Optimal Test Charts

Thorsten Braun[1] and Jodi Alejandro[2]

Keywords: characteristics, delta-e (ΔE), error, measurement, mathematical

Abstract

This paper will discuss the analysis of determining the ideal number of color patches to produce 'optimal test charts'. A test chart consisting of 8,000 patches does not necessarily mean that the fingerprint or characteristic of an output system is more accurate than using a test chart with less patches such as 500. How much quality is really gained by increasing the number of patches?

We will introduce a simple formula predicting the expected quality when varying the test chart size and discuss the results.

Introduction

Currently, there is not a standard on the optimal number of color patches for test charts. A single test chart (of reasonable size) can be evaluated to analyze the quality for a given output system. Based on a few parameters it can be calculated how many patches a test chart should have to achieve a given quality.

Choosing the wrong test chart size may cause serious problems. If test charts are too small the production quality may be critical. On the other hand, if test charts are too large it may cause high costs with no benefits.

Procedure

A variety of downscaled data sets (or subsets) were used to model the output system and were compared to the original data set. This process was used to evaluate the quality of test charts with fewer patches such as the one, two, and three color combinations. For a visual representation of the process, refer to Appendix 1, Figure 3. The results were then inputted into spreadsheets to create graphs to display the maximum, average, sigma, 95th percentile and approximation of the 95th percentile (see Equation 1).

---

[1]ColorLogic GmbH; [2]CrossXColor, Inc.

As a criterion for the quality the dE-2000 method was selected and a tolerance of 1.0 dE-2000 or less was used for the target deviation tolerance. A <1.0 dE-2000 indicates that the color difference is barely noticeable. A 1.0 - 2.0 dE-2000 indicates a color difference that is noticeable, but only if it is through close observation (Habekost, 2013). We will see that it is possible to achieve a dE-2000 of 1.0 or less, especially with fewer color channels.

In order to evaluate the best possible quality, we took test chart data with large amounts of color patches, which were gathered from printed test charts. Existing measurement data sets from customers were used (with approval). Using manipulated process colors and gradients, we also printed more extreme data of one, two, three, and four color (1CLR, 2CLR, 3CLR, 4CLR) combination test charts consisting of red, blue, yellow, and black gradients. Some data sets were optimized to remove or correct any printing errors or anomalies.

The 1CLR test charts were gradients from white to a primary color using 52 color patches. The 2CLR combinations used black and red; black and blue; black and yellow; blue and yellow; blue and red; finally, red and yellow. Each of these test charts had 600 color patches. The 3CLR combinations used black, red, and blue; black, red, and yellow; black, blue, and yellow; and blue, yellow, and red. Each of the three color combination test charts had 1,890 color patches. The 4CLR test chart using black, red, blue, and yellow had 5,050 color patches.

These custom test charts were created using manipulated gradients to produce complex data sets. The size of the color patches was 8.0 mm (width) x 8.0 mm (height) with no gap size between the patches and a random distribution of the patches was applied. This allowed an equal number of patches per page. The test charts were printed on proofing paper white semi-matte using an Epson SCP5000SE printer with the color management print driver turned off.

Measurement data based on the following color spaces: RGB, CMYK, and combinations of multicolor (five, six, and seven), were supplied to us from clients. The number of patches from each of these test charts varied from 1,495 to almost 10,000. The measurement data were also optimized to compensate for production and measurement issues.

Results & Discussion

To determine the optimal number of patches for a test chart depends on the print production and the desired quality. Two criteria were used to determine the optimal number of patches: the 95th percentile and <1.00 dE-2000. Varying the criteria showed the same characteristic and the following results are also applicable for other criteria (such as average, maximum, dE-76).

The results of the test chart analysis were inputted into spreadsheets to create graphs to visualize common characteristics. In Figure 1, three curves are displayed, which represents the typical characteristics from the majority of measurement data gathered. The green curve represents the 95th percentile which shows that there is a higher dE-2000 with fewer patches, but decreases with additional patches until it reaches a certain point in which it plateaus. In the 95th percentile, each value is calculated and the approximation is a specific curve applying the values.

