Evaluation of Current Color Management Tools: Image Quality Assessments

Hendrik Büring [a], Patrick G. Herzog [a], Eggert Jung [b]
[a] Aachen Univ. of Technology, Technical Electronics Institute, Aachen, Germany
[b] NexPress GmbH, Kiel, Germany

Abstract

Color Management (CM) Systems are widely used in conjunction with ICC profiles to obtain Color fidelity today. For this purpose, each device is characterized using a suitable CM tool to generate an ICC profile. Those tools provide acceptable to high quality, actually.

CM tools of distinct manufacturers generate different output even if the same parameters are given. This is also the case for an externally generated measurement text file, characterizing the printing device and media, which is imported into each CM tool. Obviously the manufacturers use different techniques to fit the measurement data into a model of the device to be characterized, thus each CM tool has its own strengths and weaknesses.

This study is intended to complete our investigation of CM tools (ICC v2) with the focus set on image quality. For this purpose, human observers were asked to judge the reproductions with respect to a given original. The evaluation was carried out using natural as well as artificial images with various image content. From the visual assessment data, an interval scale was calculated using psychophysical and statistical methods.

The results are compared with and related to previous analyzations where synthetic test images were used to check the ICC profile tables with ΔEab and ΔE94 taken as the criteria.

Introduction

ICC Color Management (CM)[1] is widely used today for many color imaging applications, e.g. cross media image reproduction, desktop publishing or printing on demand. In order to minimize the differences in color reproduction, each device has to be characterized e.g. by measuring a printed testchart or some displayed test colors, respectively. The resulting measurement data is used by CM tools to generate ICC profiles.

Actually those tools provide acceptable to high quality, in contrast to an investigation carried out only two years ago.[2] As one might expect, CM tools of distinct manufacturers generate different output even if the same parameters are given.

There are three reasons for this observation:

- First, the manufacturers use different device models to fit the measurement data into,
- second, different gamut mapping algorithms are used to map the images and
- third, the perceptual rendering intent is intentionally defined in a loose manner.

Thus each CM tool has its own strengths and weaknesses.

In order to evaluate the quality of CM tools, a number of aspects can be defined, such as accuracy, consistency, smoothness or invertability.[2,3] These properties determine the image quality (IQ) which can be achieved by using an ICC profile to map images between the device color space and the profile connection space, i.e. CIELab with D50 as whitepoint.

In this study, four CM tools (designated A, B, D, E) are used. The same measurement data, depending on the printing device and media, is used for the generation of each ICC profile. The intention is to verify the validity of the previously defined IQ aspects by comparing the simulated results with human observer judgements.

Profile Testing

The IQ is checked using generated L*a*b* slices and several artificial color images as well as natural images, containing pastel as well as saturated colors taken from the ISO sRGB SCID image data.[4] To perform the test, each image is converted to L*a*b* using the standard sRGB profile and the B2A-transform of the ICC profiles sequentially.

The resulting CMYK images are then proofed on an inkjet printer with photo quality paper and inks. The resulting prints are judged in paired comparison experiments by ten observers, followed by statistical evaluation, to rank the appearance and pleasantness of the reproductions.

These evaluations are done for an digital electrophotographic output device with medium gamut size. The following charts and natural motifs were used as test images, see fig. 1:

- 'Patches' (S6), 'Vignettes' (S7) and 'Flowers' (N2), taken from ISO sRGB-SCID
- Kodak Color Evaluation Target (CET)
- Lab image 'Group', provided by NexPress

