



Assessing color reproduction tolerances in commercial print workflow

Giordano B. Beretta, Eric Hoarau, Sunil Kothari, I-Jong Lin, Jun Zeng

HP Laboratories
HPL-2011-231

Keyword(s):

printing, digital printing, commercial printing, commercial print automation, workflow

Abstract:

Except for linear devices like CRTs, color transformations from colorimetric specifications to device coordinates are mostly obtained by measuring a set of samples, inverting the table, and looking up values in the table (including interpolation), and mapping the gamut from input to output device.

The accuracy of a transformation is determined by reproducing a second set of samples and measuring the reproduction errors. Accuracy as the average predicted perceptual error is then used as a metric for quality. Accuracy and precision are important metrics in commercial print because a print service provider can charge a higher price for more accurate color, or can widen his tolerances when customers prefer cheap prints.

The disadvantage of determining tolerances through averaging perceptual errors is that the colors in the sample sets are independent and this is not necessarily a good correlate of print quality as determined through psychophysics studies. Indeed, images consist of color palettes and the main quality factor is not color fidelity but color integrity. For example, if the divergence of the field of error vectors is zero, color constancy is likely to take over and humans will perceive the color reproduction as being of good quality, even if the average error is relatively large. However, if the errors are small but in random directions, the perceived image quality is poor because the relation among colors is altered.

We propose a standard practice to determine tolerance based on the Farnsworth-Munsell 100-hue test (FM-100) for the second set and to evaluate the color transpositions—a metric for color integrity—instead of the color differences. The quality metric is then the FM-100 score. There are industry standards for the tolerances of color judges, and the same tolerances and classification can be used for print workflows or its components (e.g., presses, proofers, displays). We generalize this practice to arbitrary perceptually uniform scales tailored to specific applications and present an implementation.

In essence, we propose to extend the color discrimination test procedures used to evaluate human observers, to mechanical and electronic color reproduction devices.

Assessing color reproduction tolerances in commercial print workflow

Giordano B. Beretta, Eric Hoarau, Sunil Kothari, I-Jong Lin, Jun Zeng

Hewlett-Packard Laboratories, Printing and Content Delivery Laboratory
1501 Page Mill Road, Palo Alto, USA

ABSTRACT

Except for linear devices like CRTs, color transformations from colorimetric specifications to device coordinates are mostly obtained by measuring a set of samples, inverting the table, and looking up values in the table (including interpolation), and mapping the gamut from input to output device.

The accuracy of a transformation is determined by reproducing a second set of samples and measuring the reproduction errors. Accuracy as the average predicted perceptual error is then used as a metric for quality. Accuracy and precision are important metrics in commercial print because a print service provider can charge a higher price for more accurate color, or can widen his tolerances when customers prefer cheap prints.

The disadvantage of determining tolerances through averaging perceptual errors is that the colors in the sample sets are independent and this is not necessarily a good correlate of print quality as determined through psychophysics studies. Indeed, images consist of color palettes and the main quality factor is not color fidelity but color integrity. For example, if the divergence of the field of error vectors is zero, color constancy is likely to take over and humans will perceive the color reproduction as being of good quality, even if the average error is relatively large. However, if the errors are small but in random directions, the perceived image quality is poor because the relation among colors is altered.

We propose a standard practice to determine tolerance based on the Farnsworth-Munsell 100-hue test (FM-100) for the second set and to evaluate the color transpositions—a metric for color integrity—instead of the color differences. The quality metric is then the FM-100 score. There are industry standards for the tolerances of color judges, and the same tolerances and classification can be used for print workflows or its components (e.g., presses, proofers, displays). We generalize this practice to arbitrary perceptually uniform scales tailored to specific applications and present an implementation.

In essence, we propose to extend the color discrimination test procedures used to evaluate human observers, to mechanical and electronic color reproduction devices.

Keywords: printing, digital printing, commercial printing, commercial print automation, workflow

1. INTRODUCTION

1.1 Problem statement and industry needs

Generally, state-of-the-art ICC-based commercial color print workflows aim to a perfect color match between contract proofs and the final output of a digital press. In the real world, perfection does not exist: processes are based on physics and feedback loops, which entail measurements and their errors, in addition to the tolerance of the physical process itself.

In our simulation-based approach to commercial color print workflows,¹ we assume knowledge of the tolerances of all stages in the print and finishing process, because modeling and simulation allows scheduling and resource allocation according to quality requirements. The predictability of print manufacturing quality allows to charge as much as possible for print jobs, or conversely to reduce quality to what the market can afford. In either case, the print service provider's (PSP) profits can be maximized.

Giordano B. Beretta, Eric Hoarau, Sunil Kothari, I-Jong Lin, Jun Zeng, "Assessing color reproduction tolerances in commercial print workflow," Color Imaging XVII: Displaying, Hardcopy, Processing, and Applications; Reiner Eschbach, Gabriel G. Marcu, Alessandro Rizzi, Editors, Proc. SPIE 8292, Paper 8292-32 (2012)

© Copyright 2012 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic electronic or print reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

DOI not yet available (look for proceedings 8292 on spiedl.org).

The ICC standard covers the encoding of the data specifying the characterization of a device, i.e., the mapping of intrinsic device coordinates to device-independent colorimetric values, as shown in Fig. 1. This data can be a function (when there is a model), or it can be a table of various sizes (when there is no model), and either one can be colorimetric or spectral. Depending on a device's deviation from linearity, the choice of the method and the precision of the data, quality can vary widely. Last but not least, the accuracy depends on the skills and tools of the profile's creator, which also vary widely.² Skilled printers will look at the ICC profiles provided with a job and, in case of quality issues, will ask the client to provide better profiles. For the output profiles they will most likely use their own carefully prepared ICC profiles.

