

Automatic Selection of Scanner Color Profiles

Miguel A. López-Álvarez, Johan M. Lammens, and Michel Encrenaz. Hewlett-Packard, Large Format Printing Division; Sant Cugat; Spain.

Abstract

We propose an automatic procedure for color scanner profile selection, which greatly improves the color quality of the final scanned images, and also the user experience. By just analyzing a small number of initial rows of the scanned image, our algorithm is capable of distinguishing among the available color profiles, and then it uses the one that better fits with the substrate of the original. The selection can be done “on the fly” with no time penalty for the user, and the color results obtained are improved when compared to a scanner workflow where no color profiles are used.

Introduction

Large format scanners have to deal with several kinds of media, printing technologies, and contents of the original plots to be scanned. Traditionally, the user was asked about these three parameters before starting the scan or copy (in the case of MFPs [multi-function printers]) process. By doing this, the imaging pipeline in the MFP can be adjusted to provide the best results in terms of color and image quality (IQ). However, for those users whose color and IQ requirements are not exceptional –or for those who are not experts in scanners- this step of settings selection is sometimes frustrating.

Some modern scanners can now automatically detect the content and the printing technology of the original, and then apply the correct image processing settings to the final job. By doing this, the scanned image results in better quality in terms of line sharpness, objects contrast, or areas grain, for example. Nevertheless, no method for the automatic selection of media type to be scanned is used today, and perhaps this is the most important factor influencing the color quality of the final job.

Since these scanning devices are based on RGB channels color capture, and since they are frequently calibrated using one single target (usually an IT8.7/2 color chart printed on traditional silver-halide photographic media), some well-known issues appear (like capture metamerism, non von Kries scanner sensors, different geometries regarding the color measuring device and the scanner, etc) resulting in inaccurate color capture for some types of originals, particularly those that are made using different materials than the calibration chart previously used.

In order to obtain a reasonably good color quality in the copies and scanned files, some scanner manufacturers provide the customers with a way to recalculate the color calibration any time in the future. The problem here is that, again, a single color target is provided for the calibration. Hence, the color quality would be accurate for those originals of the type of the color calibration chart, but very poor for any other kind of plot. Some scanner manufacturers prefer to use IT8.7/2 charts, while others prefer to use inkjet charts printed on matte paper. A trade-off between usability and color accuracy has to be made when selecting the

color target for scanner calibration, if a single color correction is to be done.

Designjet T2300 eMFP is the first large format MFP with an integrated scanner, which provides the customer with 4 different input ICC (International Color Consortium[1-3]) profiles covering a wide range of use cases (matte, glossy, recycled and translucent media), so that optimized color management is available for various types of media and the most frequently used printing technologies for each paper. An additional advantage of the ICC approach is flexibility, which is very useful mainly for two reasons. First, additional ICC profiles can be added at any time, if the manufacturer wishes to improve the color accuracy of new media or printing technologies. Second, ICC profile workflow permits to isolate the printer and the scanner color management and we can easily use the same scanner with different printers, or vice versa, with very good color accuracy.

As an example of the better color accuracy achieved by ICC profiles at the scanner, in Table 1 we show some data from a set of experiments developed with the Designjet T2300 eMFP, where color accuracy clearly improves when we use a specific ICC profile for a given kind of media and printing technology. We show a column at the right with data from the same large format MFP hardware (based on a contact image sensor scanner [4]), but using 3x3 color correction matrix to build a transformation from device RGB into sRGB color space [5,6], instead of ICC profiles at the scanner.

Table 1. Mean (and standard deviation) values showing color accuracy (expressed as dE2000 color differences [7,8]) of some color test targets (in rows) copied on HP Coated paper by using two different color workflows.

Target	ICC profile workflow	3x3 matrix to sRGB
IT8.7/2	4.5 (11.7)	5.7 (11.6)
6 ³ pigment coated	3.4 (6.4)	4.2 (8.1)
6 ³ pigment glossy	6.2 (13.3)	7.0 (14.2)
6 ³ laser plain	3.1 (6.0)	3.5 (6.7)
6 ³ dye glossy	3.0 (9.4)	4.8 (10.6)
6 ³ dye coated	2.1 (4.5)	3.1 (7.0)

In Table 2 we show how the influence of the media type on the final color accuracy is more important than the printing technology, at least in our scanner. For example, the “6³ pigment glossy” target in the third row of Table 2, is copied with good

accuracy when using the ICC profile built with the glossy IT8.7/2 target (which is printed using silver halide technology), and not so good when using the ICC profile built for matte paper even if it is built with a color target printed with the same pigment inkjet technology. In fourth row of Table 2 we can also see how the color target printed on plain matte paper with laser technology is also well copied when using the ICC profiles built with matte paper and inkjet technology (both dye-based and pigmented inks), hence demonstrating the higher importance of the substrate over the printing technology when dealing with scanners ICC color profiles.

