Optimal Linear RGB-to-XYZ Mapping for Color Display Calibration

Brian Funt, Roozbeh Ghaffari and Behnam Bastani
School of Computing Science, Simon Fraser University
Vancouver, Canada

Abstract

Color display calibration, in part, involves mapping input RGB values to corresponding output values in a standardized color space such as CIE XYZ. A linear model for RGB-to-XYZ mapping is based on a 3-by-3 linear transformation matrix *T* mapping data from (linearized) RGB to XYZ. Such a mapping is often determined by least squares regression on the difference between predicted and measured XYZ values. However, since displays are calibrated for viewing by human observers, it likely would be better to optimize relative to a perceptually uniform color space. Two new methods are proposed which optimize the total error relative to CIELAB or CIEDE2000. The first method uses weighted least squares with weights based on the rate of change of CIELAB coordinates as a function of change in XYZ. The second method uses Nedler-Mead nonlinear optimization to minimize directly in CIELAB or CIEDE200. Experiments based on calibrating 2 CRT monitors, 3 LCD monitors and 2 LCD projectors show significantly better results than the standard least squares calibration.

Introduction

To make effective use of a CRT or LCD monitor it needs to be color calibrated in the sense that the color displayed at a given pixel can be reliably predicted for each possible input. In what follows we will assume that all pixels behave identically, that the input will be called RGB, that the output will described in terms of CIE XYZ, and that there is no interaction between the RGB channels. We further assume that any nonlinearity, in the form of a gamma correction function, say, has been removed. In other words, that the output intensity varies linearly as a function of the input RGB. In this context, the calibration problem then involves finding a mapping from linearized RGB to XYZ.

A linear model for RGB-to-XYZ mapping is based on a 3-by-3 linear transformation matrix *T* mapping data from (linearized) RGB to XYZ. In the typical display calibration procedure, many samples of the RGB input space are sent to the display and corresponding XYZs are measured from the screen. Since there are many more RGB-XYZ pairs than the 9 entries of *T,* its solution is over constrained. One common way[8] of determining *T* is to solve for the *T* that minimizes the least squares error between predicted XYZ values and the corresponding measured values. Although the least squares method works well, it suffers from the fact that the quantity being minimized—namely, distance in XYZ—is not the quantity that matters most—namely, distance in a perceptually uniform color space such as CIELAB ΔE [5] or CIEDE2000.[3]

We propose two new methods for calculating *T.* Each is based on more directly optimizing the error measure that we would truly like to minimize. The first extends the standard least squares solution to a weighted least squares solution in which the weights are defined (in an intuitive sense) to be inversely proportional to the approximate size of the MacAdam color discrimination ellipse that would surround each XYZ. The regions of XYZ space in which the ellipses are small correspond to places where human subjects are sensitive to color differences, so the method places more importance on minimizing the XYZ error there and less importance elsewhere. Since the weights applied in the weighted least squares directly relate to CIELAB $L^*a^*b^*$, we will refer to this method as LabLS and the original unweighted least squares method as ULS.

The second method is based on optimizing the ΔE error directly using the Nelder-Mead simplex algorithm.[9] Unfortunately, the Nelder-Mead algorithm (not to be confused with the linear programming simplex algorithm) finds a local minimum and does not guarantee a globally optimal solution. However, by using the results of the LabLS method as the starting condition, we have found experimentally that it converges reliably to excellent solutions. This method will be referred to as DEM since it is based on a ΔE minimization.

Unweighted Least Squares (ULS)

The unweighted least squares method is included for comparison. Before applying standard least squares fitting to calculate *T,* the measured XYZ are first normalized to the 0-1 range. In other words, the XYZ corresponding to RGB = (0,0,0) is mapped to (0,0,0) and similarly RGB = $(R_{max},G_{max},B_{max})$ is mapped to XYZ = (1,1,1). *T* is then calculated via the pseudo-inverse using equation (3) where *W* is the identity matrix of appropriate size, *N* is an nx3 matrix containing the linearized input RGBs and *M* is an nx3 matrix containing corresponding normalized tristimulus values XYZ.