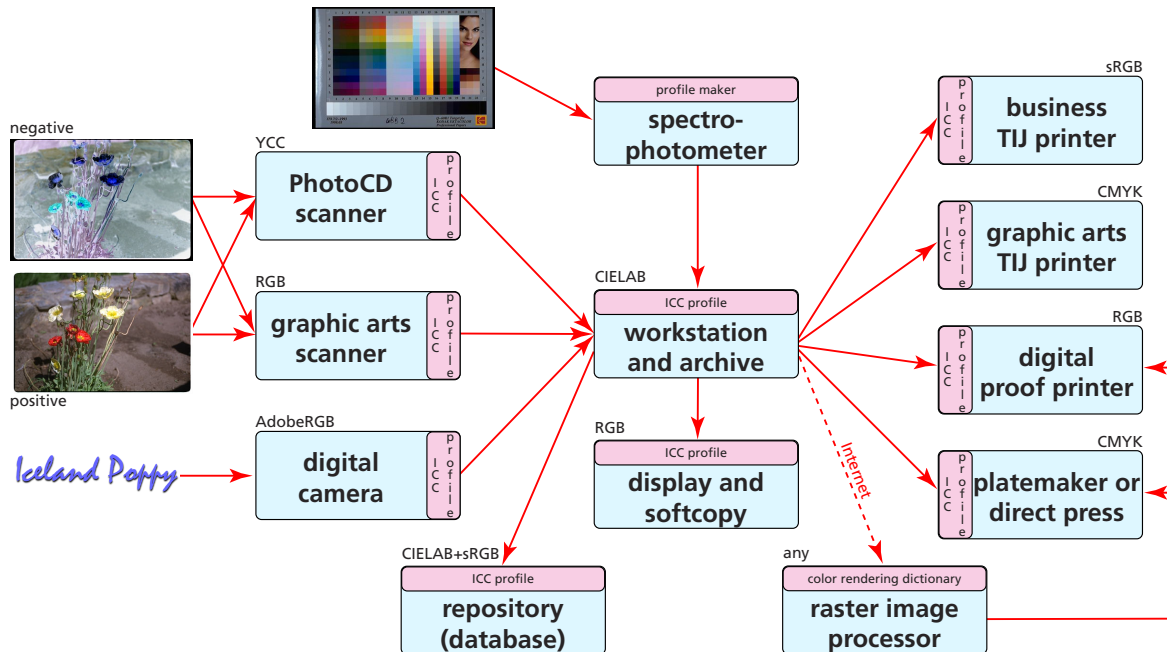


Figure 1: With ICC profiles, a workflow comprising n color devices only requires data for $2n$ color transformations instead of n^2 . (Photograph by William R. Hewlett used with permission.)

A printed piece is not a perfect replica of an original, and CIE 1976 ($L^*a^*b^*$) errors in the range $3 < \Delta E < 7$ are routine in commercial color print processes that are run by skillful operators. In the real PSP world, larger errors can be experienced, especially when market pressures force down production costs. When a service license agreement (SLA) is negotiated and during the constraint resolution phase between requirement job ticket and capabilities job ticket, precise knowledge of color tolerances is essential for profit maximization (for example, a job can be run on a less accurate press or by a less skillful operator) in order to meet a profitable deadline.

Typically the tolerances are calculated as mean CIE 1976 ($L^*a^*b^*$) errors. The disadvantage of this method is that the colors in the sample sets are independent and this is not a good estimation of print quality as determined through psychophysics studies.³ Indeed, images consist of color palettes and the main quality factor is not color fidelity but *color integrity*, i.e., the ability to preserve the relations among the colors in a palette. For example, if the divergence of the field of error vectors is zero, color constancy is likely to kick in and humans will perceive the color reproduction as being of good quality, even if the average error is relative large. However, if the errors are small but in random directions and clustered, the perceived image quality is poor.

EXAMPLE 1. Consider a catalog selling denim garments. One page might have jeans in original, washed, and stone-washed denim, while a second page might have corresponding jackets. The specific hue of each garment is less important than being able to match jackets to jeans while preserving the rank of the processes. As used herein, color integrity is preserved in a reproduction when a color scale is ordered, and this order is preserved in the reproduction.

1.2 Proposed solution

In a paper we wrote a few years ago [4, §2.2], we postulated that the old goal of color fidelity in color reproduction had become an oxymoron in the age of mobile computing and opportunistic printing. We argued for color integrity in the sense of Evans’ consistency principle [5, p. 600], [6, §2.2]. In short, we proposed that color integrity could be preserved when the vector field of error vectors is divergence-free and no color name boundaries are crossed.

That same year in November, at the Color Imaging Conference, Lindsay McDonald organized a panel discussion to debate our proposal. Half of the delegates, including Dr. Robert W.G. Hunt, saw merit in our idea. By the way of record, Tony Johnson debunked as an urban legend the still widely circulating story of Land’s End receiving a large number of returns due to color matching problems. In reality returns are mostly due to design and style, and rarely to color.

In the meantime, the color name boundaries* are still unknown and, as far as we know, nobody has ever implemented a system based on color integrity. In this paper we propose a solution that is half-way between color fidelity (CIELAB $\Delta E \leq 1$) and color integrity. To begin, we consider a uniform color scale that we reproduce on the system to characterize. We then count the color transpositions and score them. If a color reproduction system has a score of zero (no transpositions), it is considered to have the smallest tolerance, independent of ΔE .

1.3 Current practices

The standard practice for establishing color tolerances is described in ASTM standard D 3134.⁸ This standard practice explains that for product specification, usually tolerances between specimens and a standard are agreed upon between the purchaser and the seller. However, at times the purchaser is either unknown or unequipped to determine tolerances. In such cases, D 3134 recommends the seller shall set tolerances unilaterally, based upon such considerations as the economics of the production process, known buyer requirements, and similar products [8, §7.4.1].

D 3134 recommends to express color differences through the CIE 1976 ($L^*a^*b^*$) color difference equations [8, §7.2.1.1]. More specifically, D 3134 [8, §7.4.5.3 and §7.4.5.4] states that tolerances are preferably set experimentally by classifying specimens visually against a standard,[†] then measuring the specimens and standard and plotting the values on a CIELAB Δa^* , Δb^* constant lightness ($L^* \equiv k, k \in [0, 100]$) plane. For the L^* tolerance, D 3134 states that a rough approximation is the lightness difference ΔL^* for colors that match the standard in hue and saturation.

In general, the tolerance is not a sphere but an ellipsoid or an ovoid, with axes aligned to the correlates for lightness, chroma, respectively hue. For simplicity, in this work we just assume a spherical tolerance; it is straightforward how this is generalized.

It is clear from the wording of D 3134, that the main application are large solid colors, such as for example automobiles, appliances, colored masonry, etc. In the case of complex colors, such as printed images, one would measure a target image consisting of a number of colors, like for example a Munsell color chart, and average the differences. Sometimes, the RMS is calculated instead of the average.

Similarly to the above, the CIE defines indices to be used for determining tolerances of such quantities as the color rendering of light sources or the metamerism of a process. They are defined as average differences calculated using the CIE 1976 ($L^*a^*b^*$) color difference equations or their more modern derivatives.

The 1989 Xerox Color Encoding Standard—written largely by Robert R. Buckley—had an attribute for each color to specify the tolerance of that color [9, §2.4, page 2–5]. The tolerances were specified in terms of CIE 1976 ($L^*a^*b^*$) color-difference units ΔE_{ab}^* . In the product implementation, this attribute could be used, for example, to select the default printer based on its capability to achieve the prescribed tolerance. Other later standards, like the ISO/ODA Color Addendum had similar provisions.

In the end, color prints are not bought and read by spectrophotometers but by humans, or “observers” in color science parlance. While the tolerances of spectrophotometers can be determined for example with the BCRA tiles,² those of humans can be determined more appropriately with color discrimination tests like the Farnsworth-Munsell 100-hue test, the HVC Color Vision Skill Test, and the Japanese Color Aptitude Test.¹⁰

* At the time we were considering the ISCC–NBS color name boundaries,⁷ but now we know they are too empirical.

