Estimating Spectral Reflectances from XYZ Data

Kristyn Falkenstern and Phil Green, London College of Communication (UK); and Marc Mahy, Agfa Graphics (Belgium)

Abstract

Color management is the method of communicating color information between devices with predictable and consistent results. A traditional color management workflow communicates color based on tristimulus colorimetry. This method of communication reduces the accuracy and uniqueness of the color information.

Convexity is used in this work to find the most extreme set of spectral reflectances that all result in a single CIE XYZ value. Once the entire convex set is determined analytically, a method is needed to deduce the most optimal reflectance. A number of different techniques are implemented, primarily involving quadratic programming.

The analytical method used to find all spectral reflectances for a given CIE XYZ set is evaluated using a set of training spectral curves that are an average of several measurements of a test chart printed according to a standard printing condition. The CIELAB values are compared to the training set under different viewing conditions. For example, the maximum CIE XYZ color difference within the convex set under the second viewing condition, D50 simulator, was ΔE_{ab} 20. All CIE XYZ values derived from the convex set of reflectances under the original viewing conditions resulted in the same CIE XYZ value as the starting data.

Two different approaches to finding an optimal curve are evaluated. The first technique uses two different smoothness constraints to estimate an optimal reflectance curve. The next approach uses an aim curve to estimate reflectances. The curves generated from this technique were very close to the aim and gave no difference to the original CIE XYZ values. A media adjusted scaled aim curve was then used to further improve the results.

After the reflectance curves are generated, a workflow needs to be in place to utilize the additional color information. A spectral CMS extends the functionality of a colorimetric approach to include both spectral data of the color data and the viewing conditions, as well as a dynamic CMM.

Introduction

When color matching using tristimulus colorimetry, two different spectral curves can result in identical tristimulus values under a given set of viewing conditions (e.g. a given illuminant and observer), but not necessarily for various illuminants or another set of color matching functions (CMFs). Hence, the uniqueness of the color data is not preserved. Spectral data enables the user to make a better spectral match that will reduce the color difference between two colors when viewed under different viewing conditions.

A printing state is characterized by giving the relationship between a sample of CMYK digital input values and the corresponding measured colorimetric values. [1] The resulting values are referred to as a characterization data set. Characterization data sets for standard printing conditions are provided by industry organizations such as FOGRA, ANSI CGATS and Japan Color, and available through the ICC Characterization Data Registry

Published characterization data sets generally only contain CIE XYZ and CIELAB data, and since these data are commonly the result of processes such as averaging and smoothing of colorimetric data, there are no spectral data associated with them. This lack of spectral data in characterization data sets and ICC profiles is a barrier to implementing a spectral color management system (CMS). Measured or estimated spectral reflectances are needed in order to match print and proof under different viewing conditions.

Spectral Estimation Background

Several different methods of estimating the spectral reflectance curves from the CIE XYZ values have been developed. Encoding is a common technique. It reduces the number of dimensions to minimize metameric matches. The interim connection space, LabPQR, is an example of an encoding technique. [2]

Recently, a mathematical method was developed using metamer sets. This mathematical method determines the set of physically realizable surface reflectances that achieve the same tristimulus values. [3]

Proposed method

The aim of this work is to develop an analytical estimation technique by calculating the convex set in spectral space. This unique solution includes all possible reflectance curves, and then an optimal curve is determined

First, the tristimulus equation is inverted. This becomes the convex set. The convex set is then constrained to physically possible reflectances. To simplify the estimation process, fluorescence was not considered in this work.

Once the convex set is determined and constrained, several different methods to derive the optimal curve from within the convex set are compared. These include two smoothness constraints and the use of an adjusted aim curve.

The following data sets were used in this work:

FOGRA39 – a standard characterization data set available in the ICC Characterization Data Registry, in XYZ

FOGRA39-DNS – measurements of a test chart printed according to the same printing condition as FOGRA39, in spectral reflectance [4]

TR001 – a standard characterization data set available in the ICC Characterization Data Registry, provided by ANSI CGATS, in XYZ [5].