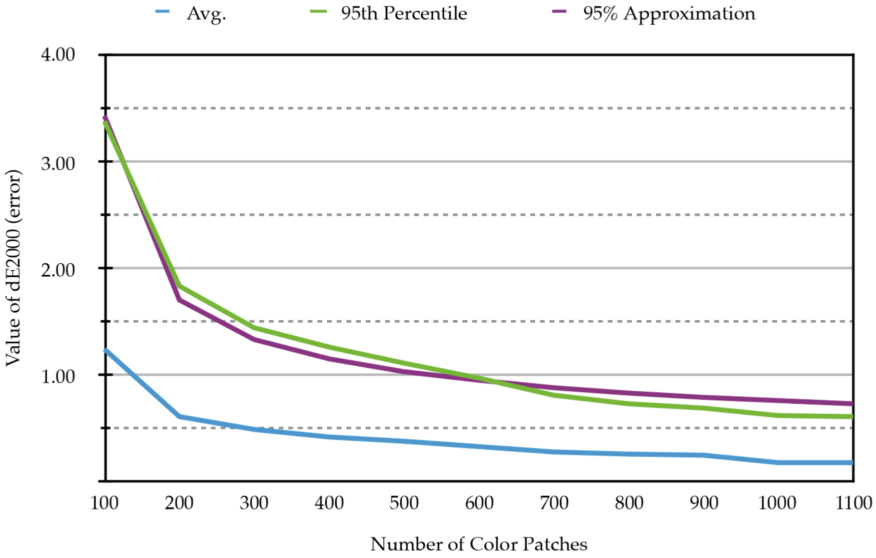

Figure 1. *Common curve characteristics*

For the 1CLR test charts, the measurement data of <1.0 dE-2000 was achieved with only 14 color patches. This means that a test chart with single color gradients would only need 14 color patches and not 52 in order to achieve a <1.0 dE-2000.

Increasing the number of color channels significantly increased the number of required patches in test charts to achieve a certain quality. For the 2CLR test charts, a <1.0 dE-2000 was achieved at about 280 color patches. The 3CLR test data already required about 1,500 color patches. Analyzing the 4CLR data showed that about 4,000 color patches were required to reach <1.0 dE-2000.

We analyzed 10 RGB test charts, which produced various results. In some cases, the limit of 1.0 dE-2000 could be reached with only 600 patches. In other cases, a quantity of 1,5000 patches were not sufficient, but in all cases the shape of the curves were similar.

The same could be observed analyzing the CMYK and multicolor data sets: the number of required color patches varied from system to system. However, we could see a typical relation between the number of patches and the resulting deviation between upscaled test chart data and actual measurement data.

In the majority of the test cases, the error could be estimated by a simple formula as shown in Equation 1. The factor k represents the constant, n number of patches, g is the system complexity, and e is the error. The error is the deviation/difference between predicted results and real measurements.

$$e = \frac{k}{n^g}$$

Equation 1. *Describes the relationship between the error and the number of patches.*

This formula uses two parameters, an exponent indicating the complexity of the output system and a scaling factor. The formula can easily be reversed to calculate the number of patches a test chart should have to achieve a given quality for an 'Optimal Test Chart.' For example,

Count	Max	Avg	Sigma	95%-Percentile	95%-Approximation
2	12.39	7.80	4.15	12.23	12.17
4	4.44	1.65	1.42	4.05	4.76
6	4.43	1.61	1.45	4.04	2.75
8	2.10	0.84	0.66	2.02	1.86
10	1.26	0.45	0.32	1.07	1.38
12	1.32	0.34	0.27	0.77	1.08
14	0.62	0.26	0.16	0.53	0.87
16	0.62	0.24	0.17	0.53	0.73
18	0.84	0.24	0.17	0.53	0.62
20	0.84	0.23	0.17	0.49	0.54
22	0.84	0.22	0.18	0.50	0.47
24	0.84	0.21	0.19	0.50	0.42
26	0.82	0.19	0.17	0.49	0.38
28	0.50	0.17	0.15	0.45	0.34
30	0.50	0.15	0.15	0.45	0.31

Table 1. *A data subset from an original data of 52 patches.*

Table 2 displays the color combinations, its respective process colors, the data for the factor (*k*) and exponent (*g*) variables. For RGB and CMYK, the table contains a representative data set as an example, only.

Test Charts	k = factor	g = exponent
R - 1 CLR	30.51	1.33
B - 1 CLR	59.57	1.62
Y - 1 CLR	21.20	1.14
K - 1 CLR	39.70	1.71
RK - 2 CLR	42.69	0.74
BK - 2 CLR	33.75	0.68
YK - 2 CLR	27.39	0.65
BY - 2 CLR	49.16	0.70
BR - 2 CLR	100.00	0.71
RY - 2 CLR	17.69	0.59
RBK - 3 CLR	100.00	0.71
RYK - 3 CLR	60.47	0.63
BYK - 3 CLR	56.51	0.63
BYR - 3 CLR	60.53	0.57
RBYK - 4 CLR	32.86	0.50
CMYKB - 5 CLR	53.12	0.50
CMYKR - 5 CLR	28.55	0.42
CMYKOG - 6 CLR	100.00	0.59
CMYKOGB - 7 CLR	100.00	0.57
CMYK	11.57	0.38
RGB	88.13	0.68