[†] I.e., the starting point is the graphical interpretation of historical data (batches accepted, rejected, or considered borderline) [8, §7.2.5.1]. This graphical interpretation and the rating are accomplished by multiple observers.

2. COLOR TRANSPOSITIONS IN A PALETTE

As noted in the footnote on the previous page, the observer is a critical element. ASTM standard guide E 1499¹⁰ provides criteria for the selection, evaluation, and training of human visual-sensory observers. Visual observation is a two-stage process: stimulus detection by the eye and interpretation in the brain. Current instruments can only perform the first stage, therefore the screening and training of observers is very important.

For example, at in one company all researchers working on color imaging and printing were encouraged to affix their color discrimination score outside their office doors, in another people working on color printing had to periodically pass the HVC Color Vision Skill Test. In our lab, almost all members have been administered the Farnsworth-Munsell 100-hue test. This is in accordance to §7.1 in E 1499.

In §6.2, E 1499 discusses a number of color vision tests commonly used in industry. Of these, it singles out Farnsworth-Munsell 100-hue test as a color discrimination test that measures color discrimination directly and in detail. Our method consists of administering this test to a color reproduction system in order to assess its ability to preserve color discrimination.

2.1 The Farnsworth-Munsell 100-hue test set

The Farnsworth-Munsell 100-hue test is described in Farnsworth’s original paper.¹¹ The basic idea is to generate a one-dimensional scale in a perceptually uniform color space and to rank the scale elements (Munsell Color Standards, or Munsell Sheets of Color), thus introducing an order. The Munsell Color Standards are then randomized and placed on a two-dimensional surface in a controlled station. The subject being tested is asked to sort the Munsell Color Standards in a linear order. This order is then compared to the rank of the Munsell Color Standards.

The scale elements are chosen to have constant value and chroma, changing only in hue, as shown in Fig. 2 reproduced from the Farnsworth paper. An observer with so-called “Golden Eyes” will be able to sort all Munsell Color Standards according to their rank. Any other observer will be unable to discriminate among two or more of the Munsell Color Standards and will place these in random order. The transpositions provide a measurement of the observer’s color discrimination ability.

2.2 Computing transpositions

The computations can be easily determined by printing the rank order on the back sides of the Munsell Color Standards. The test administrator simply flips over the tray with the Munsell Color Standards and records the rank in the order of the Munsell Color Standards.

EXAMPLE 2. *Results of a subject’s test. Top row: Munsell Color Standards rank. Bottom row: order in which the Munsell Color Standards are sorted by the subject*

2	3	4	5	6	7	8	9	10	11	12	13
2	3	5	4	6	7	8	11	9	13	10	12

2.3 Scoring the test

Scoring is straightforward and described in detail in the Farnsworth-Munsell 100-hue test manual.¹² The score for each of the Munsell Color Standards is the sum of the differences between the number (rank) of that element and the numbers of the Munsell Color Standards adjacent to it. Because in the case of no transposition the sum of the differences is 2, we subtract 2 from each individual score. The total score is obtained by summing the individual scores.

EXAMPLE 3. *Scoring Ex. 2. The third row contains the differences (assuming that the next Munsell Color Standards to the right is 15) and the fourth row the scores*

2	3	4	5	6	7	8	9	10	11	12	13
2	3	5	4	6	7	8	11	9	13	10	12
2	3	3	3	3	2	4	5	6	7	5	5
0	1	1	1	1	0	2	3	4	5	3	3

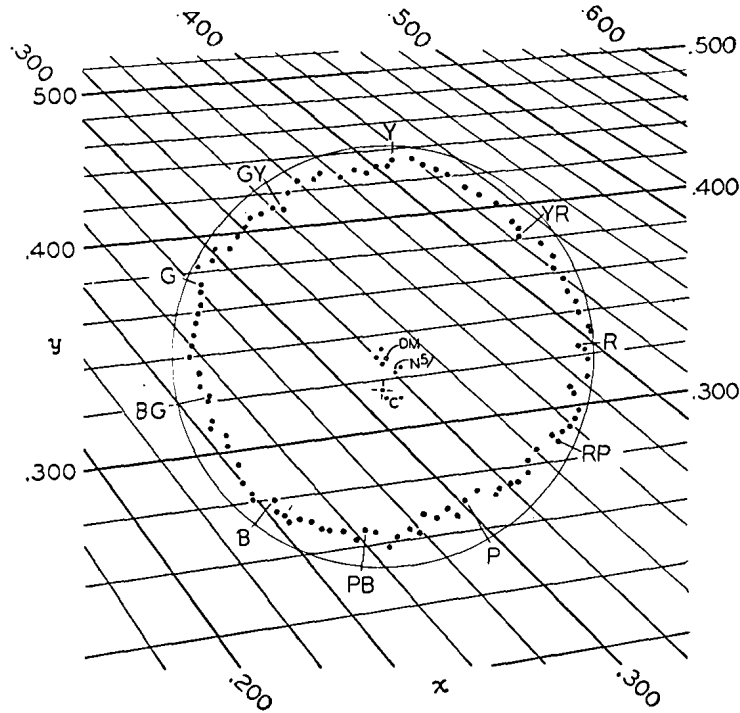


Figure 2: Position of the 100 Munsell Color Standards at 5/5 on a perspective projection of the I.C.I. Standard Observer Diagram .¹¹

In the case of human observers, a subject with a total score of zero is often said to have *Golden Eyes*. About 16% of the population (exclusive of color defectives) has been found to make 0 to 4 transpositions, or total error scores in $[0, 16]$ and is designated as having *Superior Discrimination*. About 68% is designated *Average Discrimination* and scores in $[20, 100]$. Finally, about 16% of the population scores higher than 100 and is designated *Low Discrimination*.¹²

METHOD 1. *In the case of color transformations for digital presses, we would print the Munsell Color Standards and measure them. We would then sort them by hue and score them like in the case of humans. If two Munsell Color Standards have the same measured hue, we would order them at random.*

In a psychophysics experiment to establish the ground truth, a panel of experts would establish a quality scale in function of total score.

The scores can be represented graphically, as shown in Fig. 3. This graphical representation shows that this particular observer of average discrimination performs poorly mostly in the green to blue range. Therefore, for example a jeweler would not employ this person to sort turquoise gems.

In the practice of the Farnsworth-Munsell 100-hue test a different graphical representation is used, where the scores are plotted in a polar diagram as shown in Fig. 4. The symmetry of the scores allows for a quick diagnosis of color vision deficiencies. In the case of Fig. 4 both observers have normal color vision, but the observer at left has average color discrimination, while the observer at right has superior color discrimination.

