# Precision of Color Matches and Accuracy of Color-Matching Functions in Cross-Media Color Reproduction

*Mark D. Fairchild and Richard L. Alfvin*
*Munsell Color Science Laboratory, Center for Imaging Science*
*Rochester Institute of Technology, Rochester, New York*

## Abstract

A visual experiment was completed to evaluate the degree of observer metamerism in color matches between typical color reproduction media. The results illustrate that current CIE standard colorimetric observers provide reasonable estimates of average color matches and that the range of color mismatches encountered in cross media color reproduction due to observer variability can be very large, on the order of 10 CIELAB units.

## Introduction

Due to differences in the spectral characteristics of the primary colorants used in various imaging systems, all cross-media color matches are metameric. Since metameric matches are dependent upon the characteristic spectral responsivities of the human visual system, variations in color-matching functions will result in variations in metameric matches. A metameric color match perceived by one observer may appear to be a significant mismatch to another observer. This phenomenon is known as observer metamerism.

Pobboravsky[1] studied the effect of observer metamerism on color matches between CRT displays and prints. He performed a set of theoretical calculations based on the color matching data used in the CIE recommended technique for assessing observer metamerism.[2] His conclusion was that observer metamerism was not a significant problem in these types of matches. He verified this conclusion with qualitative visual judgements of acceptability. These data, while useful, might well underestimate the importance of observer metamerism. The research described in this paper was undertaken to better quantify observer metamerism.

The following sections describe an experiment designed to explore the effects of observer metamerism in cross-media color matching by quantifying the precision and accuracy of three sets of color-matching functions and the magnitude of observer variability found in hard-copy to CRT color-matches. The effects of observer met-amerism for both intra- and inter-observer color matches are quantified and compared. In addition, the results are compared with the CIE recommendations on observer metamerism.[2] Uncertainties derived by Nimeroff, Rosenblatt, and Dannemiller[3] for the 1964 CIE 10° Supplemental Standard Colorimetric Observer color-matching data are also compared with the experimental data.

## Experimental

A visual experiment was designed to permit observers to make critical color matches between color prints or transparencies and a CRT display. Seven color prints and seven color transparencies were prepared as fixed matching stimuli. The seven colors included red, green, blue, gray, cyan, magenta, and yellow. The color print samples were produced with a Fujix Pictrography 3000 color printer. The color transparencies were imaged with an MGI Solitaire $8_{xp}$ film recorder using 4 × 5 Ektachrome 100 Plus Professional film. The chromaticities of the fixed hard-copy samples illuminated with a fluorescent D50 simulator were designed to effectively sample the color gamut of the Sony Trinitron CRT display used to generate the soft-copy color matches.

A simple optical apparatus, consisting of an equilateral glass prism mounted on an optical bench, allowed observers to simultaneously view both the soft and hard-copy matching stimuli. A diffuser was placed in front of the CRT display to eliminate the appearance of scan lines. The fixed hard-copy stimulus and the adjustable soft-copy stimulus were presented in a vertical symmetric bipartite field. The color-matching stimuli were presented as solid colors appearing self-luminous in a darkened room.

The hard-copy stimuli were illuminated by a GTI Soft View D50 fluorescent light booth designed for viewing both reflective and transmissive materials. The light booth was adjusted to yield equal luminance from the reflective and transmissive gray hard-copy stimuli. The color of the soft-copy image produced with the CRT display was adjusted by the observer to match each fixed hard-copy stimulus. Observers used a computer mouse to independently adjust the color appearance attributes of the soft-copy image along a CIELAB L* vector, and an a*-b* plane. The CRT display was controlled with a Pixar II image computer with 10-bits per RGB color channel resolution.

Twelve male and eight female observers between the ages of twenty-one and fifty-six participated in the color-matching experiment to assess inter-observer variability. Each of the twenty observers successfully passed a screening for congenital color vision deficiencies. A 22-year old male observer performed the color-matching experiment twenty times to assess intra-observer variability. The 5 × 5 cm matching field subtended a visual angle of 2.9°. After a match was attained, a Photo Research 650 telespectroradiometer was used to measure the spectral radiance of both the hard-and soft-copy stimuli from the observers point of view. The order of color matches was randomized for each observer.

## Results and Discussion

Three different sets of color-matching functions were used in conjunction with the spectral radiant power distributions, recorded from both the hard- and soft-copy stimuli, to calculate tristimulus values for each of the observer color matches. The three sets of color-matching functions were the 1931 CIE 2° standard colorimetric observer, the 1964 CIE 10° supplemental standard colorimetric observer, and the 1955 Stiles-Burch 2° mean observer.

