Measuring the Quality of ICC Profiles and Color-Management Software

By Abhay Sharma and Paul D. Fleming

Accurate device profiles are the bedrock upon which all color-managed workflows rest. We describe some repeatable, quantitative experiments for measuring and comparing the accuracy of several mainstream profiling products.

The growth in color management means that there are now many different software packages that can make ICC profiles. But how do we know which is the most accurate and which represents the best value? There is no agreed standard among the software vendors, and most users are unable to independently assess the quality of profiles. The WMU Profiling Review is a "consumer report" that provides an independent, objective assessment of current software products and ranks them according to their colorimetric accuracy. An ongoing project of Western Michigan University's Department of Paper and Printing Science and Engineering, the Review is released biannually on January 1 and July 1.

What we tested

In this edition of the Review, we made profiles for the following types of devices: scanner, printer, LCD and CRT. The products covered in this review are:

- Fujifilm ColourKit ProfileMaker
- GretagMacbeth Eye-One Match
- GretagMacbeth ProfileMaker
- ICS basICColor
- Monaco EZColor
- Monaco Profiler
- X-Rite ColorShop

Generic profiles and Apple's Display Calibrator utility were also tested. The following devices were used: Umax Astra 4000u (scanner), Apple Studio Display (17-inch flat panel), Mitsubishi Diamond Plus (CRT) and Epson Stylus Pro 5000 (ink-jet printer). In most instances, testing was done with Mac OS 10.2.2, ColorSync 4.1, Photoshop 7.0.1 and the ACE CMM.

Testing philosophy

It is important to have a quality measure for ICC profiles, because this indicates how well a device has been characterized and therefore how accurate the color is likely to be in a color-managed workflow. It is important for software vendors to publish a merit figure and for the industry to agree on how the figure is calculated. Some vendors quote a delta-E (ΔE) merit figure and, often, programs will write out a file with statistics. However, there is no indication to tell us how these figures are calculated and whether everybody is measuring the same thing in the same way. The aim of the WMU Profiling Review is to establish some baseline assessment for ICC profiles and thus assist user choice, raise the standard of profiling software and promote the wider acceptance of ICC color management.

The assessment of ICC profiles and color reproduction is a complex issue involving everything from color science, psychophysics and image analysis to "preferred" reproduction styles. The approach adopted in this work is to evaluate the accuracy of profiles using the colorimetric intent. This does not provide an all-encompassing result, but it does provide an indicative set of metric figures that can be used to make valid cross-vendor comparisons. In the future, we aim to encompass subjective aspects of image reproduction; in such instances, the perceptual rendering intent would be used.

It is inevitable in a survey of this type that some vendors fared better than others, but this should not be taken as an endorsement of any particular product or manufacturer. Due to the nature of the tests, some tables are listed in order of merit (with the best first), while other tables are in alphabetical order.

The test procedure for each category of profile is described in the report. Enough detail is provided for a skilled user to replicate our results. For each product tested, the default settings were always used; no attempt was made to alter the vendor's starting recommendations. Simple CIELAB ΔE was used in all cases.

Scanner profile results

Agfa, Fujifilm and Kodak IT8.7/2 reflection test targets were scanned on a Umax Astra 4000u scanner,

SCANNER PROFILE RESULTS (WITH THE BEST FIRST)

Scanner profile quality Umax Astra 4000u	Agfa IT8.7/2 Chart *Mean (Max) ΔE*	Fujifilm IT8.7/2 Chart *Mean (Max) ΔE*	Kodak IT8.7/2 Chart *Mean (Max) ΔE*	Final result *Average ΔE*
Fujifilm ColourKit ProfileMaker 3.0.4 (OS 10)	1.11 (4.36)	0.90 (3.52)	0.88 (4.47)	0.96
ICS basICColor 1.5 (OS 9)	1.04 (8.17)	1.12 (4.60)	1.12 (11.40)	1.09
Monaco Profiler 4.5 (OS 10)	1.25 (11.31)	0.91 (4.40)	1.19 (9.02)	1.12
GretagMacbeth Eye-One Match 1.1 (OS 10)	1.12 (3.60)	0.97 (3.90)	1.32 (4.78)	1.14
GretagMacbeth ProfileMaker 4.1 (OS 10)	1.15 (3.59)	1.12 (2.86)	1.22 (4.91)	1.16
Monaco EZColor 2.5.1 (OS 10)	1.32 (10.21)	0.95 (4.37)	1.20 (7.61)	1.16
Generic Umax scanner profile1	29.80 (44.55)	28.93 (42.03)	29.38 (46.67)	29.37

[1]The generic profile was obtained as part of the Umax scanner driver, Umax VistaScan 3.5.4.

and profiles were made in different profiling packages using batch reference data.