2.4 Generalization to arbitrary perceptually uniform color scales

The color scale with the Munsell Color Standards does not have to be a circle of constant chroma and value. Any one-dimensional scale with equally and critically spaced colors can be used. This leads to the creation of application-specific tolerance methods. For example, a scale of complexions could be constructed for assessing the quality of color transformations when used to print portraits in advertisements for skin care products. In paint applications, it is often very difficult

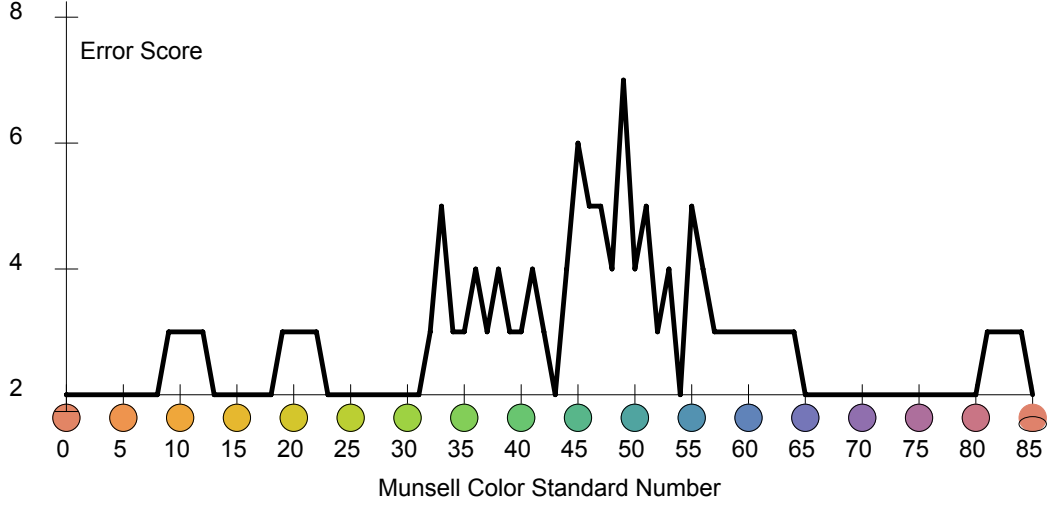


Figure 3: Error scores of a human observer with a total score of 68, average discrimination.

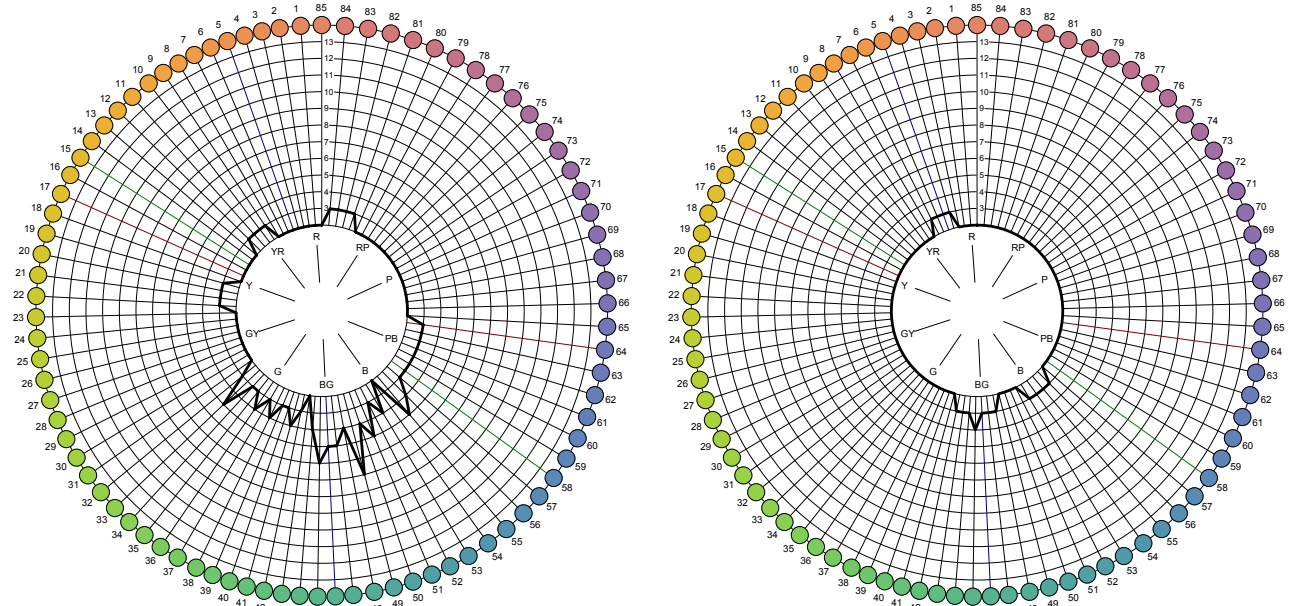


Figure 4: Left: Same error scores as in Fig. 3, but plotted in a polar diagram; there is no diametrical symmetry, suggesting the observer has normal color vision. Right: diagram of an observer with total score 16, superior discrimination.

reproduce correctly the off-whites. An off-white scale could be used to assess the capability of a print workflow to process satisfactorily this kind of print jobs.

3. ADVANTAGE OF INTEGRITY VS. FIDELITY

It has been shown that color accuracy is not a good predictor for the quality of print jobs.³ We propose a different method to characterize the tolerance of a digital commercial press. This method is based on the ability of a color reproduction process to preserve the order of colors in a one-dimensional palette. The method is similar to what is done to test human observers using the Farnsworth-Munsell 100-hue test.

If the color sequence is the same as the Munsell Color Standards in the Farnsworth-Munsell 100-hue test, the tolerance of a print process can be compared directly to the performance of humans. For example, if the commercial print workflow

comprises a step in which embedded profiles and tone reproduction curve manipulations are inspected, in the case of simple images the job can be printed with a press having a tolerance scoring in [20, 100], while jobs with images that contain high quality ICC profiles and image enhancements would be printed on a press with a tolerance scoring in [0, 16] to keep customers happy.

In a similar scenario, if an embedded ICC profile would produce a poor reproduction, the workflow software can swap out that ICC profile with a more accurate profile, as suggested at last year's conference.¹³ Falkenstern et al.¹⁴ recently proposed to select print workflows based on image content. In this case, our method can be used to ascertain that the image integrity is not compromised, aiding in the selection of the best workflow.

By using Munsell Color Standards or other standards specific to an application, such as complexion, off-white paints, or denims, a workflow's tolerances and the SLA negotiation can be optimized for the specific application. Basing a commercial print workflow on tight and reliable tolerances will maximize the press owner's profits.

4. IMPLEMENTATION

A one-dimensional color scale at a fixed step is reproduced in the crop margin of the print sheet. The reproduced scale is then measured and the transpositions of color swatches are scored. When there are no transpositions, the order in the color scale can be maintained, otherwise, the score is used to determine if the job has to be reproduced differently, for example swapping out the ICC profile or changing the algorithms for the black skeleton, gamut mapping, or halftoning. It may also be necessary to change the medium or use a better color management system.

In the current implementation, the color scale is defined in the CIE 1976 ($L^*a^*b^*$) color space, but the underlying implementation can create a scale in any n -dimensional colorimetric color space.