Colorimetric data in the FOGRA39 and TR001 data sets were computed using a D50 illuminant and CIE 1931 Standard Colorimetric Observer.

Reflectances were estimated from the colorimetric data in the FOGRA39 data set. These estimations were evaluated against the spectral data in the FOGRA39-DNS dataset (FOGRA Training Dataset).

The Convex Set

The spectral reflectance data was estimated by inverting the tristimulus equation (1).

$$\begin{aligned} X &= \sum_i I_i R_i \bar{x}_i \\ Y &= \sum_i I_i R_i \bar{y}_i \\ Z &= \sum_i I_i R_i \bar{z}_i \end{aligned} \tag{1}$$

where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the (1xn) color matching functions (CMF), I is the (1xn) illuminant vector and $\bar{R}$ is the (nxm) spectral reflectance set.

This does not result in a bijective transformation between spectral data and tristimulus values. It is an underdetermined linear system of n – 3 degrees of freedom. [6]

In the absence of fluorescence, two natural conditions are met by surface reflectances. [3] There can be no less than no light (0%) and no more than all light (100%). Thus, the space of physically possible reflectances is constrained and defined in the following equation.

$$R_\lambda \in [0,1] \tag{2}$$

All reflectance curves resulting in the same CIE XYZ value constitute a convex set. This can be shown easily in equation (3) Assume the reflectance curves R_1 and R_2 both map to the same CIE XYZ value. Then any convex combination for:

$$R = (1-k) * R_1 + kR_2 \tag{3}$$

with $0 \le k \le 1$
and R *maps to the same CIE XYZ value for all wavelengths*
$0 \le R_\lambda \le 1$

To describe a convex set, only the boundary points have to be known. Any point inside the set can be obtained by a convex combination between the boundary points of the convex set. In general, the number of boundary points can be finite or infinite. [7] A typical example of a convex set with an infinite number of boundary points is a disc in a 2-dimensional plane. In the case of spectral space, the convex set can be described by a finite number of boundary points as equation (1) without the limitations specified in equation (2), resulting in a (n-3)-dimensional subspace of the n-dimensional spectral space. The limitation of (2) can be seen as the limitations defined by (2n) hyperplanes. As the convex set is limited by a finite number of hyperplanes, the number of boundary points of the corresponding convex set is also finite.

Convexity was used to find the reflectance space for a given CIE XYZ data set. The convex set makes up all the reflectance vertices that compose the boundary of the reflectance space that all result in the same tristimulus value. These values were calculated as the first step in estimating the original spectral reflectance.

The convex set resulted in a large number of spectral boundary curves; in one case the estimation resulted in over 10,000 boundary curves. The following examples (Figures 1 and 2) illustrate the convex sets of two different colors computed from selected CIE XYZ values in the FOGRA39 data set.

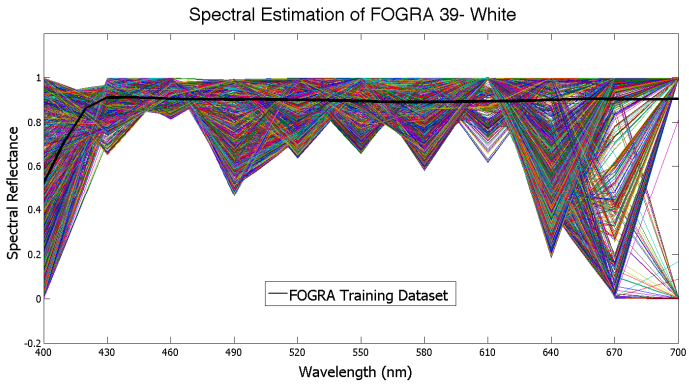

Fig. 1 The convex set of spectral boundary reflectances for paper white. This plot has over 4,000 curves. The solid black line is from the FOGRA39-DNS dataset.