Tests to measure the accuracy of the scanner profile were conducted as follows. Following profile generation, the raw scan of each IT8.7/2 chart image was opened in Photoshop. Each scanner profile was selected in turn (using Image>Mode>Assign Profile) and the image was processed to LAB using Image>Mode>Convert to Profile where the Destination Space was chosen as LAB Color. The rendering intent chosen was Absolute Colorimetric, and the CMM used was ACE. The LAB value of each patch in the chart image was calculated, and a ΔE was computed between this value and the original reference value used in profile generation. Finally, the mean and maximum ΔE were calculated over all patches of the IT8.7/2 chart.

Interpretation. The accuracy of each vendor's program is shown in the table. A lower ΔE number is preferable.

SCANNER PROFILE EVOLUTION

Scanner profile evolution	Agfa IT8.7/2 Chart *Mean (Max) ΔE*	Fujifilm IT8.7/2 Chart *Mean (Max) ΔE*	Kodak IT8.7/2 Chart *Mean (Max) ΔE*	Final result *Average ΔE*
Fujifilm ColourKit 3.0	1.11 (4.36)	0.90 (3.52)	0.88 (4.47)	0.96
Fujifilm ColourKit 2.3	1.15 (3.72)	1.23 (4.53)	1.43 (3.53)	1.27
Fujifilm ColourKit 2.2	1.17 (3.98)	1.25 (4.53)	1.42 (3.66)	1.28
Gretag ProfileMaker 4.1	1.15 (3.59)	1.12 (2.86)	1.22 (4.91)	1.16
Gretag ProfileMaker 4.0	0.85 (2.87)	0.99 (10.13)	1.23 (4.12)	1.02
Gretag ProfileMaker 3.1	0.85 (2.59)	0.97 (3.21)	1.16 (3.30)	0.99
Monaco Profiler 4.5	1.25 (11.31)	0.91 (4.40)	1.19 (9.02)	1.12
Monaco Profiler 4.0	1.19 (9.95)	0.92 (4.70)	1.19 (7.10)	1.10
Monaco Profiler 3.2	4.39(15.00)	5.04 (8.25)	4.79 (11.35)	4.74

Manufacturers are ranked in order; Fujifilm provided the best overall result, while the generic profile was worst. There is very little difference between the custom-generated profiles; with a ΔE less than 2, they are very accurate scanner profiles.

In color-management circles, it is often asked: How good is the generic profile supplied by the manufacturer? For this scanner, the generic profile, with a ΔE of nearly 30, was very poor. Note that this does not mean that the Umax scanner is poor. In fact, the scanner is a remarkably good value; the ΔE value merely tells us how well the profile has characterized the scanner.

In each case, the maximum ΔE should also be considered. The ideal program would have both a low average and a low maximum. In these tests, Fujifilm ColourKit, GretagMacbeth Eye-One Match and GretagMacbeth ProfileMaker had a low mean and a low maximum ΔE. The Monaco and ICS products had low averages, but somewhat higher maxima.

It is important that each profiling package be able to make an acceptable profile with the Agfa, Fujifilm or Kodak charts. In a few cases, the results were different across the chart types.

Though not shown in the table, the cost of the profiling packages should also be considered. Monaco EZColor has a high ranking and a very competitive retail price.

The results suggest that some vendors may be using the same core algorithms for the consumer and professional versions of their software. For example, note that GretagMacbeth's Eye-One Match and ProfileMaker produce similar results; also Monaco's EZColor and Profiler are very similar. We could conclude that these companies are using the same code in both their products.

