE XYZ tristimulus values. The XYZ tristimulus values of a series of approximately 10 levels of gray (R = G = B) colors are also measured to characterize the nonlinear electro-optical transfer functions of the display. These XYZ tristimulus values are transformed to RGB tristimulus values and then a nonlinear curve-fitting technique is used to fit a model incorporating gain, offset, and gamma terms to derive a transfer function between digital counts and radiometric scalars for each of the three CRT channels. Typical colorimetric accuracy of this characterization technique is on the order of 1.0 CIELAB unit color difference between actual and predicted colors.

Spatial Uniformity

The spatial uniformity of the CRT display was assumed to be a negligible issue. However, it is clear that displays vary in luminance across the face of the CRT and are likely to vary slightly in the other parameters of the characterization model as well. The characterization measurements were taken in the center of the display and image presentations were limited to this central area.

Film-Recorder Characterization

Once accurate colorimetric data are available for the original CRT images, it becomes necessary to convert this information into RGB digital counts that can be used to drive the film recorder to produce slides with the desired colorimetry. A model for the film-recorder colorimetric characterization and its implementation are described below. A Solitaire 8xp image recorder was used. Additional details can be found in reference 4.

Film Model and Characterization

The first stage of calibrating the film-recorder system was the characterization of the photographic film. Kodak Ektachrome 100 Plus Professional film was used throughout this project. The developed film is analyzed colorimetrically through measurement of its spectral transmittance. The spectral transmittance of the film can be treated as a combination of the spectral transmittance of the film base and the variable amounts of the cyan, magenta, and yellow image-forming dyes. This can be accurately modeled using the Beer-Bouger Law as illustrated in Eq. 1.

$$T_\lambda = T_{\lambda,g} e^{-(c_c D_{\lambda,c} + c_m D_{\lambda,m} + c_y D_{\lambda,y})} \qquad (1)$$

T_λ is the spectral transmittance of the film, $T_{\lambda,g}$ is the spectral transmittance of the base, the c terms are the concentrations, and the D terms are the unit spectral absorptivities of the cyan, magenta, and yellow dyes. Given this model of the film, it remains to determine the spectral absorptivities for the particular film dyes and develop a relationship between the RGB digital counts driving the film-recorder exposure and the dye concentrations in the resulting image.

The spectral absorptivities of each of the component dyes were determined through a statistical analysis of exposed and developed film colors. A set of 60 different single-color exposures were used. The spectral transmittance measurements of these samples were converted to spectral densities (absorptivities) and subjected to a principal components analysis. The first three characteristic vectors were rotated via an equimax rotation to assure that each of the three vectors accounted for equal amounts of the variance in the data set. The first 3 vectors accounted for over 99% of the variance. These vectors provide estimates of the dye spectral absorptivities. However, due to possible inter-image effects and a data sampling that might not contain equal variation in each of the three dye concentrations, these estimates might be biased. A second analysis was completed in which ramps of 11-steps in RGB exposure were imaged to produce samples approximately varying in only one dye concentration at a time. The first characteristic vector for each of these ramps were obtained as secondary estimates of the dye absorptivities. The first set of vectors were then rotated to the closest least-squares fit to the second set of vectors in order to produce the most robust global estimate of the dye spectral absorptivities. These were used for the remainder of the modeling.

A set of 21 single-color exposures were made in order to model the dye concentrations. These consisted of ramps of 4 exposures in which one of the RGB digital counts was varied from 64 to 255 in steps of 64 and the other two were set at 0 and a fourth ramp of 9 gray exposures with R = G = B varying from 0 to 255 in steps of 32. The spectral transmittances of each of these samples were measured and the dye concentrations were determined iteratively using a modified form of the Allen[5] tristimulus matching algorithm described by Berns[6] and the dye spectral absorptivities derived above. The results were a data set consisting of cyan, magenta, and yellow concentrations and the RGB digital counts used to produce them for the 36 colors.

Digital-Count-to-Dye-Concentration Model

The first step of the model to relate dye concentrations to exposure digital counts is the characterization of the nonlinear transformation between the two. Since this nonlinearity is made up of the nonlinear nature of photographic film, the nonlinear characteristics of the CRT used in the film recorder, and the particular nature of the film recorder look-up tables (LUTs) in use, it is extremely difficult to characterize it with an analytical model. The approach taken was to build one-dimensional LUTs that directly mapped each of the RGB digital counts to the respective CMY dye concentrations. If photographic film were a simple tri-pack with three independent layers and the RGB exposures each produced development in only one layer, then this model would be sufficient. However, neither of these situations exist. Film is designed such that development in a given layer can affect the amount of dye produced in another layer (inter-image effects) and it is difficult to design an efficient exposing system that assures single-layer exposures. Thus the model had to be enhanced. This was accomplished by adding a 3×3 matrix transformation on the dye concentrations after the linearizing LUTs in order to produce more accurate estimates of the dye concentrations. This 3×3 was obtained via linear regression while constraining each row to a sum of 1.0 to maintain gray balance. To summarize, RGB digital counts were converted to CMY dye concentrations through 3 1-D LUTs followed by a 3×3 matrix transformation. These concentrations were used to reconstruct film spectral transmittance curves which were then used to calculate tristimulus values.