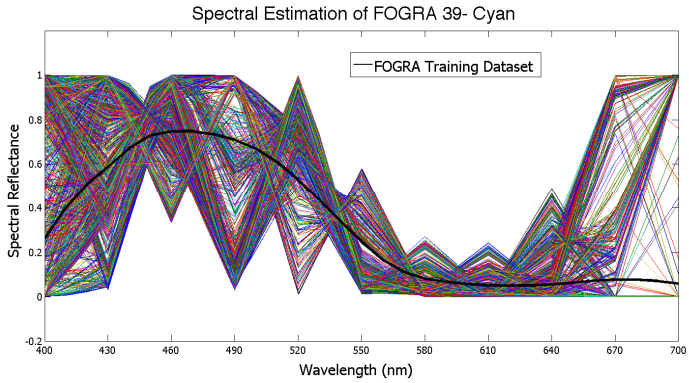

Fig. 2 The convex set of spectral boundary reflectances for primary colorant cyan. This plot has over 3,000 curves.

Evaluating the Convex Set

At any given wavelength, the maximum difference between boundary curves varied from a very small range to 100% of the total reflectance band. For both figures 1 and 2, all of the curves resulted in identical CIE XYZ values under the original viewing conditions. The entire convex set of reflectance curves were converted to CIE XYZ and CIELAB values for a second set of viewing conditions. A convex set in CIE XYZ space was obtained that is completely defined by the boundary spectra (which can be easily shown as it is the intersection between a 3-dimensional subspace and the convex set). When the illumination source was changed from CIE D50 to a D50-simulating source, these boundary spectra did not result in the same CIE XYZ value.

Figure 3 compares the two spectral power distributions of the illumination source.

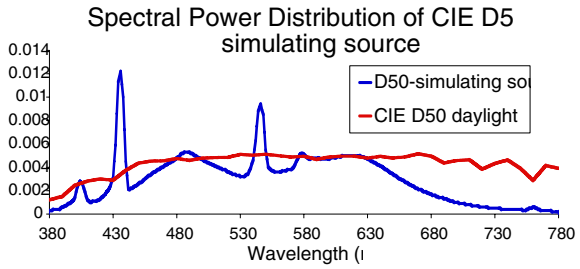

Fig. 3 Spectral power distribution of the CIE daylight illuminant and a D50-simulating illumination source.

The D50-simulating source was chosen because it is commonly used in graphic arts color reproduction.

Figures 4 and 5 illustrate the conversion in CIELAB space of the convex set of reflectances under the second viewing condition, the D50-simulating source. As the relationship between CIE XYZ and CIELAB is not linear, the set in CIELAB space is no longer convex.

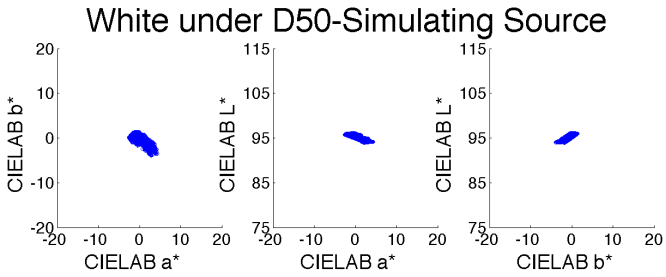

Fig. 4 The projection plots in CIELAB space for the convex set of reflectances for paper white, (D50-Sim/2°).

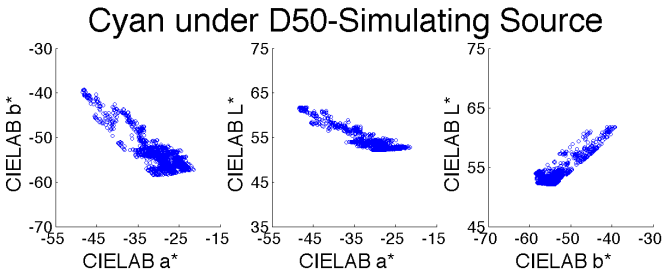

Fig. 5 The projection plots in CIELAB space for the convex set of reflectances for solid cyan, (D50-Sim/2°).

These plots illustrate that the largest difference within the convex set was in the CIELAB a* values and the smallest difference in the CIELAB L* range. This was true for all of the convex sets calculated.

Table 1 gives a statistical summary of the CIELAB results. The comparison is between the convex set averages and the FOGRA39-DNS printed samples.