Table 2. Mean (and standard deviation) values showing color accuracy (expressed as dE2000 color differences) of some color test targets (in rows) copied on HP Coated paper by using different ICC profiles for each target.

Profiles right / Targets down	ICC profile pigment matte	ICC profile dye-based matte	ICC profile silver halide glossy	ICC profile pigment glossy
IT8.7/2	6.8 (13.1)	6.1 (10.6)	4.5 (11.7)	5.6 (11.5)
6 ³ pigment coated	3.3 (6.4)	3.4 (6.5)	4.5 (7.4)	4.4 (9.3)
6 ³ pigment glossy	7.4 (15.4)	6.7 (13.2)	6.5 (12.9)	6.2 (13.3)
6 ³ laser matte	3.6 (6.8)	3.8 (7.3)	4.7 (9.6)	6.5 (9.4)
6 ³ dye coated	3.7 (8.4)	2.1 (4.5)	4.6 (7.3)	3.4 (6.2)

The drawback of this ICC approach is that customers have to select one of the available scanner profiles every time they start a new scan. If the wrong profile is selected, then the scanner's color accuracy will be as poor as in those scanners with single color chart calibration. If the right profile is selected, then the color quality of the scanned document is better than the approach not using ICC profiles, as illustrated in Table 1.

In the case of our scanner, there are four ICC profiles available for the user to select, corresponding to matte, glossy, recycled, and translucent media. These profiles have been chosen so that they cover the widest possible range of media types for the type of users this scanner is targeted. The profile for matte paper is selected as "default" since our studies show that our customers mostly scan this type of paper. Since the spectral and colorimetric characteristics of these four media types are very different, we believe that an automatic method for distinguishing them can be easily implemented.

Our method provides an automatic approach for selecting the best ICC profile for the scanner when dealing with a particular original, by just analyzing a small number of initial rows of the scanned image. Hence, users do not have to select the color profile for the scanner every time they start a new scan. Moreover, the automatic selection will be the optimum one in most of the cases, and -if not- the final color accuracy achieved will not be worse

than in the case of a scanner not using ICC profiles workflow (as shown in Table 1).

We have tested this algorithm with a set of over 100 samples of different kinds of representative media and content, as explained in results section below, in a Designjet T2300eMFP large format scanner supporting ICC profiles, and we obtained the optimum, hence noticeable better, color results than using a 3x3 matrix for color correction under the same hardware in 74% of the cases (i.e., the algorithm selected the most appropriate profile for the scanner in these cases). In another 13% of the cases, the selected profile was not the optimum, but its white point was sufficiently similar to the optimum, so that similar color results were obtained, hence still noticeable better than using a 3x3 matrix for color correction. Only in another 13% of the cases, the default profile was finally selected since our algorithm was not capable of making an accurate decision. However, similar color accuracy (never worse) to that obtained with a 3x3 matrix was finally achieved when the default profile was finally used.

Method

Our algorithm uses the first rows of the scanned image, which very often contain some white band of unprinted paper, to obtain a colorimetric estimation of the white point of the paper. This first step can be done with no time penalty, since in most large format scanners today a previous scan of about 10cm length is already done in order to automatically detect the width of the original, and the data from this previous step can be used for our algorithm of automatic scanner color profile selection.

In case the scanner has to deal with originals that show no white information in the first rows, our algorithm is capable of detecting this situation and then use a default profile (the color accuracy obtained with this default profile is still as good as in the case of color pipelines not using ICC profiles) or provide an alert to the users so that they can manually select the profile for the scanner in case they know.

Once the first rows of the sample are scanned, our algorithm follows these steps:

- 1) Obtain an averaged white sample in the scanner's device RGB (dRGB from now on; also known as native RGB) color space. This is achieved by finding the average of a small region of pixels with maximum RGB values within the scanned first rows of the original.
- 2) Convert the dRGB value of the white sample to CIELAB using the absolute colorimetric intent with each of the available scanner profiles (previously calculated from calibration charts on different, representative media and colorant combinations).
- 3) Calculate the color differences between the CIELAB value of the white sample obtained in 2) and the CIELAB value of the "white point tag" of each of the available scanner profiles. This tag is usually given in CIEXYZ coordinates, but the profile itself can be used to calculate the CIELAB values we need here (just once).
- 4) Choose that profile which results in the smallest color difference as calculated in step 3). Even if the selection is not the 'correct' one according to media type, the result will still be good enough in practice since white point differences have been minimized (we have seen some cases of matte and semi-glossy media which give color differences close to the glossy and matte profiles respectively, hence one would think that the automatic

selection could be mistaken; nevertheless, the color accuracy achieved even if the wrong profile is selected in this particular example only worsens in the range 5-10%.

5) If the smallest color difference obtained in step 4) of the algorithm is bigger than the largest color difference among the white points of all available profiles (see Table 3), then consider the white sample invalid and switch to step 6). If not, then the algorithm is finished in step 4) and the automatically chosen ICC profile is used.