The long view. Using data from previous versions of this review, it is possible to conduct some historical analysis. We can see if software is getting better over time and, perhaps, if there any benefit in paying for an upgrade. From the data in the table below, we could conclude that Fujifilm ColourKit was not changed between versions 2.2 and 2.3 but was improved in version 3.0. GretagMacbeth's ProfileMaker has remained the same between versions 3.1 and 4.0, and actually appears to be slightly less accurate in version 4.1. Monaco Profiler was greatly improved between versions 3.2 and 4.0, but has remained the same since. Of course, there may be improvements in these products that are not detected by our tests. Also, in scanner profiling, it is possible to get slightly different results each time the experiment is conducted, as the IT8 chart may be cropped differently.

Printer profile results

To test output profiling, we used an Epson Stylus Pro 5000 ink-jet printer, driving it in CMYK mode using a

Fiery RIP. In each case, the output profile was made using the vendor's proprietary chart. A GretagMacbeth SpectroScan was used in all programs except where shown. Default values were used in each program for all settings of black generation and profile quality. Next, the basic subset of an IT8.7/3 chart (182 patches) was printed and measured. This provided a set of in-gamut LAB values. This is the data that was used in the subsequent test process.

Printer profiles contain three rendering intents: perceptual, colorimetric and saturation. Each intent has a forward (Profile Connection Space to Device) lookup table and a reverse (Device to Profile Connection Space) lookup table. Our tests evaluated the combined accuracy of the absolute colorimetric intent. To start the process, the LAB values of the IT8.7/3 basic chart were put in an image and, in Photoshop, the LAB image was converted to CMYK. The ACE CMM was used, and the intent selected in Image>Mode>Convert to Profile was Absolute Colorimetric. Next, the CMYK image was converted back to LAB using Image>Mode>Convert to Profile (LAB Color). Again, the ACE CMM was used and the intent was Absolute Colorimetric. Finally, the LAB value of each patch in the resulting file was compared to the starting LAB value, and the ΔEs were averaged over all 182 patches.

The test took in-gamut LAB values and converted them to CMYK and then back again to LAB. Thus, the test gives us an indication of the accuracy of the colorimetric lookup table in the output profile. This test is often referred to as "round tripping." There are two advantages of this approach. One is that it uses separate data for training (*i.e.*, making the profile) and testing; the other is it can be conducted entirely in software. The disadvantage is that it does not separately measure the accuracy of the forward and reverse parts of the output profile.

PRINTER PROFILE RESULTS (WITH THE BEST FIRST)

Printer profile quality Epson Stylus Pro 50001	Instrument	Mean (Max) ΔE
ICS basICColor 1.5 (OS 9)	GretagMacbeth SpectroScan	1.45 (4.41)
Fujifilm ColourKit ProfileMaker 3.0.4 (OS 10)	GretagMacbeth SpectroScan	1.99 (8.65)
GretagMacbeth ProfileMaker 4.1(OS 10)	GretagMacbeth SpectroScan	2.01 (5.06)
GretagMacbeth Eye-One Match 1.1(OS 10)	GretagMacbeth Eye-One	2.29 (11.38)
Monaco Profiler 4.5 (OS 10)	GretagMacbeth SpectroScan	2.35 (6.36)
Generic Profile[2]	Unknown	3.58 (12.31)
Monaco[3] EZColor 2.5.1 (OS 10)	HP Scanjet 7400c	5.44 (27.38)

[1] Epson Stylus Pro 5000 using Fiery RIP SPv1.3 in CMYK mode with Epson Photo Paper, batch Y1JL0U744.
[2] Generic profile downloaded from www.cgs.de/de/icc.html.
[3] Monaco EZColor does not require a conventional measuring instrument.

Interpretation. The results produced by our vendors, with an average ΔE less than 3, are very good and are likely to produce excellent results in all printer-based workflows. However, keep in mind that a large maximum ΔE has the potential to cause problems in particular image colors.

While photographic images are normally processed using the perceptual intent, the colorimetric intent is used during the facsimile reproduction of

Version History

Updates to the WMU Profiling Review occur bi-annually on January 1 and July 1.