Average Color Differences of the solid inks (CMYK)
(comparison is between the Convex Set and the printed samples)

		Median	Mean	Maximum Difference	Minimum Difference
Average of the Convex Set and Printed Sample	CIELAB ΔL*	0.24	0.74	2.58	0.15
	CIELAB Δa*	1.39	1.34	2.09	0.41
	CIELAB Δb*	1.38	1.69	4.54	0.11
	ΔE_{2000}*	3.98	3.85	13.08	1.31
Spectra Within the Convex with the maximum difference from the Printed Sample	CIELAB ΔL*	3.63	4.19	7.82	1.93
	CIELAB Δa*	15.34	17.54	35.02	6.92
	CIELAB Δb*	20.35	21.01	34.07	5.85
	ΔE_{2000}*	6.29	8.26	13.08	3.84

Table 1 Summary of results comparing the convex set to the printed sample (D50-Sim/2°). The second comparison is between the spectra from the cyan convex set that has the largest color difference (D50-Sim2°) from the printed sample.

These color differences are larger with CIE standard illuminants A and D65. The color differences within the convex set were also large when the CMFs were changed from the 1931 Standard Colorimetric Observer (2°) to the CIE 1964 Standard Colorimetric Observer (10°) while preserving the original illumination source.

The calculated convex sets are theoretical, and include all reflectances within the constraints described. When an actual 4-ink CMYK printer was used, the maximum CIELAB ΔE_{ab} was only 2.04 for the D50-simulating source. The ink and media used adds additional constraints to the sets that reduce the difference between boundary points in the convex set, and therefore reduce the color differences. Using the same 4-ink printer, the largest color difference of 6.29 was calculated under Illuminant A (2°).

Selecting an optimal reflectance curve

Once the convex set has been calculated, a method to deduce an optimal reflectance curve for a given CIE XYZ set is needed. Two different methods were evaluated, each with its own set of constraints. With both methods, the estimated spectral curve had to be equal to the original CIELAB value for the original viewing conditions. Quadratic programming was used to find the optimal curve with each method.

The first method used is based on van Trigt's work on smoothest reflectance functions [8]. van Trigt defines smoothness as:

$$\int \left(\frac{d^2 R(\lambda)}{d\lambda^2} \right)^2 d\lambda = minimal$$

$$\sum_{i=2}^{n} \left(R_{i+1} - 2R_i + R_{i-1} \right)^2 = minimal \qquad (4)$$

The first smoothness constraint found the curve with the smallest differences between consecutive wavelengths (see equation (4)).

$$\left(\frac{dR}{d\lambda}\right)^2 d\lambda = minimal$$

$$\sum_{i=1}^{n}\left(R_{i+1}-R_i\right)^2 = minimal \quad (5)$$

The second smoothness constraint found the curve with the fewest number of local minima and maxima (see equation (5)).

The second method uses an aim curve to determine the most optimal curve within the convex set. [9] This minimization constraint found the reflectance curve with the smallest difference from the aim (see equation (6)).

$$\sum_{i=2}^{n}\left(R_i - R_{aim}\right)^2 = minimal \quad (6)$$

The aim curve method was generally more accurate than the smoothness constraints. In the example illustrated in Figures 6a and 6b, the TR001 data set was used to estimate the original spectral reflectances. The FOGRA39-DNSdata set was used as the aim.

The CMYK test target was printed using the TR001 printing condition and measured. The following two plots demonstrate how the two methods used compare to the average measured reflectance of the TR001 data.

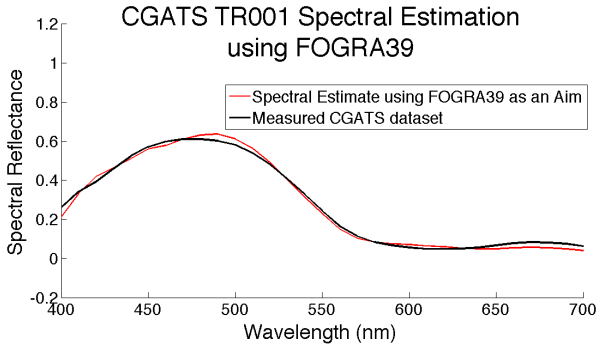

Fig. 6a The CGATS TR001 estimation compared to the measured spectral reflectance.