CIELAB coordinates were calculated using each of the three sets of color-matching functions, for both hard-and soft-copy stimuli constituting observer-determined metameric pairs. CIE $\Delta L^*$, $\Delta a^*$, and $\Delta b^*$ values were calculated for all metameric pairs. Figure 1 shows a $\Delta a^*$-$\Delta b^*$ plot of the intra- and inter-observer soft-copy color matches for the hard-copy cyan-transparency, with respect to the 1931 CIE 2° observer match point located at the origin.

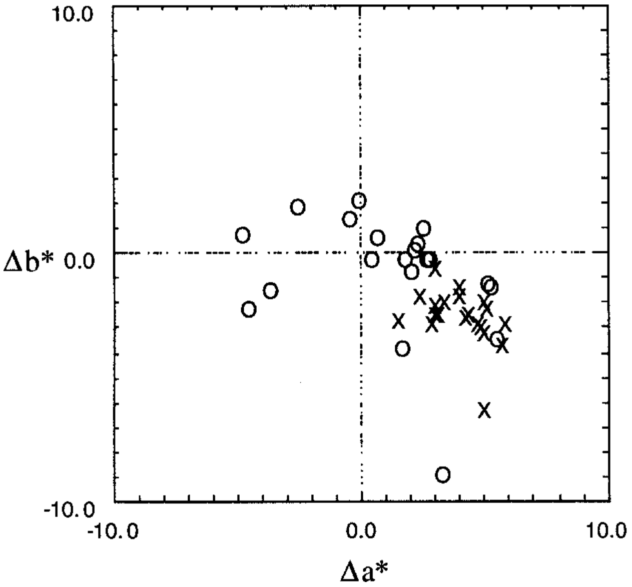

*Figure 1. Match point deviations for 20 observers (O) and 20 matches by a single observer (X)*

**Table I. Summary data for CIELAB deviations and CIE 2° standard colorimetric observer.**

| | Intra-Observer | Inter-Observer |
|---|---|---|
| $\Delta L^*$ min. | -4.14 | -14.08 |
| $\Delta L^*$ max | 4.41 | 13.70 |
| $\Delta L^*$ mean | 0.18 | 0.11 |
| $\Delta a^*$ min. | -6.98 | -15.71 |
| $\Delta a^*$ max | 6.97 | 9.76 |
| $\Delta a^*$ mean | 0.36 | 0.15 |
| $\Delta b^*$ min. | -11.23 | -19.69 |
| $\Delta b^*$ max | 8.29 | 18.51 |
| $\Delta b^*$ mean | -1.55 | -0.70 |

For all sample colors, the maximum CIELAB differences were as large as 19.7 units for inter-observer matches, and 11.2 units for intra-observer matches. Table I shows the statistics on the CIELAB deviations for the CIE 2° observer. The relatively small mean CIELAB deviations for the inter-observer data indicate that the average of all experimental color matches correlates well with the theoretical color matches of the standard observers. Similar results were obtained for the other two sets of color-matching functions.

The sample covariance matrix, defined by the sample covariances and variances of $\Delta L^*$, $\Delta a^*$, and $\Delta b^*$ values, was calculated for each combination of observer, color-matching functions, color, and medium. Assuming a multivariate normal distribution, the sample covariance matrix can be used to construct a 95% confidence region for the sample distribution. The ellipsoidal confidence region encloses 95% of the observer responses for this stimulus. A confidence region for the mean of a sample distribution can also be described by factoring in the number of observations Such a region defines the certainty with which the mean is known and becomes smaller with increasing number of observations.

The 95% confidence regions for the sample distribution and sample mean of the inter-observer cyan-transparency color matches calculated with the CIE 2° observer are plotted together in Figure 2. The larger of the two ellipses encompasses the distribution of the sample population of color matches, while the smaller ellipse describes the uncertainty of the mean color match of the sample population.

In addition to the $\Delta a^*$-$\Delta b^*$ relationship examined in Figure 2, it is important to consider the $\Delta a^*$-$\Delta L^*$, and $\Delta b^*$-$\Delta L^*$ planes because the experimental color matches involved adjustments to each of the three colorimetric dimensions. In the case of the cyan transparency, the mean color match of the twenty observers is not significantly different than the predicted color match for the CIE 2° observer. In other words, the 95% confidence region for the mean includes the origin in all three dimensions simultaneously. This evaluation can be completed statistically using a Hotelling's $T^2$ test. This test for means was used to compare the experimental mean color matches with the theoretical color matches cal- culated for each of the three sets of color-matching functions.