- Version 1.0 (released April 1, 2002). Covered scanner profiles and monitor profiles.
- Version 1.1 (April 9, 2002). Maintenance release to amend one numerical result for Monaco monitor profiling.
- Version 2.0 (July 1, 2002). Scanner profiles: Fujifilm ColourKit updated from 2.2 to 2.3 and Heidelberg ScanOpen updated from 2.1.0 to 4.0.5; ColorBlind 4.2 added to the study. Where appropriate, data for old and new versions of software were quoted. Printer profiles were new in this release; they were made using GretagMacbeth ProfileMaker 4.0, Fujifilm ColourKit 2.3, Monaco Profiler 4.0, Kodak Colorflow 2.2, ColorSynergy 4.5 and Heidelberg PrintOpen 4.0.5. Fading test data were presented for the HP Designjet 20ps and Epson Stylus Pro 5000 ink-jet printers. Monitor profile data was unchanged from Version 1.1. The product from Color Solutions called basICColor arrived too late to be included in this review. The way the prices are quoted was improved in this version of the review.
- Version 3.0 (January 1, 2003) is described in this article. New products in this version are GretagMacbeth Eye-One Match, ICS basICColor and Monaco EZColor. Other products are updated to reflect OS X versions of the software. New in this version is analysis of profiles for LCD panels. CRT monitor profile data is unchanged from Version 1.1. Prices are no longer being quoted.

To get the WMU Profiling Review, send an e-mail to abhay.sharma@wmich.edu with "WMU Profiling Review" in the subject line. You will be sent the latest version, and your name will be added to a mailing list for future updates. There are currently 500 addresses on the mailing list. TSR

LCD MONITOR RESULTS (IN ALPHABETICAL ORDER)

Apple Studio Display Monitor profile quality	Measuring instrument	Achieved gamma (Target was 1.8)	E difference in white point from a target of D50	Average ?E of 24 patch Macbeth ColorChecker
Apple Display Calibrator 4.1	Visual	1.79	5.74	5.32
Fujifilm ColourKit ProfileMaker 3.0.4	GretagMacbeth Eye-One	1.76	1.08	1.76
Generic Profile Apple Studio Display	None	1.81	17.78	12.93
GretagMacbeth[1] ProfileMaker 4.1.1	GretagMacbeth Eye-One	1.71	2.37	3.46
GretagMacbeth Eye-One Match 1.3	GretagMacbeth Eye-One	1.77	2.3	3.78
ICS [2] basICColor 1.5	GretagMacbeth Eye-One	1.77	1.26	5.26
Monaco EZColor 2.5.1	MonacoOPTIX	1.83	6.40	4.07
Monaco Profiler 4.5	GretagMacbeth Eye-One	1.84	1.39	2.24

[1] It was necessary to update to Mac OS 10.2.3 and ProfileMaker 4.1.1 to achieve this result. Very poor results were obtained with Mac OS 10.2.2 and ProfileMaker 4.1.

[2] ICS display profiles contain some gamut mapping, which this test was not designed to measure, so it is likely that the basICColor profile could produce better results than shown here. To avoid this confusion, it is recommended that vendors populate lookup tables in complete accordance with the ICC specification.

images, during soft proofing when images are evaluated on a monitor, and during proofing when press images are "returned" to the Profile Connection Space and printed on a proofing device. The colorimetric intent may also be used when legacy CMYK images are opened. So although the colorimetric intent is not normally used to process photographic images, it is used in a number of significant ICC workflows and, as such, is worth testing. Besides, it is an easily calculated profile-accuracy measurement. Nonetheless, the appearance of images is an important criterion that should also be considered in such tests.

LCD monitor results

LCD panels are becoming increasingly important in color management. Are profiling instruments and profiling packages able to accurately characterize these devices? A number of profiling packages were tested to see if they were able to achieve a requested gamma, a requested white point and an accurate reproduction of a Macbeth ColorChecker chart.

A 17-inch flat-panel Apple Studio Display was used on a Power Mac G4. Monitor profiles were made using different measuring instruments, as shown in the table. The user requested a gamma of 1.8 and white point of D50. After each profile was made, it was selected as the system profile.