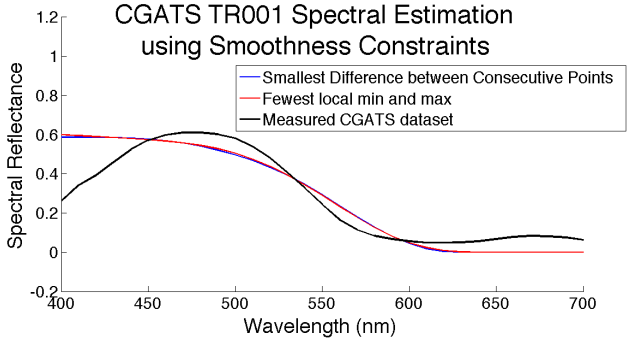

Fig. 6b Comparison plots between the two different methods used to find the most optimal curve estimate. The black lines are the measured CGATS spectral reflectance. The colored lines are the estimates.

The standard deviation was calculated and used used to compare the results of the two techniques, Equation 7.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \quad (7)$$

where $\bar{x}$ is the mean and n is the number of samples

Table 2 gives a summary of the results of the smoothness and the aim curve estimations.

CIELAB Calculated Color Differences between estimates and printed samples, Under D50-simulating source & 2° observer						
		Median	Mean	Standard Deviation	Maximum Difference	Minimum Difference
Smooth 1 Smallest Difference $\frac{d^2R(\lambda)}{d\lambda^2}^2 d\lambda = minimal$	CIELAB ΔL*	-0.45	-0.45	0.07	-0.29	-0.59
	CIELAB Δa*	0.02	0.03	0.10	0.28	-0.10
	CIELAB Δb*	-0.26	-0.30	0.15	-0.09	-0.77
	ΔE_{2000}*	0.47	0.46	0.09	0.69	0.27
Smooth 2 Fewest minimums & maximums $\frac{dR}{d\lambda}^2 d\lambda = minimal$	CIELAB ΔL*	-0.43	-0.41	0.11	-0.21	-0.65
	CIELAB Δa*	0.04	0.06	0.29	0.63	-0.40
	CIELAB Δb*	-0.24	-0.24	0.08	-0.09	-0.39
	ΔE_{2000}*	0.45	0.45	0.10	0.65	0.25
FOGRA39 Aim Curve	CIELAB ΔL*	0.06	0.07	0.06	0.19	-0.03
	CIELAB Δa*	0.13	0.11	0.10	0.24	-0.16
	CIELAB Δb*	0.30	0.21	0.15	0.41	-0.10
	ΔE_{2000}*	0.24	0.23	0.11	0.46	0.05

Table 2 CIELAB values and color differences comparing the two smoothness techniques and the aim curve technique for estimating the TR001 characteristic data.

The results of these techniques are plotting in Figure 7, along with the confidence intervals, Equation 8>

$$CI_{95} = \bar{x} \pm \frac{\sigma}{\sqrt{n}} \quad (8)$$

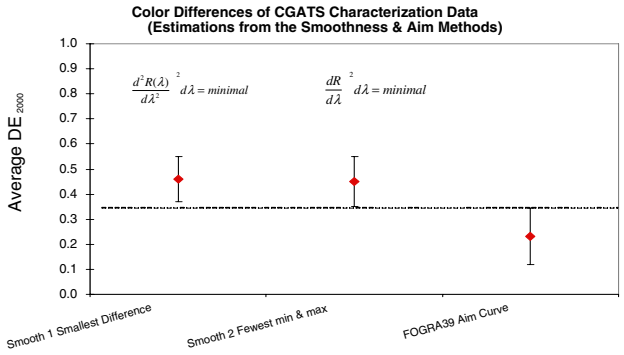

Fig. 7 Mean color differences (De_{2000}) between the two smoothness techniques and the aim curve estimate. Error bars represent 95% confidence intervals (CI_{95}) on the means across 40 samples from the characterization data.*

The results from Figure 7 suggest that there is a statistically significant difference between the two estimation techniques. At this point in the work, only the aim curve method has been used.