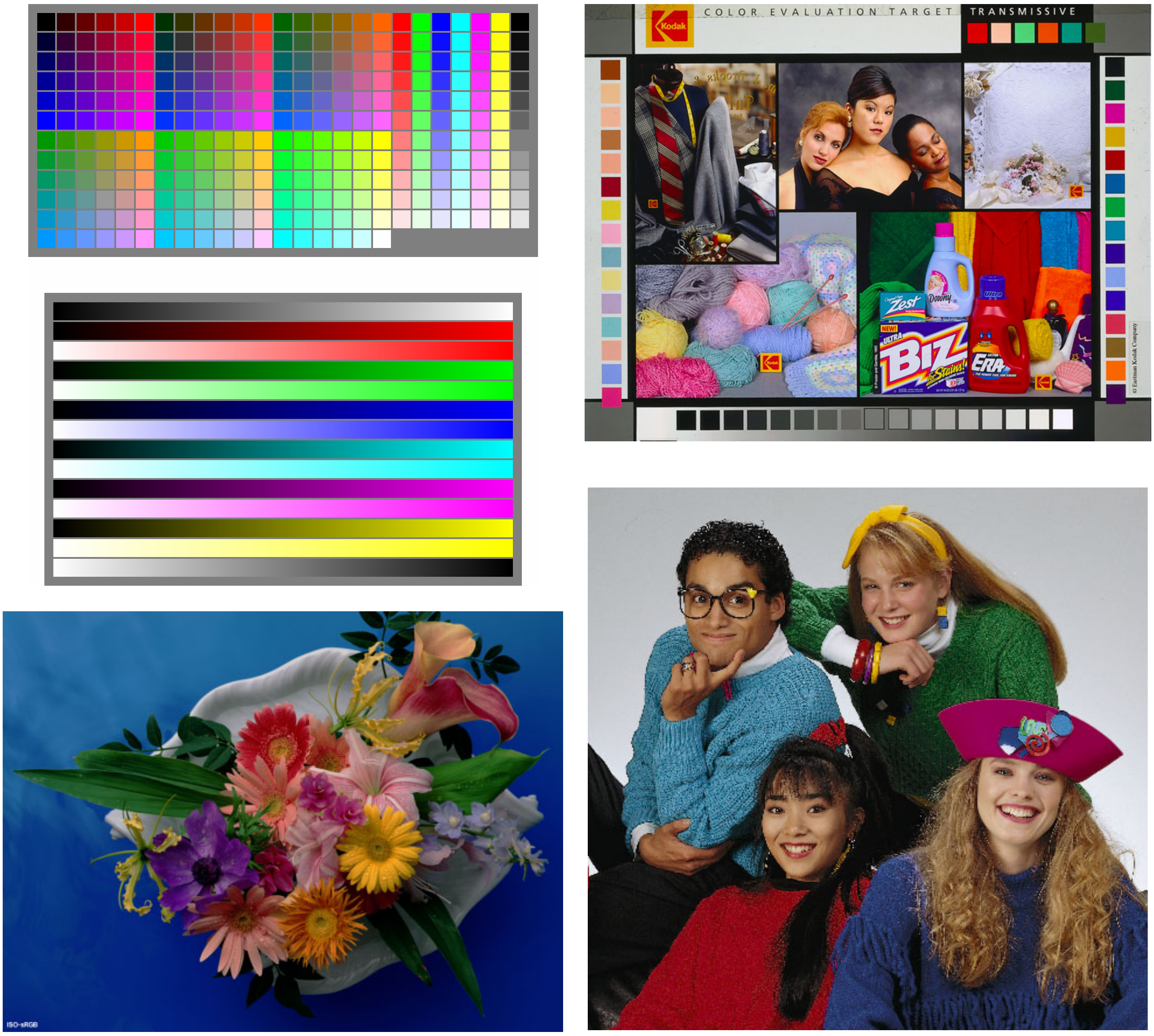

Figure 1. Images used for testing profiles. Left: sRGB SCID images S6 (top), S7, N2. Right: Lab images CET (top) and Group

Using ISO sRGB-SCID Images

Regarding the ISO sRGB-SCID-images,[4] there are two topics to be dealt with:

Firstly, the color patches and vignettes are not evenly spaced according to the definition of sRGB in the ISO/IEC standard,[5] a consequence of inconsistent definitions of various “s”RGB spaces.

Secondly, the corresponding Lab values differ from the ones calculated using the plain CIELAB formula, since a chromatic adaptation transform (CAT) is built into the 'sRGB color space' ICC profile.[6] This must be kept in mind if the image gamut is compared with the source gamut of a CRT calibrated to D65, since the shape of the gamut is rotated by the transform. Due to that, the image gamut extends partly beyond the border of the CRT gamut and gets closer to a printer output gamut, which facilitates the task of gamut mapping.

Definitions of Various “s”RGB Spaces

According to ISO/IEC 61966-2-1,[5] sRGB is defined as follows:

- The chromaticity coordinates (x,y) of the primary colors are R(0.64, 0.33), G(0.3, 0.6), B(0.15, 0.06).
- Different whitepoints for monitor (D65) and surround (D50) are given.
- Linear RGB values r,g,b are converted to normalized device control values R,G,B using the piecewise defined function R(r) = G(g) = B(b), see upper part of tab. 1.
- Use of the term "gamma" for the exponent of the power function is explicitly discouraged.

Table 1. Different nonlinear functions for converting linear values r,g,b to device control values R,G,B

ISO/IEC 61966-2-1	
R(r) =	$\begin{cases} 12.92 \cdot r & , \ r \le 0.00313 \\ 1.055 \cdot r^{1/2.4} - 0.055 & , \ r > 0.00313 \end{cases}$
ITU-R BT.709-3	
R(r) =	$\begin{cases} 4.5 \cdot r & , \ r \le 0.018 \\ 1.099 \cdot r^{0.45} - 0.099 & , \ r > 0.018 \end{cases}$

The sRGB-SCID images[4] will be provided as XYZ images as well, however the final version of ISO 12640-2 is not finished up to now. In contrast to the definition of sRGB, the XYZ images are calculated according to ITU-R BT.709-3, see ISO 12640-2, section 5.2.2. In this standard the same primaries as for sRGB are used, however the so-called "Opto-Electronic Characteristic Function (OECF)" nonlinear function differs from sRGB, see lower part of table 1.