Using Photoshop, a grayscale ramp, consisting of RGB (0, 0, 0), (15, 15, 15) and so on up to (255, 255, 255), was built and displayed on the monitor. The luminance (Y) was measured using a GretagMacbeth Spectrolino spectrophotometer, and this was used to determine the actual gamma of the display. Next a white patch (RGB 255, 255, 255) was displayed and the XYZ value of this patch was measured. The measured XYZ was then normalized to Y=100 (the color temperature is unchanged by a uniform rescaling of the XYZ values). The measured XYZ was converted to LAB for the chosen illuminant, D50, and compared to an ideal D50 white point that has an LAB of (100, 0, 0). A $\Delta E_{a,b}$ calculation was done to establish how closely each profile was able to create the requested color temperature. A $\Delta E_{a,b}$ figure was defined as:

$$\Delta E_{a,b} = (a^2 + b^2)^{0.5} = C$$

Thus, we see that the $\Delta E_{a,b}$ has a simple interpretation as the chroma, C, of the measured white point, referenced to the target white point.

From the Munsell Web site (www.munsell.com), the XYZ values for a 24-patch Macbeth Digital ColorChecker were downloaded. These were converted to LAB (D50) and displayed in Photoshop. A Spectrolino was then used to measure the LAB of the patches, and the ΔE was calculated and averaged over the 24 patches.

Interpretation. The results show that all of the products in our test can create the requested gamma for the display. In most cases, the small difference from the expected gamma of 1.8 is not significant. For the white point, a ΔE of up to 2 is very good. For the colors of the Macbeth ColorChecker, a lower ΔE is better, and a ΔE of 3 or less is likely to produce good results. Keep in mind that some colors may be out of gamut of the display.

ICC profiles nowadays can contain different lookup tables for different rendering intents—A2B0

*A number of vendors create profiles that are not in complete compliance with the latest ICC specification. This may be intentional (ICS basICColor and Creo Profile Wizard) or due to older software that has not been changed (Kodak Colorflow). Using such profiles can cause unintentional changes to an image.

For Further Reading

A. Sharma, M.P. Gouch, and D.N. Rughani, "Generation of an ICC profile from a proprietary style file," *J Imag Sci Tech*, 46:26, 2002.

A. Sharma and P.D. Fleming, "Evaluating the quality of commercial ICC color management software," presented at TAGA Annual Technical Conference, North Carolina, April 11–14, 2002.

A. Sharma, "A procedure to evaluate the accuracy of ICC profiles," presented at the ICC meeting, Zurich, June 11–14, 2002.

P.D. Fleming and A. Sharma, "Color management and ICC profiles," 56, *Gravure Magazine*, August 2002.

A. Sharma, "Buying color management software," presented at GATF Color Management Conference, December 8–10, Phoenix, AZ, 2002. **TSR**

(perceptual), A2B1 (colorimetric) and A2B2 (saturation). However, this was not always the case. In the early ICC file format specifications, scanner and monitor profiles used to have only one lookup table, called the A2B0 tag. In the 1998 specification, the A2B1 and A2B2 tags for the scanner profile were mentioned but were "undefined." In the current revision of the ICC specification (Specification ICC.1:2001-12, Version 4.0.0), the A2B0, A2B1 and A2B2 tags for all profiles are explicitly defined. All profiles can now have the A2B0, A2B1 and A2B2 tags; thus, there is no excuse for vendors to put perceptual data in the colorimetric (A2B1) tag or *vice versa*.*

CRT monitor results

CRT monitor profiles were tested to see if they were able to achieve a requested gamma and a requested white point. A Mitsubishi Diamond Plus CRT was used on a Power Mac G4. Monitor profiles were made using different measuring instruments, as shown in the table and, if offered a choice, the user requested a gamma of 1.8 and white point of D65. After each profile was made, it was selected as the system profile.

Using Photoshop, a series of patches were displayed on the monitor so that the actual white point and actual gamma of the display could be verified. The procedure and calculations were done in the same way as described above for LCD profiles.

What happens when you make a monitor profile? Generally, you stick a measuring instrument on the face of the monitor and the software displays a series of color patches. This process is measuring the inherent or factory settings of the monitor. The software then asks the user for the setting that he would like. A set of corrections is calculated that converts the factory setting to the user's desired setting, and these corrections are loaded into lookup tables in the video card's digital-to-analog circuitry. Macintosh monitor profiles (both LCD and CRT) are distinguished by the use of a vcgt tag that is used to store this correction; vcgt stands for video card gamma tag, and it has been part of the Mac OS since ColorSync 2.5.