To improve the results further, two constraints were investigated to improve the effect of the aim curve. The first constraint was scaling factor, and the second was a media normalization.

The scaling factor constraint found the aim curve with the smallest spectral difference from the convex set (Fig. 8). The scaling factor linearly adjusts the aim curve until it reaches the smallest difference between the aim and the estimate, from 0 to 10 in fractional increments.

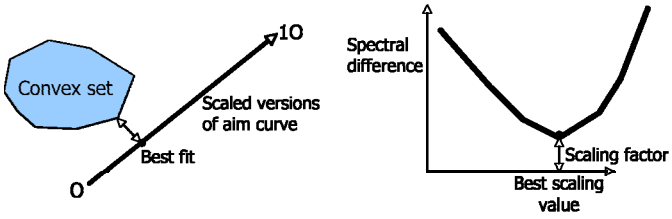

Fig. 8 An illustration explaining the scaling factor for the aim curve.

Table 3 compares the results between using the original aim and the scaled aim curve to estimate the samples printed according to the same printing condition as TR001. CIE XYZ values were calculated for both CIE D50 daylight and D50-simulating Source. The calculated CIE XYZ CIE D50 values were used to estimate the spectral reflectance curves, using the FOGRA39-DNS data set as an aim. A second set of estimates were found using a scaled FOGRA39 set. The estimates and scaled estimates were used to calculate the CIELAB set under a D50-simulating source.

CIELAB Calculations Under D50-simulating source and 2° observer

Solid Primaries	CIELAB Values from measured CGATS			CIELAB calculated from the estimates with FOGRA39 Aim			CIELAB calculated from the estimates with scaled FOGRA39 Aim			Aim Est	Scale Aim
	L*	a*	b*	L*	a*	b*	L*	a*	b*	De_{ab}*	De_{ab}*
White	88.66	4.37	-0.86	88.60	4.20	-1.32	88.65	4.37	-0.88	0.50	0.03
Black	18.64	1.83	-0.49	18.63	2.02	-0.45	18.65	2.08	-0.35	0.19	0.19
Cyan	55.30	-35.43	-45.47	55.35	-35.20	-45.53	55.32	-35.41	-45.29	0.24	0.17
Magenta	47.84	68.03	-5.22	47.52	68.16	-5.87	47.78	68.00	-5.21	0.74	0.07
Yellow	84.12	0.33	81.37	84.00	0.12	81.61	84.09	0.29	81.77	0.34	0.34

Table 3 CIELAB values and color differences of the original training set, the estimate using an aim and the estimate using a scaled aim curve.

The scaling constraint improved the results with a few exceptions. In both examples, the CIELAB ΔE_{ab} color difference between the estimates and measured is less than 1. The scaling factor is especially beneficial when the aim curve is a different technology from that used when generating estimate. Figure 9 shows the results of using the FOGRA39 data set to estimate the spectral reflectance of a newspaper print. The scaled FOGRA39 aim curve gives a much closer estimate to the original measured curve. The CMYK test target was printed on newspaper and the spectral reflectance was measured.

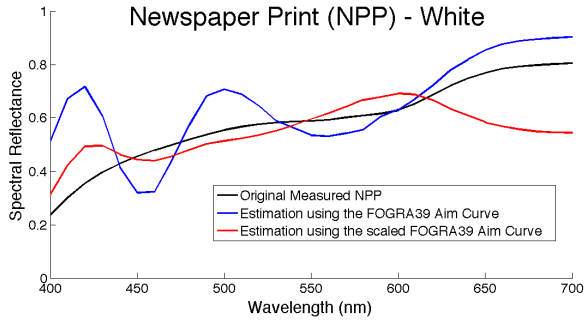

Fig. 9 Comparison plots between the estimates for newspaper print, with and without the scaling factor.

The media normalization constraint deducted the media white from the aim set. These estimated reflectances were adjusted for the reflectance of the media (Fig. 10).