Overall Performance and Implementation

An independent set of test colors was generated to evaluate the overall accuracy of the film recorder model. The average CIELAB ΔE^* (2° Observer, Projector Source) was 5.73 with a maximum of 10.78 for the full color set and the 1.67 with a maximum of 2.13 for the grays. While this performance is not as good as might be desired, it is probably near the limit of what can be achieved for such a system and was accomplished with a very simple model requiring the generation and measurement of only 21 color samples.

The color variability of the complete system was analyzed by making several series of repeat exposures. The exposure variability for repeated exposures of the same color on a single roll of film showed a maximum ΔE^* of 0.5. The processing variability for a similar set of exposures on three different rolls of film showed an average ΔE^* of 0.5 with a maximum of 1.9. The degradation of the film after 4 minutes of projection showed an average ΔE^* of 1.5 and after 16 minutes of projection an average ΔE^* of 2.6. It is clear that the sources of variability in the system quickly add up to an error magnitude similar to the performance of the characterization procedure described above. It should also be noted that the processing variability was evaluated over a 3-day period with a single photo finisher. Long-term tracking of control strips showed variation as much as 3 times larger than what was observed during this 3-day period. In addition, variability across photo finishers would be substantially larger. Thus the estimate of processing variability is conservative.

The film-recorder characterization model described above cannot be analytically inverted to allow the prediction of the RGB digital counts required to produce a desired set of XYZ tristimulus values. Thus the model was implemented using an iterative inversion procedure to build a 3-dimensional LUT that was used to transform image data.

Projector-System Characterization

The film-recorder model described above allows the spectral transmittance of the exposed and developed film to be predicted from the RGB digital counts used to make the exposure. However, the transformation from the spectral transmittance of the film to the tristimulus values measured from a screen when the slides are projected remains to be characterized. Difficulties in this step of the process are reviewed below. A Leica P-2200 projector was used throughout this study.

Theory and Measurements

In theory tristimulus values of a projected image would be calculated by multiplying the transmittance of the film by the spectral power distribution of the projector source and the spectral reflectance factor of the screen and then integrating with each of the CIE color matching functions. The one variable to be addressed would be the proper measurement geometry (total or regular) for the spectral transmittances. This theory was evaluated by making measurements of the spectral transmittances of slides in the projector system and comparing them to measurements made in a spectrophotometer using both total and regular geometries. Discrepancies were discovered that were spectrally selective, functions of the slide color, and not correlated with geometry differences. The causes of these discrepancies were determined and addressed as described below.

Thermochromism

The major cause of discrepancies between the spectral transmittance measurements made in the projector and those made in a spectrophotometer was determined to be thermochromism — a change in color with temperature. The measurements made in the spectrophotometer were at room temperature (20°C) while the slide temperature in the projector was approximately 52°C. Significant color shifts are expected in any colored material upon such large temperature fluctuations. A series of colored slides were measured both in the spectrophotometer (cool) and in the projector (hot). The color shifts were found to be significantly color dependent and typically resulted in a decrease in chroma with increasing temperature. The changes averaged a ΔE^* of 4.6 with a maximum of 10.3. The color shifts were reversible upon cooling. Details of the analysis of thermochromism in slide film can be found in reference 7.

This complication was addressed by measuring the color samples used to characterize the film through principle components analysis in the projector system. This resulted in dye spectral absorptivity curves for hot film which differed from those for cool film. The "hot film" curves were used when modeling the system. In addition, all of the measurements used for the con-struction of the digital-count-to-dye-concentration model were made in the projector system. Had this procedure not been followed, the overall accuracy of the film-recorder characterization would have decreased by a factor of two.

Staining

Slide film also is degraded upon projection through dye fading and the production of a yellow stain. This was evaluated by measuring color samples after various durations of projection. After the first 4 minutes of projection the average color shift was 1.5 CIELAB units. After an additional 8, 16, and 32 minutes, the average color differences increased to 2.0, 2.6, and 3.2 units respectively. Thus the color shifts induced in the slide film by prolonged projection can be significant. Reference 7 contains additional details.