For CIE 1976 ($L^*a^*b^*$) the constructor is as follows.

```
1 public ColorScale (ColorOpLab start, ColorOpLab end, double step) {
2     if (step < 0.1) step = 0.1;
3     double distance = end.deltaE76(start);
4     double t, stretch, distanceFinal;
5     ColorOpLab startFinal, endFinal;
6     int swatches = (int) Math.ceil (distance/step);
7     if (swatches < 3) swatches = 3;
8     distanceFinal = step * swatches;
9     stretch = (distanceFinal - distance) / (2 * distanceFinal);
10    startFinal = new ColorOpLab (-stretch, start, end);
11    endFinal = new ColorOpLab (1.0+stretch, start, end);
12    swatch = new ColorOpLab[swatches+1];
13    for (int i = 0; i <= swatches; i++) {
14        t = (double) i / (double) (swatches);
15        swatch[i] = new ColorOpLab (t, startFinal, endFinal);
16    }
17 }; // ColorScale constructor
```

The parameters are the start and end points of the scale in CIE 1976 ($L^*a^*b^*$) space and the fixed step between swatches in the scale. The step is clamped at the lower end to 0.1 units, as a value below 0.3 is not meaningful anyway, being the typical tolerance of a spectrophotometer across units.

For simplicity, it is not necessary that the length of the scale be specified to be a multiple of the step size. The constructor will extend the scale on both ends to fit a whole number of steps of the specified size. The scale length as specified is calculated on line 3 of the above listing. On line 6 we calculate the whole number of elements in the scale, which must be at least 3 to make sense (line 7).

On line 8 we calculate the extended scale size and on line 9 we calculate by how much we need to extend the scale at each end, and on lines 10 and 11 we calculate the new start respectively end points.

On line 12 we allocate the array containing the CIE 1976 ($L^*a^*b^*$) values and in the loop on lines 13–16 we calculate the CIE 1976 ($L^*a^*b^*$) value for each swatch. The class *ColorScale* then contains a method *scaleElement* to return the color coordinates of the i -th elements of the color scale.

We now discuss the CIE 1976 ($L^*a^*b^*$) color model operator class constructor called in lines 10, 11, and 15. This method is implemented in the super-class *ColorimetricOp*, which is not restricted to three dimensions and which allocates a new color point as follows.

```
1 public ColorimetricOp (double t, ColorimetricOp p0, ColorimetricOp p1) {
2     value = new double[p0.value.length];
3     for (int i = 0; i < value.length; i++)
4         value[i] = (1.0 - t) * p0.value[i] + t * p1.value[i];
5 };    // ColorimetricOp
```

The loop is on the dimensions of the color model operator; in the case of CIE 1976 ($L^*a^*b^*$) these are L^* , a^* , and b^* . As the inner loop in line 4 suggests, we use the parametric representation of line segments. The parametric equation of a line is defined in terms of its end points and a parameter $t \in [0, 1]$ for points in the line segment. This parameter t is a convenient scalar quantity introducing an order. More formally, a line segment from point P_0 to point P_1 is defined by the parametric equation

$$P(t) = (1 - t)P_0 + t(P_1) \quad (1)$$

A point at the fraction of a segment is found by setting t to this fraction. For example, consider the two colors with CIE 1976 ($L^*a^*b^*$) coordinates (80, 15, 2) respectively (85, 20, 45). The above ColorScale constructor will generate a scale of 23 swatches spaced by 2 units and with these CIE 1976 ($L^*a^*b^*$) coordinates:

(79.9, 14.9, 1.7), (80.2, 15.2, 3.7), (80.4, 15.4, 5.7), (80.6, 15.6, 7.7),
 (80.8, 15.8, 9.6), (81.1, 16.1, 11.6), (81.3, 16.3, 13.6), (81.5, 16.5, 15.6),
 (81.8, 16.8, 17.5), (82.0, 17.0, 19.5), (82.2, 17.2, 21.5), (82.5, 17.5, 23.5),
 (82.7, 17.7, 25.4), (82.9, 17.9, 27.4), (83.1, 18.1, 29.4), (83.4, 18.4, 31.3),
 (83.6, 18.6, 33.3), (83.8, 18.8, 35.3), (84.1, 19.1, 37.3), (84.3, 19.3, 39.2),
 (84.5, 19.5, 41.2), (84.7, 19.7, 43.2), (85.0, 20.0, 45.2)

The points in this scale are plotted in Fig. 5. The non-uniformities are due to rounding the coordinates to one decimal digit. The swatches colored with their CIE 1976 ($L^*a^*b^*$) values yield the color scale shown in Fig. 6 below.

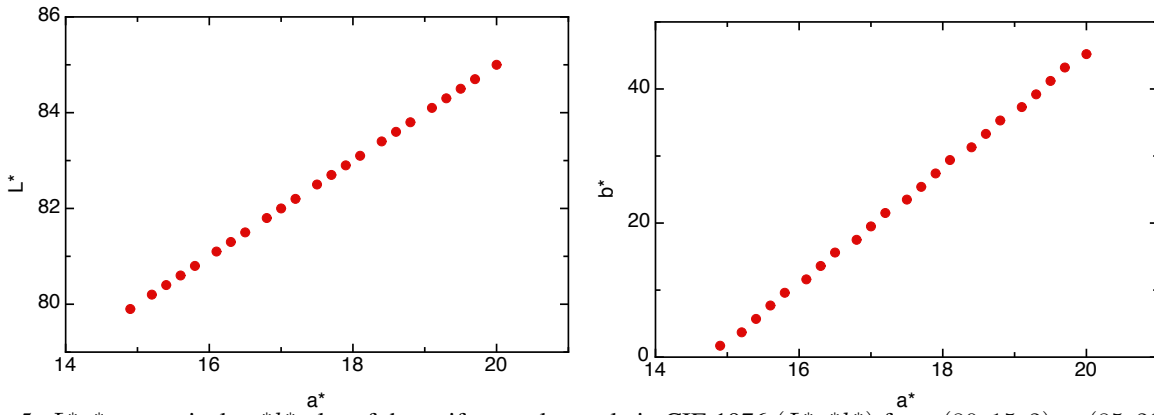


Figure 5: L^*a^* respectively a^*b^* plot of the uniform color scale in CIE 1976 ($L^*a^*b^*$) from (80, 15, 2) to (85, 20, 45) in steps of $2 \Delta E$. The non-uniformities are due to rounding the coordinates to one decimal digit.



Figure 6: A uniform color scale in CIE 1976 ($L^*a^*b^*$) from (80, 15, 2) to (85, 20, 45).

Fig. 7 shows the color coordinates of Fig. 5 when they are printed on color laser beam printer for the office market. The measurements were performed with a Gretag SPM100-II spectrophotometer with white backing, 2° observer, and illuminant D_{50} . The numerical values are shown in Tab. 1, which displays the color of each swatch, first the CIE 1976 ($L^*a^*b^*$) values sent to the printer, then—prefixed with r —the measured CIE 1976 ($L^*a^*b^*$) values.