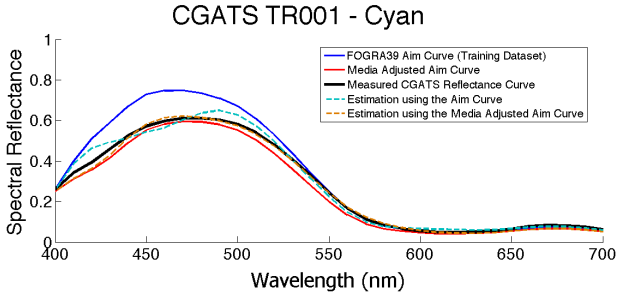

Fig.10 Comparison plots between the estimates with the media normalization and without for newspaper print.

Figure 11 illustrates the different combinations of the aim curve when the original spectral reflectance of the media is unknown. The estimation shown in Figure 9 was used for the media normalization in Figure 11. The estimation that used neither constraint resulted in the largest difference to the measured reflectance. All of the estimations resulted in identical CIE XYZ values under the original viewing conditions.

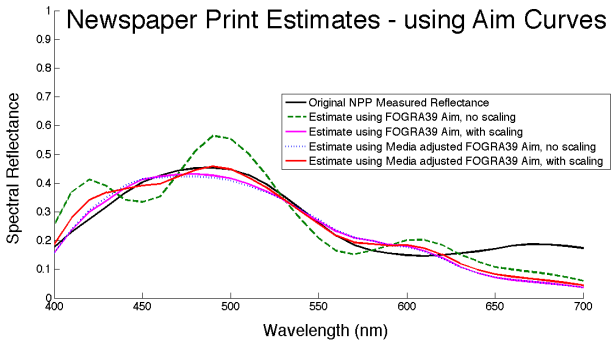

Fig. 11 Newspaper spectral estimation with scaling, without, with media normalization and without, compared with the measured reflectance.

The final Figure compares the different estimation aim curve techniques. With both medias, all aim techniques resulted in low ΔE values of less than 1. The adjusted (scale, normalization or both) aim curve usually improved the color difference value.

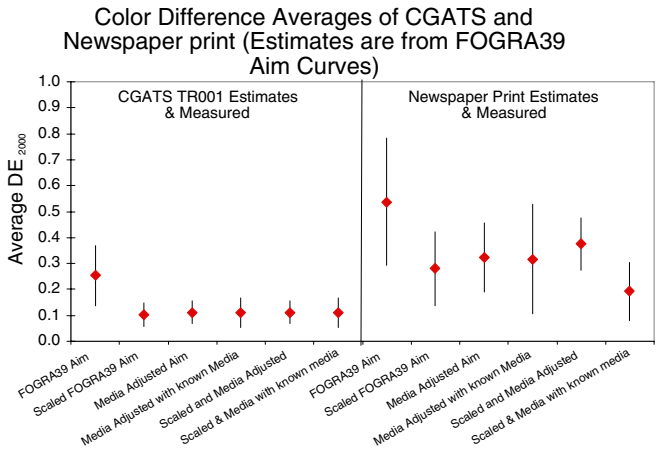

Figure 12 compares the aim curve adjustment techniques of the TR001 and Newspaper print CMYK test targets.

Into Practice - Metameric Matches

With CMYK printers, many different CMYK combinations can map to a single CIELAB value, depending on the amount of black ink used. (Fig. 13).

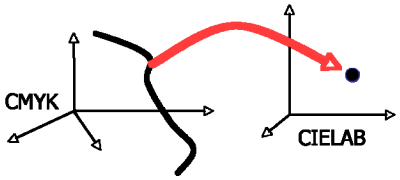

Fig. 13 different combinations of CMYK percentages can all result in the exact same CIELAB value.

In the final test, CMYK values were found that were had the same CIELAB values under one viewing condition but had the maximum color difference under a second viewing condition. A pair of CMYK test patches can be used to check the viewing illumination over time. Once the pair no longer match, the illumination source needs to be replaced.