CIELAB-weighted Least Squares (LabLS)

The weights for the LabLS method are based on human sensitivity to color differences as encapsulated in CIE $L^*a^*b^*$ ΔE. The idea is to evaluate the rate of change in $L^*a^*b^*$ as a function of change in XYZ. A given (X,Y,Z) maps to a specific $L^*a^*b^*$ and for an incremental change to (X+dX, Y+dY, Z+dZ) there is a corresponding change in $L^*a^*b^*$. Since this is in 3-dimensions, it is natural to measure the change in volume. Since the underlying ΔE varies only as the cube root of the volume of the $L^*a^*b^*$ change, we base the weights on the cube root of the volume change. Whether the volume is increasing or decreasing does not matter, so the sign of the change is ignored. This is formalized in Equation (2) as the cube root of the absolute value of the Jacobian determinant.

$$W(X,Y,Z)=\left\|\begin{matrix}\frac{\partial L}{\partial X} & \frac{\partial a*}{\partial X} & \frac{\partial b*}{\partial X}\\ \frac{\partial L}{\partial Y} & \frac{\partial a*}{\partial Y} & \frac{\partial b*}{\partial Y}\\ \frac{\partial L}{\partial Z} & \frac{\partial a*}{\partial Z} & \frac{\partial b*}{\partial Z}\end{matrix}\right\|^{1/3} \qquad (2)$$

For each measured XYZ, the corresponding weight is calculated and arranged along the diagonal of a matrix *W*. *T* is then calculated using equation (3), where *M* is an nx3 matrix containing normalized XYZ tristimulus values and *N* is an nx3 matrix containing the linearized RGBs. This use of the Jacobian of $L^*a^*b^*$ is similar to that proposed by Balasubramanian[1] in the context of color printer calibration and by Sharma and Trussell[11] in the context of color scanners.

$$\begin{aligned}N_w &= W \times N\\ M_w &= W \times M\\ T &= (N_w^T N_w)^{-1} N_w^T M_w\end{aligned} \qquad (3)$$

ΔE Minimization (DEM)

The problem with minimizing ΔE directly is that it is not linearly related to XYZ. The LabLS method overcomes this problem by approximating the non-linearities by weights in the original linear space; however, ideally it would be better to solve the non-linear optimization problem and minimize ΔE directly. Nedler-Mead simplex[9] search is a directed search method for multi-dimensional non-linear regression. We used the Matlab function *fminsearch* to find the 9 components of matrix *T* that minimize the total color difference error. The error can be either that of CIEDE2000 or CIELAB ΔE.

The solution depends on the starting conditions. It is shown below that the Nedler-Mead simplex search outperforms other models when it starts with an initial solution found by LabLS.

Display Calibration Experiments

To test the effectiveness of the two proposed methods, LabLS and DEM, in comparison to standard least squares, we calibrated 2 CRT monitors, 3 LCD monitors and 2 LCD projectors (see Table 1). Data was collected using a Photo Research SpectraScan 650 spectroradiometer in a dark room with the spectroradiometer at a fixed distance (1/2 meter for monitors and 2 meters for projectors) and set perpendicular to the center of the display surface. Before beginning each test, the display settings were re-set to their factory defaults. The display brightness was then adjusted using a gray-scale calibration pattern until all shades of gray were visible.

The data collection was based on a test suite involving randomization and repetition, as suggested by Cressman et al.,[2] and included long integration times for the CRT monitors. The entire input RGB space was sampled on a 10-by-10-by-10 grid. For each input RGB, the spectrometer was used to measure the XYZ of the color displayed.

Table 1. Display Devices for Calibration Testing

Name	Description
CRT1	Samsung Syncmaster 900NF
CRT2	NEC Accusync 95F
LCD1	IBM 9495
LCD2	NEC 1700V
LCD3	Samsung 171N
PR1	Proxima LCD Desktop Projector 9250
PR2	Proxima LCD Ultralight LX

Each of the RGB channels must first be linearized so that the output intensity is linearly related to the input. There are several techniques for representing the linearization function. These include GOG, S-Curve, Polynomial model and look-up tables. The results here are based on look-up tables determined by principal components analysis of the response of each channel followed by linear regression.[2] The look-up tables are also offset and scaled so that zero input maps to black and maximum input maps to white. In other words, RGB = (0,0,0) implies XYZ = (0,0,0) and RGB = $(R_{max},G_{max},B_{max})$ implies XYZ = (1,1,1).

Results

The ULS, LabLS and DEM methods of computing the RGB-to-XYZ transformation matrix, *T,* were applied to the 1000 measured, linearized and normalized RGB-XYZ pairs. In each case, the resulting *T* is evaluated by using it to map the

1000 RGB inputs to XYZ and measuring the difference between the predicted and measured values. The difference is calculated in terms of the average CIEDE2000 and CIELAB $\Delta E_{L^*a^*b^*}$ error measures.

Table 2 shows the performance of the three linear models based on CIEDE2000. Table 3 gives the error in CIELAB $\Delta E_{L^*a^*b^*}$ In both tables, 'mean,' 'stdev' and 'max' are the average error, standard deviation and maximum error over the 1000 test samples. The percentage improvement in error relative to unweighted least squares is labeled 'change'.

None of the methods requires more that a few seconds of computer time to solve for *T*.