Spatial Uniformity

The spatial uniformity of the projection system was not addressed in the characterization. All measurements were made in the central portion of the image area and the image sizes were restricted to this area to minimize variation due to spatial non-uniformity. The uniformity of the system was evaluated and found to be mainly a luminance fall-off from the center to the corners of approximately 10% — less than that typically found on CRT displays.

Summary Recommendations

Several procedures were followed to assure the most accurate colorimetric presentation of the projected slide images. All slides were preconditioned in the projection system for 4 minutes prior to any measurements or visual evaluations. The largest portion of the yellow stain formation took place in the first four minutes. By preconditioning the slides, these shifts were avoided in the measurements and observations. Each slide was placed in the projector for 75 sec. prior to being viewed by any observer or measured. This allowed the temperature and therefore thermochromic shifts to stabilize. After any slide was visually evaluated 8 times (16 minutes projection) it was discarded and replaced with a new slide. After this duration of projection, the colors have shifted enough to be perceptibly different than when the slide was first viewed.[8] In addition, all the slides used in the experiments described below were exposed and processed on a single day to minimize processing variability.

The entire characterization process is dependent on the slide projector being used. Thus accurate characterization for projected slides must include characterization of the projection system, not just the film recorder.

Color-Appearance Transformations

Once the CRT and film-recorder systems are colorimetrically calibrated, the issue of color appearance modeling must be addressed since the CRT images and projected slides are typically viewed with different white points, luminance levels, and surrounds. Simple reproduction of tristimulus values results in unsatisfactory color reproduction since these variables are not taken into consideration. Color-appearance models must be used. Several models were implemented and psychophysically evaluated as described below.

Implementation

The models evaluated were RLAB,[9] Hunt,[10] CIELAB,[11] and von Kries.[12] The model published by Nayatani *et al.*[13] was not included since it was previously shown to produce unacceptable images and due to all of the complications described above the experimental variables had to be minimized.

Two CRT set-ups were evaluated. The first had white-point chromaticities of CIE illuminant D65 at a luminance (white) of 53 cd/m^2 while the second had white-point chromaticities of CIE illuminant D93 at a luminance of 60 cd/m^2. The CRT images were always viewed with a gray background filling the remainder of the CRT display and a dim surround consisting of a room illuminated with cool-white fluorescent lighting at an illuminance of 345 lux. None of the ambient light was allowed to reflect off the CRT face. The slide projector system had a white-point CCT of 3863K. The slide images had a gray background and were viewed with a dark surround. The CRT and projected images were arranged such that they subtended the same visual angle (approx. 7°) from the observers' position.

Each model was implemented as published with no specific optimization for the viewing conditions. The RLAB and Hunt models were implemented for a dim surround on the CRT and dark surround for the slides and it was assumed that there was no cognitive "discounting the illuminant" for either display. (This also means that the Helson-Judd effect was predicted to occur according to the Hunt model.) A medium gray background (20%) was used and implemented in the Hunt model. The von Kries and CIELAB models only account for changes in white point.

Psychophysical Tests

The CRT images were treated as originals and the desired reproductions for each condition were produced according to each of the 4 models. Three different pictorial images were used. Problems of gamut mapping were avoided by gamut-compressing the original images such that no attempts were made to produce out-of-gamut colors. For one image, the prediction of the Hunt model did not allow a reproduction to be made without excessive gamut compression, thus the Hunt model was not used for that particular condition.

Slides were produced that each contained a pair of reproductions. Each possible pair was produced. A total of 15 observers performed 4 phases of observations. For each CRT white point, a preference and a matching experiment were completed. The preference experiment was performed first. Observers were asked, for each pair, which of the two images they preferred. This was done with no knowledge of the original images. In the matching experiment, observers first studied the CRT original, and then made judgments of the slides by choosing which of each pair was the best color match to the original. The preference and matching data were analyzed using Thurstone's Law of Comparative Judgments to derive interval scales and uncertainties of image preference and match quality for each CRT white point and image.

Results

Table I shows the rank order of each model's performance in the four experimental phases averaged over all three images. Scale values that are statistically distinct with 95% confidence are given different ranks. In each experiment, the RLAB model performed best. The Hunt model performed worst for the color matching experiments, but better for the image preference experiments. The results were significantly image dependent. However, the RLAB model performed best in all cases. The image preference judgments are significantly different than the image matching judgments, affirming that observers are indeed making distinct judgments. Again, the RLAB model was best in both circumstances. The observers were also asked to respond whether or not the best-matching image was an acceptable reproduction of the CRT original. For the RLAB model, observers thought the reproductions were acceptable 79% of the time. The acceptability of the other models was substantially lower.