What did the results show? In terms of the gamma value, the results fell into two camps. Profiles with a vcgt tag produced a gamma of 1.8 as requested by the user, while the generic profile without a vcgt produced a gamma of around 3.0, which is the inherent gamma of the display. This situation happens whenever "PC" monitor profiles and "Mac" monitor profiles are used together.

The difference between the color of the profiled monitor and the desired white point of D65 was also calculated for each vendor and is shown in the table.

Why is my favorite program not listed?

Attempts were made to contact all vendors and invite them to participate. No product has been refused entry into this review. If your favorite software is not listed, please contact the vendor and get them to talk to us.

CRT MONITOR RESULTS (IN ALPHABETICAL ORDER)

Mitsubishi Diamond Monitor profile quality	Measuring instrument	Achieved gamma (Target was 1.8)	DE difference in white point from a target of D65	vcgt tag
Apple ColorSync Monitor Calibrator 3.0.4	Visual	2.03	12.76	Yes
Fujifilm ColourKit ProfileMaker 2.2	X-Rite DTP92	1.77	3.35	Yes
Generic Profile Mitsubishi monitor	None	2.99	3.83	No
GretagMacbeth ProfileMaker 4.0	GretagMacbeth Spectrolino	1.81	0.58	Yes
GretagMacbeth Monaco Profiler 4.0	Spectrolino	1.85	0.68	Yes
X-Rite ColorShop 2.6.2	X-Rite DTP92	1.76	5.81	Yes

Pictographics (ColorSynergy) withdrew its product and results from the rankings. Creo (Profile Wizard) entered the review process but, having seen the preliminary results, withdrew its data prior to publication. Kodak (Colorflow) also felt that the data shown in the tables was too simplistic and was not indicative of the true quality of its product and so withdrew from the process. The following vendors were invited to participate but chose not to enter: Agfa (ColorTune), EFI (ColorProfiler), Heidelberg (ScanOpen, PrintOpen, ViewOpen), ITEC (ColorBlind) and Pantone (OptiCAL, PhotoCAL).

The testing procedure for so many different products and so many different types of profiles is a lengthy process that incurs considerable labor and materials costs. In order to meet some of these costs, vendors are charged $1,000 per year to enter. The fee is waived in certain instances. The report continues to be made available free to users.

Conclusion

Our intention in launching the WMU Profiling Review was to develop a set of standard, repeatable and quantitative measurement techniques for assessing the accuracy of ICC profiles, and thus for judging the quality of profiling tools. Then, by applying our methods over time, we could assess what progress is being made. Now, after three iterations of the process, we think that we've partially achieved that goal. A by-product of the research has been to identify technical issues and shortcomings with software products. Often, when these concerns are raised with the vendors, the vendors respond by improving their products or procedures, thus benefiting all color-management users.

Acknowledgments

We are grateful to Dr. Iris Mangelschots of GretagMacbeth for generous hardware and software donations. We are grateful to Tom Dlugos and Bill Owens of X-Rite for their support. The authors are pleased to acknowledge the help of our student, Ben Starr, in producing the results. TSR

As with any research project, it is possible to criticize our methods, premises and conclusions. We welcome your comments and your suggestions for better testing protocols. (One good way is to send a Letter to the Editor.) The science and technology of color management is growing more important, and appropriate ways of evaluating the latest technologies are part of that growth. TSR

About the authors

Abhay Sharma has a B.S. in Imaging Sciences from the University of Westminster, U.K., and a Ph.D. in Physics from King's College, London. He worked as a senior research scientist for FujiFilm Electronic Imaging before joining Western Michigan University as an associate professor in color imaging. His book, Understanding Color Management, is being published by Delmar Thomson Publishing in 2003. Dr. Sharma is a member of the ICC working group that is looking at the issue of profile quality assessment. Contact him at abhay.sharma@wmich.edu.

Paul D. Fleming has a Ph.D. from Harvard University. Dr. Fleming is director of the Digital Imaging Research Group at Western Michigan University and a member of the SWOP Review Committee.