Table 2. Calibration Error in CIEDE2000

		ULS	LabLS	DEM
CRT1	mean	1.01	0.95	0.89
	change		5%	11%
	max	3.17	2.89	2.7
	stdev	0.49	0.44	0.41
CRT2	mean	0.78	0.76	0.69
	change		2%	12%
	max	3.12	3.08	3.09
	stdev	0.41	0.46	0.36
LCD1	mean	0.44	0.43	0.42
	change		3%	5%
	max	3.12	3.00	3.04
	stdev	0.36	0.35	0.36
LCD2	mean	1.35	1.32	1.26
	change		3%	7%
	max	4.29	4.17	4.22
	stdev	0.66	0.67	0.68
LCD3	mean	1.59	1.52	1.48
	change		4%	7%
	max	5.13	5.09	5.09
	stdev	0.71	0.74	0.78
PR1	mean	0.64	0.59	0.58
	change		8%	10%
	max	3.46	3.64	3.62
	stdev	0.43	0.43	0.47
PR2	mean	0.87	0.83	0.82
	change		5%	6%
	max	2.67	3.11	3.14
	stdev	0.46	0.42	0.43

White Point Preservation

White point preserving color correction minimizes the least square error in tristimulus space while constraining white to be predicted precisely.[8] The white point is preserved in a constrained least-squares formula with a Lagrange multiplier term on the error in white. Table 4 shows how white varies from the ideal when the white point is not constrained.

Table 3. Calibration Error in CIELAB ΔEL*a*b*

		ULS	LabLS	DEM
CRT1	mean	2.14	1.96	1.70
	change		8%	20%
	max	8.54	7.15	4.57
	stdev	1.39	1.16	0.90
CRT2	mean	2.27	1.84	1.57
	change		19%	31%
	max	17.77	8.97	4.71
	stdev	2.10	1.22	0.76
LCD1	mean	0.86	0.82	0.79
	change		4%	8%
	max	3.21	3.03	2.95
	stdev	0.55	0.53	0.55
LCD2	mean	2.86	2.73	2.53
	change		5%	12%
	max	7.48	7.83	7.88
	stdev	1.44	1.41	1.39
LCD3	mean	3.65	3.28	3.10
	change		10%	15%
	max	7.25	8.42	9.45
	stdev	1.43	1.44	1.67
PR1	mean	1.73	1.54	1.46
	change		11%	16%
	max	7.60	7.14	9.63
	stdev	1.26	1.27	1.55
PR2	mean	1.98	1.91	1.88
	change		4%	5%
	max	6.29	7.91	9.11
	stdev	1.13	1.16	1.27

In a manner similar to the ULS case, LabLS and DEM can easily be extended to enforce a constraint on white (or any other color, if need be). Simply setting the weight on white in either the LabLS or DEM to equal the sum of all other colors forces white to be pure white to within 4 decimal digits. For some applications a user may prefer having some specific colors calibrated more accurately than others. Such considerations can easily be added to LabLS and DEM by including an extra matrix of weights expressing the relative importance of different colors. Tables 5 and 6 list the CIEDE2000 and CIELAB $\Delta E_{L^*a^*b^*}$ average error in the calibration when white point preservation is included. Note that in addition to the drop in the average error, often the maximum error drops very significantly with the new methods.

Table 4. White Prediction for the Three Methods.

		X Y Z
ULS	CRT1	1.0825,1.0784,1.0755
	CRT2	1.0675,1.0672,1.0711
	LCD1	0.9989,0.9984,1.0044
	LCD2	0.9302,0.9270,0.9561
	LCD3	1.0069,1.0007,1.0235
	PR1	1.0074,1.0029,1.0137
	PR2	0.9952,0.9910,1.0201
LabLS	CRT1	1.1319,1.1276,1.1246
	CRT2	1.0479,1.0487,1.0469
	LCD1	1.0078,1.0083,1.0108
	LCD2	0.9986,0.9946,1.0248
	LCD3	1.0479,1.0447,1.0643
	PR1	1.0034,1.0031,1.0185
	PR2	0.9869,0.9879,1.0232
DEM	CRT1	1.1276,1.1235,1.1229
	CRT2	1.0523,1.0514,1.0485
	LCD1	1.0073,1.0074,1.0098
	LCD2	1.0077,1.0016,1.0439
	LCD3	1.0567,1.0546,1.0762
	PR1	1.0046,1.0043,1.0207
	PR2	0.9850,0.9857,1.0267