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

Model	D65 Match	D65 Pref.	D93 Match	D93 Pref.
RLAB	1	1	1	1
CIELAB	2	4	2	4
von Kries	2	2	3	2
Hunt	4	3	4	2

The rather poor performance of the Hunt model can be traced to two factors. The first is the inclusion of the Helson-Judd effect which causes a hue shift in the gray scale as a function of lightness. The inclusion of this effect is according to how the model was published. In the Hunt model, when cognitive "discounting-the- illuminant" does not occur adaptation is incomplete and the Helson-Judd effect is present. The RLAB model does not predict the Helson-Judd effect (in fact it is never observed in practical viewing conditions), but does predict incomplete chromatic adaptation. Recently, Hunt and Luo[14] have suggested that for projected slides the Hunt model should be implemented with incomplete adaptation, but no Helson-Judd effect. This would result in predictions much more similar to the RLAB predictions. Also, the surround compensation in the Hunt model is apparently more severe than that in RLAB. This combined with the Helson-Judd effect prediction cause the Hunt model to produce images that cause much more severe gamut-mapping challenges.

Conclusion

The accurate colorimetric characterization of a CRT-to-projected-slide color reproduction system provides significant challenges. However, the results obtained are quite striking and far superior to the reproductions typically obtained with film recorders that have not been colorimetrically calibrated.

The model-based approach outlined in this paper is both accurate and efficient. The CRT display can be calibrated with only 13 colorimetric measurements and the film recorder system requires only 21 measurements after the film has been characterized. This allows for efficient recharacterization of the system when various system parameters are changed. Exhaustive measurement characterization techniques cannot be used to update characterizations as efficiently and have no implementation advantage since the models can be used to build 3-dimensional LUTs to implement image transformations.

The RLAB color-appearance model proved to be a satisfactory mechanism for transforming image color-appearance data from a CRT display viewed in a dim surround to a projected slide viewed in a dark surround. The more-complicated Hunt model could make similar predictions if its parameters were optimized. However, the added complexity of the Hunt model might not be necessary in image reproduction applications for which the precision of color-appearance judgments seems to be reduced.

Acknowledgments

This research was supported by the NSF-NYS/IUCRC and NYSSTF-CAT Center for Electronic Imaging Systems, Management Graphics, Inc., and Eastman Kodak Company.

References

1. M. D. Fairchild, Some hidden requirements for device-independent color imaging, *SID International Symposium,* 865-868 (1994).
2. R. S. Berns, R. J. Motta and M. E. Gorzynski, CRT colorimetry. Part I: Theory and practice, *Color Res. Appl.,* **18**: 299-314 (1993).
3. R. S. Berns, M. E. Gorzynski and R. J. Motta, CRT colorimetry. Part II: Metrology, *Color Res. Appl.,* **1 8**: 315-325 (1993).
4. R. S. Berns, Colorimetric characterization of the Solitaire 16 film recorder for Kodak Extachrome 100 Plus Professional: A pilot study, *Munsell Color Science Laboratory Technical Report*, May, (1993).
5. E. Allen, Colorant formulation and shading, *Optical Radiation Measurements* Vol. **2**, Academic Press, Orlando, Chapter 7 (1984).
6. R. S. Berns, Spectral modeling of a dye-diffusion thermal-transfer printer, *J. Elec. Imaging*, **2:** 359-370 (1993).
7. A. A. Lester and M. D. Fairchild, Thermochromism of Ektachrome 100 Plus transparencies upon projection, *J. Img. Sci. Tech.*, **38:** 332 (1994).
8. M. Stokes, M. D. Fairchild and R. S. Berns, Precision requirements for digital color reproduction, *ACM Transactions on Graphics*, **11**: 406-422 (1992).
9. M. D. Fairchild, and R. S. Berns, Image color-appearance specification through extension of CIELAB, *Color Res. Appl.,* **18**: 178-190 (1993).
10. R. W. G. Hunt, An improved predictor of colourfulness in a model of color vision, *Color Res. Appl.,* **19**: 23-26 (1994).
11. CIE, *Colorimetry*, CIE Publication **15**.2, Vienna (1986).
12. J. v. Kries, Chromatic adaptation, Festrchrift der Albrecht-Ludwig Universität, Fribourg, 1902 (Translation: D. L. MacAdam, *Sources of Color Science*, MIT Press, Cambridge, 1970).
13. Y. Nayatani, K. Takahama, H. Sobagaki and K. Hashimoto, Color-appearance model and chromatic adaptation transform, *Color Res. Appl.*, **15**: 210-221 (1990).
14. R. W. G. Hunt and M. R. Luo, Evaluation of a model of color vision by magnitude scalings: Discussion of collected resuls, *Color Res. Appl.*, **19**: 27-33 (1994).

published previously in the IS&T 1994 Color Imaging Conference Proceedings, page 69