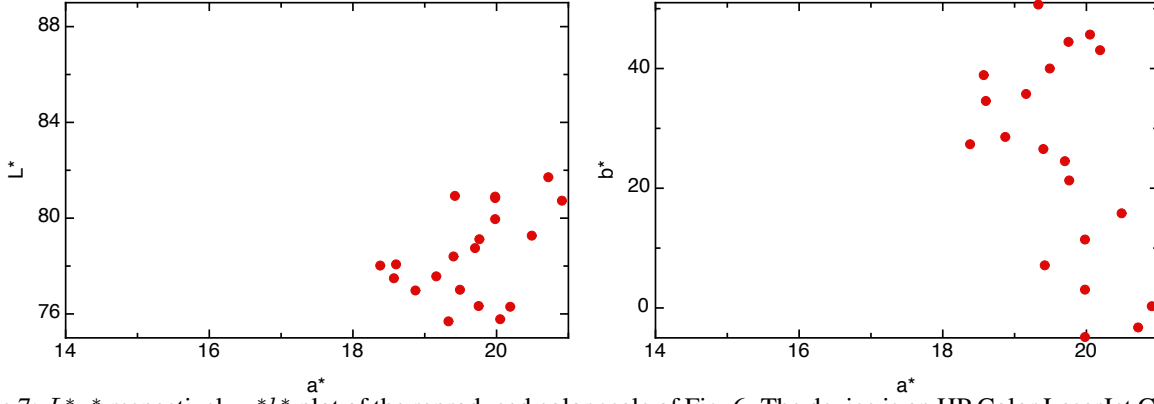


Figure 7: L^*a^* respectively a^*b^* plot of the reproduced color scale of Fig. 6. The device is an HP Color LaserJet CM6040.

order	L^*	a^*	b^*	rL^*	ra^*	rb^*
1	79.9	14.9	1.7	80.84	19.98	-4.86
2	80.2	15.2	3.7	81.71	20.72	-3.26
3	80.4	15.4	5.7	80.73	20.91	0.29
4	80.6	15.6	7.7	80.90	19.98	3.06
5	80.8	15.8	9.6	80.93	19.42	7.12
6	81.1	16.1	11.6	79.96	19.98	11.43
7	81.3	16.3	13.6	78.73	21.30	15.75
8	81.5	16.5	15.6	79.27	20.49	15.81
9	81.8	16.8	17.5	78.79	21.25	19.01
10	82.0	17.0	19.5	77.82	23.24	20.34
11	82.2	17.2	21.5	79.12	19.76	21.30
12	82.5	17.5	23.5	78.75	19.70	24.52
13	82.7	17.7	25.4	78.40	19.40	26.55
14	82.9	17.9	27.4	78.02	18.38	27.35
15	83.1	18.1	29.4	76.98	18.87	28.57
16	83.4	18.4	31.3	78.07	18.60	34.58
17	83.6	18.6	33.3	77.57	19.16	35.75
18	83.8	18.8	35.3	77.49	18.57	38.90
19	84.1	19.1	37.3	77.01	19.49	40.00
20	84.3	19.3	39.2	76.30	20.19	43.06
21	84.5	19.5	41.2	76.33	19.75	44.45
22	84.7	19.7	43.2	75.78	20.05	45.67
23	85.0	20.0	45.2	75.69	19.33	50.69

Table 1: Colors sent to the printer and measured colors.

We can see what is wrong with this printer when we plot the original versus the reproduced colors in each coordinate, shown in Fig. 8. In the range of interest—from (80, 15, 2) to (85, 20, 45)—the lightness scale is inverted and there is no modulation, just noise, on the red–green axis.

This contrasts the conventional quality metric, namely the average error ΔE_{94} , which for the data in Tab. 1 is only 5.6, not a large number in the real world of commercial or office printing.

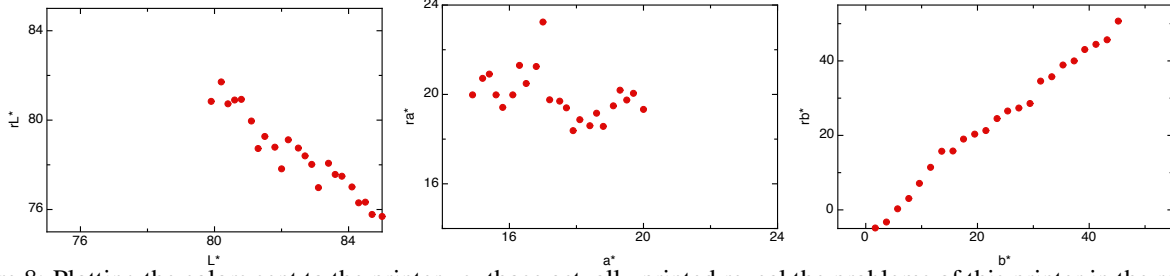


Figure 8: Plotting the colors sent to the printer vs. those actually printed reveal the problems of this printer in the range of interest from (80, 15, 2) to (85, 20, 45).

The next step is to find the permutations in a reproduction of the color scale. To this end, we project the measured colorimetric coordinates of each swatch on a line segment and calculate the parameter s for each point. If there are no permutations, the values s will be monotonically increasing.

For the line segment on which to project the measured values there are two options. The first option is to find a line through linear regression of the measured points. This method would simply assess how well the order is preserved. The second option is to project the points on the line formed by the original color scale. The second method would produce a stricter metric, because points that have different projection on the linear regression might have the same projection on the original line.

In the parametric representation, the projection of a point on a segment is determined as follows. Fig. 9 shows the problem and Fig. 10 shows the solution.

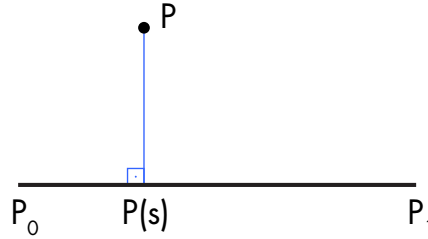


Figure 9: Given a line segment P_0P_1 and a point P , we want to find the parameter s of the point's orthogonal projection on the line segment.