As a result, the color differences designated in ISO 12640-2 are not the ones given in the synthetic images: The gray fields in the patch image S6 and the gray wedge in the vignettes image S7 should be equidistant in Lab according to section 5.4.3 and 5.4.4.

In consequence, both images S6 and S7 were reconstructed according to the definition of the sRGB piecewise function in ISO/IEC 61966-2-1.

Converting the Images from sRGB to Lab

The standard sRGB ICC-profile ("sRGB IEC61966-2.1", HP, 1998)[6] implements a chromatic adaptation using the Bradford matrix from D65 to D50, with the latter being the whitepoint of the ICC-PCS. This chromatic adaptation produces significantly different results compared with the one given in ISO 12640-2 / Annex D, which specifies a von Kries type matrix as is used in the plain CIELAB formula. The Bradford transform is only mentioned as an option: "CIECAM 97s may be applied". Anyway the sRGB profile was used due to its ubiquity.

Results

Some differences can be seen at first sight, whereas others need statistical evaluation and are best expressed as z-Scores.

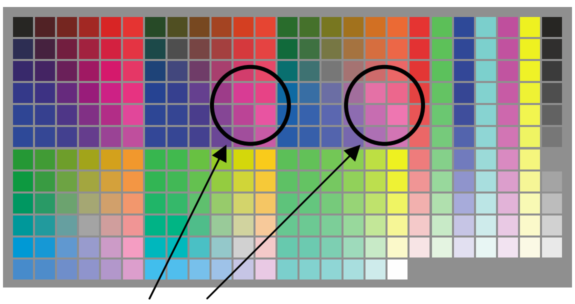

Figure 2. Side effects caused by perceptual transform of tool B

Color Patches

With the perceptual rendering of tool B, the sRGB patch image S6 showed some irregularities in the magenta region. Obviously the perceptual extension of the color space causes some strange side effects, see fig. 2.

Color Wedges

With the option "Preserve Gray Plus" of tool D both the gray and the blue wedge of the sRGB vignette image S7 get brighter, causing a perception of less saturation in the blue region. This effect can be avoided using the option "Chroma Plus" additionally.

Tool B uses small hue changes to enhance the images in a vivid and colorful manner, which is perceptible especially in the blue and magenta wedges.

z-Scores

The interval scale values and the confidence interval were graphed over the corresponding algorithm, the scales for the images investigated are shown in figure 3. An algorithm is significantly better than a second one, if its scale value on the interval scale of preference is greater than the upper bound of the confidence interval of the latter algorithm.

Obviously tool B was rated best in all cases, however in the case of the sRGB flower image this is not statistically significant. For the Group image, the level of significance could be reached by using a few more observers, since the confidence interval would shrink.

In every case, the preserve gray option of tool D produced the worst results. To calculate an image independent interval scale, the ratings of the first four images were averaged, see fig. 3 (bottom).

Discussion

The sRGB color wedges were expected to show some artifacts due to previous simulations and 3D plots of the resulting CIELab values, however this was not the case. Obviously, these wrong predictions are caused by the nonlinear function used for the XYZ to Lab transform,[6] which consists of both a linear and a nonlinear part. This does not matter only for a gray color wedge, in every other case there are at least two points where the 3D plot shows irregularities, since the three channels change from the linear to the nonlinear part at different points on the wedge. Again, this is not visible in the color image, maybe it indicates only another irregularity of CIELab color space.

One of the CM tools significantly enlarges the perceptual device space, mainly into the blue region. Thus nearly the whole sRGB space is mapped onto the gamut of the output device using this particular perceptual rendering intent. This results in higher chromas with minor hue changes, noticeable when reproducing sRGB images on a printing device with a gamut of medium size. A side effect is a slight nonuniformity, being only visible in an artificial sRGB image (sRGB SCID S6). Nevertheless, the images being reproduced using this special perceptual enlargement were judged to be the best choice in most cases.