Table II shows the results of Hotelling's $T^2$ test comparing the mean color matches made by the group of twenty observers with the hypothetical means predicted using each of the three sets of color-matching functions. The check marks (√) indicate that the mean color matches are not significantly different from the specified standard observer at the $\alpha = .05$ level. The Stiles-Burch and CIE 2° color-matching functions outperformed the CIE 10° observer in terms of predicting the mean color match for a population of color normal observers. The relatively poor performance of the CIE 10° observer is not unexpected, considering that the experimental stimulus was restricted to a 3° visual field. Although the standard observer color-matching functions were not able to predict the mean color matches of the group of twenty observers for every sample, it is important to note that in every case the predicted color matches were contained in the ellipsoids that defined the 95% confidence regions for the sample distributions of inter-observer color matches. Therefore, the hypothetical color matches determined with the standard observers are representative of a member of the population of inter-observer color matches determined in this experiment.

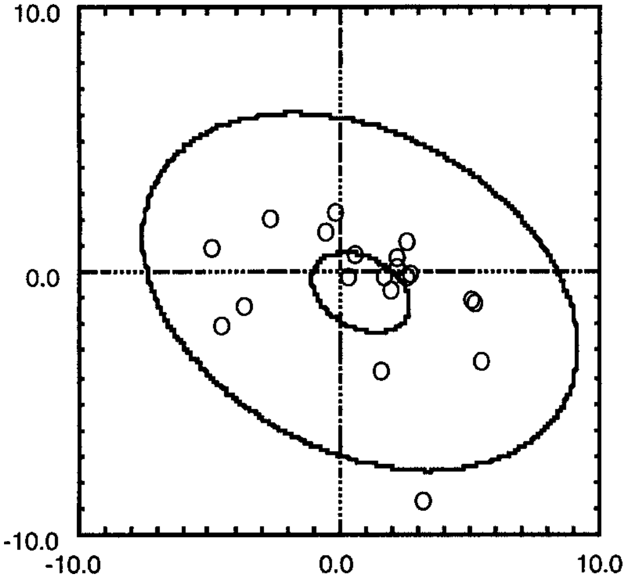

*Figure 2. 95% confidence regions for the cyan transparency.*

**Table II. Results of Hotelling's $T^2$ test for means on the inter-observer data.**

| | Transparency | | | Print | | |
|---|---|---|---|---|---|---|
| | S-B 2° | CIE 2° | CIE 10° | S-B 2° | CIE 2° | CIE 10° |
| Red | √ | √ | | √ | √ | √ |
| Green | √ | √ | √ | √ | √ | √ |
| Blue | | | | | | |
| Gray | | √ | | √ | √ | √ |
| Cyan | | | | √ | √ | √ |
| Magenta | √ | √ | √ | √ | √ | |
| Yellow | √ | √ | √ | | | |

Hotelling's $T^2$ test for means comparing the mean color matches made by the single observer with the hypothetical mean color matches predicted with the three sets of color-matching functions were also completed. The means were found to be significantly different for all color matches except the green transparency and the 2° observers. The predicted mean color matches were contained in the ellipsoids defining the 95% confidence regions for the sample distributions of intra-observer color matches in only 26 out of 42 cases. It is apparent from these results that the single observer is not characteristic of the standard observers examined in this experiment. It is not unreasonable to expect any single observer to deviate significantly from any given average or standard observer.

## Predicted Ranges of Color Mismatches

The CIE Standard Deviate Observer,[2,4] was designed to represent the individual variations among color-normal observers. It includes a methodology for evaluating a range of color mismatches for metameric pairs by a statistical confidence ellipse in a u'v' chromaticity diagram. Instead of transforming the XYZ covariance matrix to a u'v' covariance matrix as recommended by the CIE,[2] the XYZ covariance matrix was transformed to an L*a*b* covariance matrix for comparison with the ranges of the intra- and inter-observer mismatches from the experiment described in this paper.

Nimeroff, Rosenblatt, and Dannemiller derived fundamental estimates of variances and covariances in spectral tristimulus values, based on color matching for individual observers.[3] A variation of the CIE methodology to define a range of color mismatch is used to construct 95% confidence regions of color mismatches based on these tristimulus value uncertainties.

Figures 3 shows four bivariate 95% confidence regions (each shifted to the origin) for measured and predicted ranges of color mismatch for the cyan transparency. The four regions are defined by the inter-observer sample data, the intra-observer sample data, the CIE standard deviate observer recommendations, and the Nimeroff *et al.* tristimulus uncertainties. The intra- and inter-observer ellipses defined by experimental data are significantly larger than the predicted CIE, and Nimeroff *et al.* based ellipses. This especially true in the L* dimension (not shown in Figure 3).