In Fig. 10 we introduce the two vectors $v = (P_1 - P_0)$ and $w = (P - P_0)$, and α is the angle between the two vectors. We obtain

$$s = \frac{|w| \cos \alpha}{|v|} = \frac{w \cdot v}{v \cdot v} \quad (2)$$

The scoring is implemented in the following simplified code.

```
1 public int getTotalScore () {
2     final int length = swatch.length;
3     ColorimetricOp v, w;
4     double[] par, sorted;
5     int[] position;
6     double tolerance;
7     par = new double[length];
```

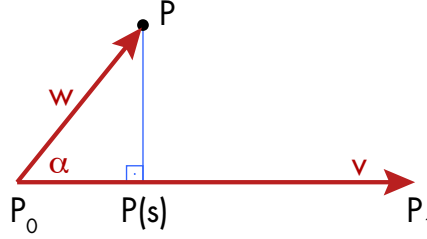


Figure 10: Introducing the two vectors v, w the parameter s can be calculated as the ratio of two scalar products.

```

8  sorted = new double[length];
9  position = new int[length];
10 scores = new int[length];
11 tolerance = 0.1 / swatch[length-1].deltaE76(swatch[0]);
12 for (int i = 0; i < length; i++) {
13     v = swatch[length-1].minus(swatch[0]);
14     w = measured[i].minus(swatch[0]);
15     par[i] = w.dot(v) / v.dot(v);
16     par[i] = (round (par[i] / tolerance)) * tolerance;
17 }
18 arraycopy (par, 0, sorted, 0, length);
19 sort(sorted);
20 for (int i = 0; i < length; i++) {
21     position[i] = binarySearch (sorted, par[i]);
22     for (int j = 0; j < i; j++)
23         if (position[i] == position[j]) position[i] = position[i] + 1;
24 };
25 for (int i = 0; i < length; i++) position[i] = position[i] + 1;
26 scores[0] = abs(position[0]-position[1]) - 1;
27 scores[length-1] = abs(position[length-1]-position[length-2]) - 1;
28 for (int i = 1; i < length-1; i++)
29     scores[i] = abs(position[i-1]-position[i]) + abs(position[i]-position[i+1]) - 2;
30 totalScore = 0;
31 for (int i = 0; i < length; i++)
32     totalScore = totalScore + scores[i];
33 return (totalScore);
34 }; // getTotalScore

```

In line 11 we provide for the instrument tolerance. This tolerance is usually given in CIE 1976 ($L^*a^*b^*$) units and for a good instrument it is rarely better than 0.1 units. In the calculation we transform this value into units of the segment corresponding to the color scale in the parameter representation.

In lines 13 and 14 we compute the vectors v, w of Fig. 10 and in line 15 we solve equation 2. Line 16 is important: we discretize the parameter's value according to the measurement's tolerance. Arguably, this is a rather aggressive value and it might turn out that one CIE 1976 ($L^*a^*b^*$) unit—which loosely corresponds to a just noticeable difference (JND)—is a more appropriate tolerance value; this would have to be determined experimentally though a psychophysics experiment.

Now that we have projected the measured values on the segment in CIE 1976 ($L^*a^*b^*$) corresponding to the color scale, we have to calculate their relative position. In lines 18–25 we simply sort the parameters and find their position through a binary search. In lines 22–24 we handle the swatches that cannot be distinguished. In principle, we should randomize their position, but here we simply order them as they are found by the binary search. Finally, in line 25 we restate the positions so they are unit-relative instead of zero-relative.

At this point we are ready to score the measurements. In lines 26 respectively 27 we score separately the first and last measurements, because they have only one neighbor. In lines 28 and 29 we score the other measurements. The total score is obtained by adding up the individual scores.

We keep an array of the individual scores for each color, so an application program can query them and plot them. Such a plot is very useful in diagnosing for which colors there is a problem. For convenience, we also provide a simple function to rate the scores by relating them to the Farnsworth-Munsell 100-hue test data.

```

1 public String getScoreRating () throws ScaleException {
2     String rating = "Unscored Device";
3     double normalScore = (double) totalScore / swatch.length * 100;
4     if (normalScore == 0) rating = "Golden Device";
5     if (normalScore > 0) rating = "Superior Device";
6     if (normalScore > 16) rating = "Average Device";
7     if (normalScore > 100) rating = "Poor Device";
8     return (rating);
9 }; // getScoreRating

```

For the printer in Tab. 1 the score is 98, which for just 23 swatches makes it a poor printer in the color region of those particular swatches.

We verify the algorithm with better quality output devices. The first device is an old laptop computer with a rather dim display (about 100 cd/m^2). The second device is a more modern office LCD display, but rather low cost and with a relatively small color gamut. The measured data is shown in Tab. 2.

order	laptop			LCD display		
	L*	a*	b*	L*	a*	b*
1	79.731	16.004	1.438	79.840	16.148	1.809
2	80.319	16.098	3.309	80.113	15.871	3.324
3	80.275	16.520	5.797	80.200	16.273	5.707
4	80.575	16.699	7.254	80.574	16.367	7.406
5	80.656	16.457	9.895	80.572	16.418	9.645
6	81.054	16.922	11.480	81.054	16.922	11.512
7	81.210	17.734	13.254	81.195	17.441	13.418
8	81.501	17.148	16.230	81.428	17.047	16.031
9	82.025	16.965	17.973	82.043	16.699	18.078
10	82.100	17.488	19.598	82.185	17.238	20.000
11	82.739	17.859	22.535	82.657	17.695	22.414
12	82.634	17.781	23.910	82.591	17.496	23.996
13	82.795	17.617	26.156	82.856	17.215	26.066
14	82.725	17.566	27.074	83.133	18.238	27.641
15	82.821	17.125	28.727	83.073	18.043	29.246
16	82.893	16.660	30.832	83.345	17.785	31.340
17	82.766	16.566	32.684	83.267	17.547	33.477
18	83.052	15.750	35.082	83.436	16.898	35.379
19	82.753	16.039	36.195	83.371	17.977	37.016
20	82.870	15.598	37.836	83.545	17.248	38.941
21	83.021	15.188	39.023	83.735	16.773	40.328
22	83.090	14.711	41.602	83.774	16.961	42.648
23	83.243	14.305	42.793	84.076	16.777	44.176

Table 2: The colors listed in Tab. 1 are measured when displayed on an old laptop LCD display and a modern office grade LCD display.

Comparing this data with that of the printer in Tab. 1 it looks much better, because the L^* and b^* values are almost monotonic, while the a^* coordinates still look problematic, but less than in the case of the printer. This is immediately apparent when the data is plotted in Fig. 11. The three graphs are for each of the CIE 1976 ($L^*a^*b^*$) coordinates and the legend is as follows: red circle: laptop display, blue diamond: LCD display, green square: printer.

The plots suggest the source of the problem. All three considered devices are office devices, not graphical art devices. Devices designed for office use are optimized for ruggedness and low total cost of ownership. Usually an sRGB gamut is sufficient and the brightness does not have to be particularly high. Fig. 11 suggests that the devices are not able to reproduce the bright vivid oranges.

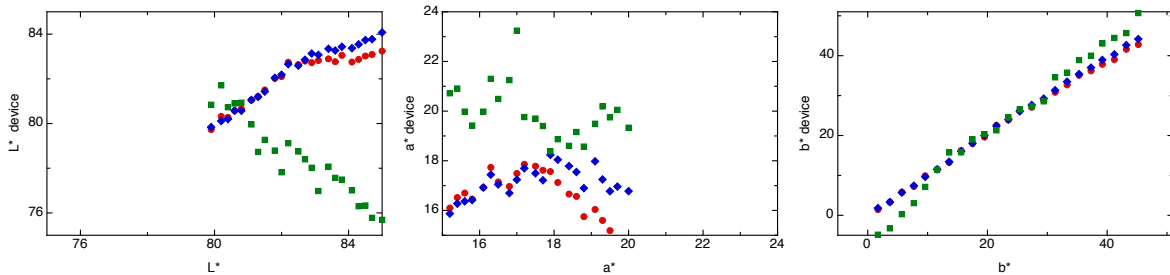


Figure 11: Plotting the colors sent to a device vs. those actually displayed reveal the problems in the range of interest from (80, 15, 2) to (85, 20, 45). Legend: red circle: laptop, blue diamond: LCD, green square: printer.