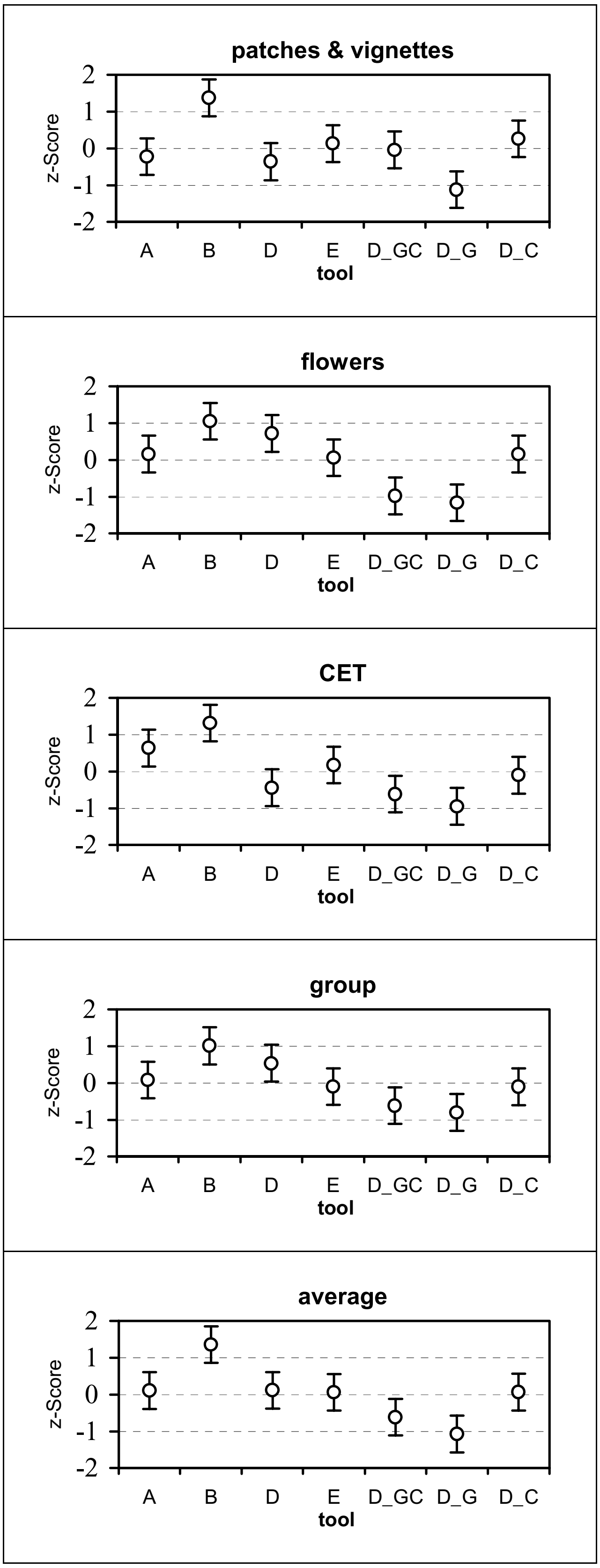

*Figure 3. z-Score values for different images. D_GC designates tool D with 'preserve **g**ray' and '**c**hroma plus' options.*

Conclusion

In this study, the performance of existing CM tools which are used to generate ICC profiles, is evaluated. For this, human observers are asked to compare printed reproductions with an original and an interval scale is constructed based on Thurstone's Law of Comparative Judgement.[8] These results are compared with former studies, where profile quality aspects were investigated e.g. by testing the integrity of the tables within the profiles.

Emerging from the judgements on the IQ of the profiles, consequences for further development of the ICC standard and influence on the evolution of Gamut Mapping is expected, which will help to make high image quality broadly available.

References

1. ICC, File Format for Color Profiles, International Color Consortium Specification ICC.1:1998-09 (1998).
2. Büring H., Herzog P.G. and Jung E., Investigating Quality Aspects of Current Color Management Tools, Proc. SPIE Electronic Imaging Conference, 4300, 300-308 (2001)
3. Büring H., Herzog P.G. and Jung E., Investigating Inverse Gamut Mapping of Current Color Management Tools, Proc. CIC9: The Ninth Color Imaging Conference, Scottsdale, AZ, 263-266 (2001).
4. ISO, Graphic technology – Prepress digital data exchange – Standard colour image data – Part 2: XYZ/sRGB encoded image data (XYZ/SCID), ISO/CD 12640-2
5. ISO/IEC 61966-2-1 Multimedia systems and equipment – Colour measurement and management – Part 2-1: Colour management – Default RGB colour space – sRGB
6. Nielsen M. and Stokes M., The Creation of the sRGB ICC Profile, Proc. CIC6: The Sixth Color Imaging Conference, Scottsdale, AZ, 253-257 (1998).
7. CIE, Colorimetry, CIE Publication No. 15.2, Vienna (1986).
8. Thurstone L. L., A law of comparative judgment. Psychological Review 34, 273 (1927).
9. BARTLESON C. J., Optical Radiation Measurement. Vol. 5 - Visual Measurements, Bartleson C. J. and Grum F. (eds.), Academic Press Inc. (1984)

Biography

Hendrik Büring studied electrical engineering at the Aachen University of Technology, Aachen, Germany. He achieved his diploma degree in 1998. Since then he has been working at the Technical Electronics Institute of the Aachen University of Technology where his interests include color management and gamut mapping. Currently, he is pursuing his PhD degree.