Discussion – Spectral CMS

To have a more flexible color matching system for different viewing conditions, reflective samples are better represented by reflectance values rather than the corresponding CIE XYZ or CIELAB values. In a spectral CMS, it would be the task of the CMS to convert the colors for different viewing conditions from source to destination. Not only would this enable conversions for the standard CMF sets such as the 2° and 10° observer, but color vision deficiencies could possibly be simulated as well.

Currently, the ICC approach is used worldwide for different color management applications. A spectral CMS could be an extension of the ICC colorimetric approach with additional functionality that includes a dynamic color management module (CMM) and spectral data. [10]

The CMM intelligence needs to be adopted to include conversions from spectral data to colorimetric data, as well as a number of advanced spectral operations to provide the functionality required by different applications. A dynamic CMM would have the ability to build and rebuild color look-up tables as needed. This would extend the current functionality of the CMM to include spot color processing and spectral estimation.

Spectral matches can only be obtained if more data is exchanged between profiles. Typically, the CMM should be extended to support color information such as spectral data and the percentage of black colorant in CMYK-CMYK conversions. The color reproduction under different viewing conditions would be more consistent and the unwanted effects of metamerism would be reduced.

A prerequisite for a spectral CMS is the accessibility of the spectral data. In the ICC approach the profile is the logical container of the spectral information. Both the spectra viewing conditions and the spectral color measurements need to be stored in public ICC tags; however, in a number of cases this data is not available.

Conclusions

The spectral CMS framework extends the colorimetric approach to include spectral data. The spectral data will offer the user more flexibility. The CMM must be dynamic in that it is capable of rebuilding transforms for different viewing conditions and making spectral estimates from CIE XYZ or CIELAB color information.

There are theoretical color differences of up to 20 between colors computed under a CIE D50 illuminant and a D50 simulator. An actual 4-ink CMYK printer can produce color differences of 2.04 ΔE_{ab}.

The use of quadratic programming and a scaled aim curve (of similar printing technology) results in very small color difference of less then 1 ΔE_{ab}. Scaling the aim reduces or maintains the already small color differences. The media normalization is a very useful constraint if the spectral data of the media is available. The spectral data would be easy to include in a traditional CMS and would greatly improve estimation results. This option should be included in a dynamic CMM.

Some of the applications of this work might include: proofing, spot color rendering and metameric matches.

References

[1] ISO 12642-2:2006, Graphic technology - Input data for characterization of 4-colour process printing - Part 2: Expanded data set

[2] Derhak, M. and Rosen, M., (2006), Spectral Colorimetry using LabPQR – An Interim Connection Space, Journal of Imaging Science Technology, 50, pp. 53 – 63

[3] P. Morovic, Metamer Sets, Ph.D. thesis, University of East Anglia, Norwich, UK, 2002.

[4] Kraushaar A. (2006), DNS – "Drucken nach Standard", http://www.fogra.org/research-de/dns/home/
[5] CGATS, Recommended Industry Practice, Color characterization data set development – Analysis and reporting, 2006
[6] G. Wyszecki, Evaluation of Metameric Colors, Journal of Optical Society of America AS, vol. 48, Issue 7, pp.451-452, 1958.
[7] H. Hindi, A Tutorial on Convex Optimization, American Control Conference, Proc. of the 2004, Volume 4, issue 30, pp. 3252-3265, 2004.
[8] C. van Trigt, Smoothest reflectance functions. I. Definitions and main results, journal of Optical Society America A, vol. 7, pp. 1891-1904, 1990.
[9] P. Urban and A. Kraushaar, Deriving Optimal Spectra for Standard Offset Primaries, 34th International Research Conference of iarigai Advances in Printing and Media Technology, Grenoble, France, 2007.
[10] ICC, Draft of Color Management Implementation Classification, White Paper #25, 2008.

Author Biography

Kristyn Falkenstern has just completed her MSc at the London College of Communication. Her recent research has concentrated on various methods of finding spectral reflectance estimates.

Prior to LCC, Kristyn worked for 6 years in an Image Science team at the Vancouver R&D facility of Hewlett-Packard. Her primary responsibilities were designing, implementing, and improving image quality testing methods for the entire HP enterprise.

Kristyn holds a BS in Imaging & Photographic Technology from RIT.