6) Select the default profile as the automatically chosen (in the case of our scanner, this default profile is the “Matte Inkjet” profile, since the color accuracy obtained with it in all the cases is good enough, and also it is the most used profile by the kind of users of this product), but always give the user the option to override the automatic choice.

Basically, our algorithm benefits from the fact that each scanner profile is built for media with different white points. If we use this method to estimate the white point of the scanned plot, then we can automatically select the scanner profile that is intended for the closest white point, hence providing good color results.

Results

In Table 3, we show the color differences values (in dE76 or CIELAB units) between the white points of the four scanner ICC profiles used in our MFP. The distance in Lab* color space between the white points of the profiles is big enough to allow our algorithm to distinguish among the different media types. According to our procedure, if the distance between a sample’s estimated white point and the nearest profile’s white point is bigger than the biggest distance shown in Table 3, we consider that the estimation is not correct (as explained in step 5 above).

Table 3. dE76 color differences between the scanner’s ICC profiles white points.

	Matte paper	Recycled	Photo	Tracing
Matte paper	0	7.4	4.3	17.3
Recycled	7.4	0	5.9	13.3
Photo	4.3	5.9	0	18.4
Tracing	17.3	13.3	18.4	0

In order to illustrate the workflow of our algorithm, we show in Table 4 an example developed with four different media as input. We scanned a white area in each of the samples (which correspond to typical samples where each of the scanner profiles was built), converted the RGB values of the white area into Lab* values using each of the four available ICC profiles, and the measured the color differences between these Lab* values and the Lab* data corresponding to each of the profile’s white point tag. We obtained the color differences shown in Table 4, where we show how the optimum ICC profile for the scanner is always selected by following our method. Hence, the customer always gets the optimum color results without the need of selecting the scanner profile manually.

Table 4. dE76 color differences obtained after applying our proposed algorithm with 4 media samples corresponding to the four scanner ICC profiles available.

Profiles right / Targets down	Matte paper	Recycled	Photo	Tracing
Matte paper	1.1	5.8	3.3	17.7
Recycled	2.9	1.5	2.0	13.9
Photo	5.1	7.1	1.5	19.3
Tracing	4.7	4.2	7.0	1.5

The experiment described in Table 4 was repeated with various sets of different media samples and printed content, which were considered to be fairly representative of the originals used in this kind of large format scanner. We obtained reasonably accurate results, concluding that in 74% of the cases the optimum profile was automatically selected by following our method, while in 13% of the cases the automatic selection –although was not the optimum- led to an ICC profile with similar white point to the optimum one, and hence the color accuracy of the final copy was almost as good as the optimum one.

In another 13% of the cases, our method was not capable of finding a white sample in the image from which obtaining an estimation of the paper type was possible. Hence, in these cases the default profile was finally used. Coincidentally, the default profile was the optimum ICC profile in 33% of these failing cases (4% of the total). In the rest of cases where the default profile was finally selected, the color accuracy achieved was comparable to a copy of the same original in a workflow which does not use ICC profiles.

References

- [1] Specification *ICC. 1:2004-10 (Profile version 4.2.0.0) Image technology colour management: Architecture, profile format, and data structure*, International Color Consortium, 2004.
- [2] Specification *ICC.1:2001-04 File Format for Color Profiles*. International Color Consortium, <http://www.color.org/>
- [3] P. Green. *Color management: understanding and using ICC profiles* (Wiley-IS&T, 2010).
- [4] R. G. Gann. *Desktop scanners: Image quality evaluation* (Hewlett-Packard professional books, Prentice Hall, 1998).
- [5] <http://en.wikipedia.org/wiki/SRGB>.
- [6] H. R. Kang. *Color technology for electronic imaging devices* (SPIE Optical Engineering Press, 1997).
- [7] R. W. G. Hunt. *Measuring Colour* (Fountain Press, 3rd edition, 1998).
- [8] M. R. Luo, G. Cui, and B. Rigg. “The development of the CIE 2000 colour-difference formula: CIEDE2000”, *Color Res. Appl.* **26** 340-350 (2001).

Authors Biography

Miguel A. López-Álvarez obtained a B.Sc. in Physics in 2003, a B.Sc. in Electrical Engineering, and a Ph.D. in multispectral color science in 2007 by the University of Granada (Spain). Since 2008 he works in Hewlett-Packard as a color and digital imaging engineer, focusing on color management and image quality of large format printers and scanners.

Johan M. Lammens studied linguistics (BA) and computational linguistics (MA) at the University of Antwerp, and worked on text-to-speech

conversion at the Technical University of Eindhoven (NL) and the University of Utrecht (NL). He obtained his MS and PhD degrees in computer science at SUNY/Buffalo, and held a postdoctoral position at the University of Pisa (IT). Since 1995 he has been with HP R&D in Barcelona (ES), currently in the position of Program Manager.

Michel Encrenaz has an Engineering degree from ENIT (France) and a Ph.D. from Université Bordeaux I. He has been working in HP (IPG – Large Format Printer division) for the last 12 years in writing system integration, and as Chief Engineer.