The question is: are office devices of sufficient quality for our job or do we need graphic arts grade devices? In the case of the printer, can we use a CMYK process or do we need to use a more sophisticated technology, such as HP's 6-color Indichrome®. Or might we get away using higher grade media? These are some questions we can answer with the proposed color reproduction tolerance tool.

For the laptop computer display the average error ΔE_{94} is 1.2 and for the LCD display it is 0.7. In today's thinking, a color engineer would suggest that the laptop display is borderline and that the LCD display is perfectly fine even with the gamut mapping issue.

Our color reproduction tolerance tool scores the laptop display with 16 and the LCD display with 8. Both would be considered average devices and inadequate to faithfully reproduce the color scale in Fig. 6. If the job is a display for cosmetics or for food items, better color reproduction technology would be required, such as extended gamut inks, Indichrome, or a spot color. Reproducing the target with the color scale and evaluating it with the proposed tool will permit finding the most economical but adequate process. In any case, with a normalized score of 98, the first device will hardly be adequate.

In a job there might be multiple color ranges that must be reproduced leaving reproduced color palettes intact. An example can be complexion and denims, another example can be tomato sauce, pesto and pasta. The color scales may have a different number of color swatches due to a different size critical area and increment size. To be able to compare the scores of different scales, and to be able to compare the scores with those of the well-understood Farnsworth-Munsell 100-hue test scores, the scores must be normalized. In the above examples, the absolute scores of 98, 16, and 8 correspond respectively to 426, 70, and 35.

The color reproduction tolerance tool can be used in two modes, as shown in Fig. 12. In physical mode, the swatches are measured on the print using a spectrophotometer. In simulation mode, the swatches are extracted from the raster image and their color is converted from CMYK to CIELAB using a different color management system from the one used in the RIP and a high quality ICC device characterization profile for the simulated configuration.^{15,16}

REFERENCES

- [1] Zeng, J., Lin, I.-J., Dispoto, G. J., Hoarau, E., and Beretta, G. B., "On-demand digital print services: A new commercial print paradigm as an IT service vertical," in [Annual SRII Global Conference, Connecting Services to Science & Engineering], (March 2011).
- [2] Beretta, G. B., "Spectrophotometer calibration and certification," External HPL-1999-2, HP Laboratories (January 1999). <http://www.hpl.hp.com/techreports/1999/HPL-1999-2.html>.
- [3] Phillips, J., Bajorski, P., Burns, P., Fredericks, E., and Rosen, M., "Comparing image quality of print-on-demand books and photobooks from web-based vendors," *Journal of Electronic Imaging* **19**, 011013 (January 2010). <http://dx.doi.org/10.1117/1.3271167>.

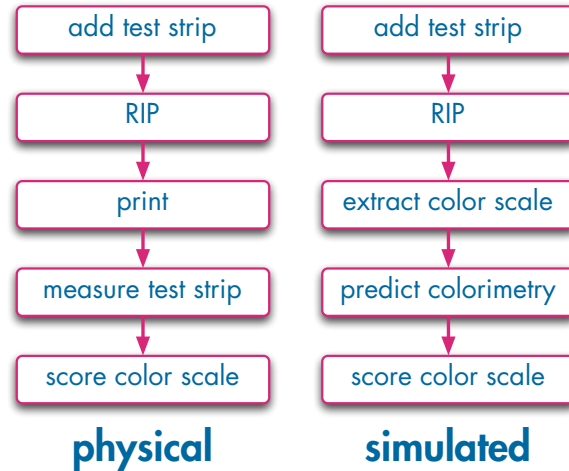


Figure 12: Workflow in physical respectively simulation mode.

- [4] Beretta, G. B., “Internet’s impact on publishing,” in [*Color Imaging: Device-Independent Color, Color Hard Copy, and Graphic Arts II*], Beretta, G. B. and Eschbach, R., eds., **3018**, 200–211, SPIE (April 1997). <http://dx.doi.org/10.1117/12.271592>.
- [5] Evans, R., “Visual processes and color photography,” *Journal of the Optical Society of America A* **3**, 579–614 (November 1943).
- [6] Beretta, G. B. and Buckley, R. R., “Trends in color imaging on the internet,” in [*9th Congress of the International Colour Association*], Chung, R. and Rodrigues, A., eds., **4421**, 471–474, AIC, SPIE, Rochester, NY (June 2001). <http://dx.doi.org/10.1117/12.464620>.
- [7] Kelly, K. L. and Judd, D. B., [*The ISCC–NBS Method of Designating Colors and a Dictionary of Color Names*], vol. 553, National Bureau of Standards, Washington (November 1955).
- [8] ASTM Committee E12.11, “Standard practice for establishing color and gloss tolerances.” ASTM Standard D 3134–97 (2003) (December 2003).
- [9] Xerox System Institute, [*Color Encoding Standard*], vol. XNSS 289005, Xerox Corporation, Sunnyvale (May 1990). Xerox System Integration Standard.
- [10] ASTM Committee E12.11, “Standard guide for selection, evaluation, and training of observers.” ASTM Standard E 1499–97 (2003) (December 2003).
- [11] Farnsworth, D., “The Farnsworth-Munsell 100-hue and dichotomous tests for color vision,” *Journal of the Optical Society of America* **33**, 568–578 (October 1943).
- [12] Farnsworth, D., *The Farnsworth-Munsell 100-hue test for the examination of color discrimination*. Munsell Color, Macbeth, Division of Kollmorgen Instrument Corp., 405 Little Britain Road, New Windsor, NY, 2nd ed. (1957). Manual.
- [13] Beretta, G. B., Dispoto, G. J., Hoarau, E., Lin, I.-J., and Zeng, A., “ICC profiles: are we better off without them?,” in [*Color Imaging XVI: Displaying, Processing, Hardcopy, and Applications*], Eschbach, R., Marcu, G. G., and Rizzi, A., eds., **7866**, 78660R, SPIE, San Francisco Airport, California (January 2011). <http://dx.doi.org/doi/10.1117/12.880240>.
- [14] Falkenstern, K., Bonnier, N., Brettel, H., Felhi, M., and Viénot, F., “Adaptively selecting a printer color workflow,” in [*Nineteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications*], Woolfe, G. and Ferweda, J., eds., 205–210 (November 2011).
- [15] Beretta, G. B., “Meta-palette.” Xerox PARC Video (June 12 1988). <http://youtu.be/sR8UzOQF9VQ>.
- [16] Beretta, G. B., “Graphical user interface for editing a palette of colors.” U.S. Statutory Invention Registration H1506 (December 5 1995).