Table 5. Calibration Error in CIEDE2000 with White Point Preservation

		WP_ULS	WP_LabLS	WP_DEM
CRT1	mean	3.37	3.29	2.91
			2%	14%
	max	11.38	8.14	7.01
	stdev	1.47	1.53	1.36
CRT2	mean	1.89	1.74	1.43
			8%	25%
	max	11.25	5.89	3.21
	stdev	1.34	0.87	0.52
LCD1	mean	0.47	0.47	0.47
			-1%	0%
	max	3.07	3.04	3.06
	stdev	0.36	0.36	0.36
LCD2	mean	1.57	1.50	1.46
			5%	7%
	max	4.24	4.01	4.26
	stdev	0.81	0.73	0.71
LCD3	mean	2.22	2.15	1.84
			3%	17%
	max	10.34	6.49	5.15
	stdev	1.52	1.04	0.86
PR1	mean	0.89	0.77	0.67
			13%	25%
	max	4.97	4.41	4.39
	stdev	0.64	0.48	0.48
PR2	mean	1.41	1.32	1.18
			6%	16%
	max	6.57	3.48	3.01
	stdev	0.84	0.58	0.55

Table 6. Calibration Error in CIEDE2000 with White Point Preservation

		WP_ULS	WP_LabLS	WP_DEM
CRT1	mean	7.41	6.97	3.56
			6%	52%
	max	36.93	22.39	6.30
	std	4.06	3.40	1.42
CRT2	mean	4.65	4.02	2.59
			14%	44%
	max	36.72	17.88	19.35
	std	4.56	2.87	1.87
LCD1	mean	0.86	0.88	0.83
			-2%	3%
	max	2.89	2.85	3.03
	std	0.58	0.56	0.56
LCD2	mean	3.34	3.10	2.98
			7%	11%
	max	8.50	6.74	8.05
	std	1.53	1.29	1.35
LCD3	mean	4.94	4.69	3.60
			5%	27%
	max	36.69	22.25	8.94
	std	5.01	3.11	1.56
PR1	mean	2.67	2.23	1.70
			16%	36%
	max	19.12	10.58	8.94
	std	2.31	1.66	1.36
PR2	mean	3.77	3.43	2.49
			9%	34%
	max	26.14	16.51	12.18
	std	3.27	2.20	1.40

Conclusion

The performance of a 3x3 linear color calibration model can be improved by optimizing for the transformation matrix in spaces other than XYZ. One alternative (DEM) is to minimize directly in CIELAB space, but this involves nonlinear optimization. Another alternative (LabLS) is to optimize using weighted least squares regression in a CIELAB-weighted version of XYZ. Experiments in calibrating with 5 different displays show that both methods significantly reduce calibration errors as measured in terms of average and maximum CIELAB ΔE or CIEDE2000 error.

Acknowledgement

Funding was provided by the Natural Sciences and Engineering Research Council of Canada.

References

1. Balasubramanian, R., Color Transformations for Printer Color Correction, *Second Color Imaging Conference: Color, Science, Systems and Applications*, IS&T/SID, pp.62-65, Nov. 1994.
2. Cressman, W., Bastani, B., and Funt, B., "Calibrated Colour Mapping Between LCD and CRT Displays: A Case Study" *Proc. Second European Conference on Color in Graphics, Imaging and Vision*, Achen, April 2004
3. CIE, Improvement to Industrial Color-Difference Evaluation Publication, *CIE Publication No. 142-2001*, Central Bureau of the CIE, Vienna, 2001
4. [CIEDE2004] Sharma, G., Wu, W., Dalal, E. N., The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations, Submitted to *Color Research and Application,* 2004
5. Wyszecki, W.S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, Second Edition, 1982, pg. 167.
6. Fairchild MD, Wyble DR. Colorimetric Characterization of the Apple Studio Display (Flat Panel LCD). *Munsell Color Science Laboratory Technical Report*, 1998, http://www.cis.rit.edu/mcsl/research/PDFs/LCD.pdf
7. [Day2002] Day E. Colorimetric Characterization of a Computer Controlled (SGI) CRT Display, *Munsell Color Science Laboratory Technical Report*, 2002, http://www.art-si.org/PDFs/CRT_Colorimetry_Ap_2002.pdf
8. Finlayson G.D., Drew M.S., White-Point Preserving Color Correction, *Fifth Color Imaging Conference: Color, Science, Systems and Applications*, IS&T/SID, pp.258-261, Nov. 1997
9. Lagarias, J.C., Reeds, J. A., Wright, M. H. and Wright, P. E., "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions," *SIAM Journal of Optimization*, Vol. 9 Number 1, pp. 112-147, 1998.
10. [MAC1942] MacAdam DL. Visual sensitivities to color differences in daylight. *J. Opt Soc Am* 1942;32:247–274.
11. Sharma, G. and Trussell, J.H., "Figures of merit for color scanners," *IEEE Trans. Image Processing*, Vol. 6, No. 7, pp.990-1001, July 1997.