The loss of luminance variance associated with the normalization of the color-matching function data used by the CIE, and Nimeroff *et al.* to assess observer variability results in a severe compression of the predicted variance in the L* dimension. The CIE and Nimeroff *et al.* based confidence ellipses predicted for each of the thirteen other color-medium samples were found to be similarly compressed.

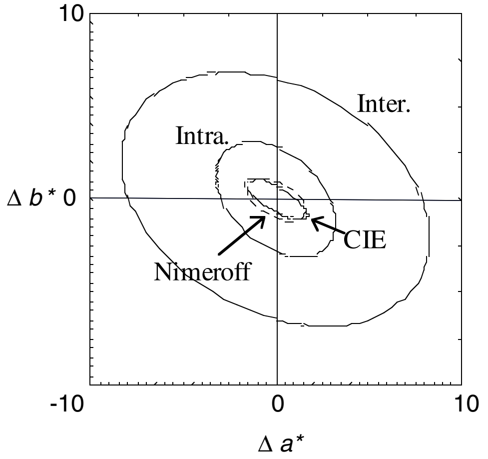

*Figure 3. Experimental and predicted ranges of color mismatches*

## Monte Carlo Simulation

Unfortunately, the data required to derive an accurate technique for predicting observer variability do not exist. The required data are the spectral covariance functions for color-matching functions that have not been normalized. The time, effort, and cost required to collect such data are prohibitive and it is not likely that the data will become available any time soon. However, much has been learned in recent years about the properties and magnitudes of the physiological causes of variation in color vision.[5] It is possible such data could be used in a Monte Carlo approach to simulating color matching functions for a population of observers. If such a simulation was successful, a set of spectral covariance data useful for predicting the observed results might be derived. A Monte Carlo simulation of individual color matching functions is currently underway. Results will be presented at a future time.

## Conclusion

A visual experiment was designed and performed to measure the magnitude of uncertainties associated with both intra- and inter-observer variability in cross-media color matching. The range of color mismatch was found to be as large as 19 CIELAB units for hard- to soft-copy color matches. The results indicate that the variability of inter-observer color matches is approximately twice as large as the variability of intra-observer color matches for the metameric pairs of soft and hard-copy media examined in this experiment. All metameric match points for the three standard observers were found to be inside the 95% confidence ellipses of the sample distributions of inter-observer color matches. The majority of the match points for the 2° standard observers were contained within the 95% confidence ellipses of the sample means of inter-observer color matches. These results suggest that the existing CIE Standard Colorimetric Observers are a reasonably good representation of the population of normal trichromats.

The experimentally established ranges of intra- and inter-observer color mismatch were found to be significantly larger than the ranges of color mismatch calculated from CIE recommendations on observer metamerism and the ranges of color mismatch calculated from the estimated spectral tristimulus uncertainties derived by Nimeroff *et al.* The CIE method, and Nimeroff *et al.* uncertainties are based on normalized color-matching function data which appears to have significantly diminished the variance component associated with luminance, resulting in regions of color mismatch unnaturally compressed in the $L^*$ dimension. The CIE recommendations on observer metamerism should be reviewed and further research should be performed to better quantify the variability associated with observer metamerism in practical cross-media color matching.

A more complete review of the experimental results and analysis described in this paper is available in reference 6.

## Acknowledgements

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

## References

1. I. Pobboravsky, “Affect of small color differences in color vision on the matching of soft and hard proofs”, *TAGA Proceedings*, 62-79 (1988).
2. *Special Metamerism Index: Change In Observer*, CIE Publ. No. **80**, Central Bureau of the CIE, Vienna, (1989).
3. I. Nimeroff, J.R. Rosenblatt, and M. C. Dannemiller, “Variability of spectral tristimulus values”, *J. Res. NBS* A **65**: 475-483 (1961).
4. North, M. D. Fairchild, “Measuring color-matching functions. Part II. New data for assessing observer metamerism”, *Color Res. Appl.* **18**: 167-169 (1993).
5. V.C. Smith and J. Pokorny, “Chromatic-discrimination axes, CRT phosphor spectra, and individual variation in color vision”, *J. Opt. Soc. Am. A* **12**: 27-35 (1995).
6. R. L. Alfvin and M. D. Fairchild, “Observer metamerism: Precision of color matches and accuracy of color-matching functions”, *Color Res. Appl.* **22**: 174–188 (1997).

published previously in the IS&T 1995 Color Imaging Conference Proceedings, page 18