Table 2. *Parameters for various process colors*

Legend:

R - Red	B - Blue	K - Black	M - Magenta
G - Green	Y - Yellow	C - Cyan	O - Orange

Using the formula, it is easy to calculate the expected quality for a given number of patches or to calculate the required number of patches to achieve a given quality. For example, to determine the error of <1.0 dE-2000 of a single color of red gradients, we see that the $k = 30.51$ and $g = 1.33$. The variable n is the number of patches and in this case, we may check the predicted quality for 10 patches. We can input the following parameters into the equation (see Example 1).

$$e = \frac{30.51}{10^{1.33}}$$

Example 1

Thus, the resulting error of a patch set with 10 can be estimated as $e = 1.43$, which is close to the actual value. We can determine the error for various quantities of patches.

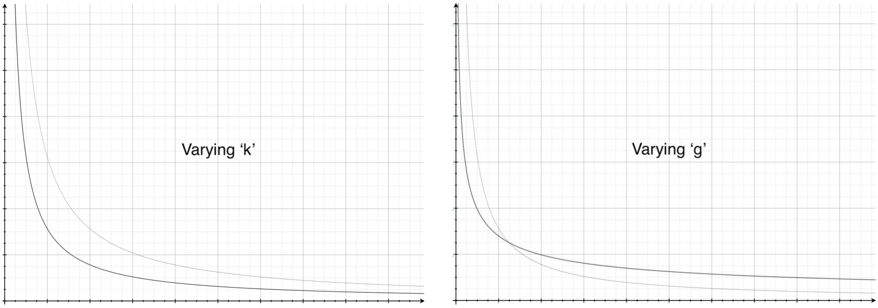

Figure 2. *A general characteristic of the factor and system complexity curves.*

It can be noticed that the factor varies significantly from data set to data set whereas the exponent depends mainly on the number of process colors. In Figure 2, the two parameters have different impact on the curves. The parameter g defines the shape of the curve - the higher the value the steeper the curve. This also indicates that increasing the number of patches has a strong impact on the error, but up to a certain point. The factor scales the curve proportionally. It can be understood as a scaling factor for the expected error. For output systems with the same characteristic, the error is directly related to the gamut size.

The number of possible color combinations increases exponentially by the number of process colors. Therefore it may be assumed that also the number of patches to fingerprint an output system may increase exponentially by the number of process colors. However, the analyzed systems indicate something different: The exponent g changes significantly for systems with up to three process colors, only. The use of more than four process colors does not change the exponent significantly. The factor k varies significantly by the gamut size.

Initially, we expected that increasing the quality of multicolor test charts would require an exponentially higher number of patches compared to CMYK data. But as we can see from the exponent g this is not the case. This implies that the number of required patches achieving the same quality for multicolor compared to CMYK data is proportional! The most important observation is that the number of patches does not increase exponentially by the number of process colors. This means that even for multicolor it is possible to get an accurate fingerprint.

Use Cases

This methodology allows to estimate the potential quality of different test chart sizes. There are two main use cases:

1. Accuracy of a fingerprint. Given the measurements for an output process we can estimate how precise this process can be simulated. This is an important indicator e.g for proofing. The accuracy of a proof is limited by the accuracy of test chart.

2. Calculation of an optimal chart size. Obviously, the optimal patch number depends on the output process and the required quality. To determine the optimal patch number a standard chart can be printed and evaluated. Based on the parameters we can calculate a recommended patch number to match the quality requirements. Note that this number may be lower or higher than the chart size used for evaluation.

Conclusion

A high number of color patches does not necessarily equate to a significantly better quality. Depending on the number of color channels, output process and requirements an optimal test chart could have as little as 600 patches to as many as 4,000.

Based on the measurements of a single test chart a simple formula can be used to estimate the expected error when varying the number of patches. Using this formula one may directly calculate the optimal number of patches for a given maximum error.

Optimal test charts can be created and used for different printing applications such as textile, offset, digital, and flexo printing. By determining an optimal test chart it would save time, paper and ink. It would not be necessary to print 5,000 or more color patches for a CMYK test proof when 2,500 or 3,500 patches would achieve the same result.

References

1. Habekost, Martin
   2013 "Which color differencing equation should be used?" *International Circular of Graphic Education and Research, Science & Technology, No. 6*, pp. 20–33

2. Tatol, Maciej and Mokrzycki, Wojciech
   2011 "Color difference Delta E-A survey", *Machine Graphics and Vision*, pp. 15, pp.18–19

Acknowledgements

The authors would like to acknowledge the following individuals for their contribution of measurement data for our analysis: Charles Brown, Michael Strickler, and Joseph Tschudi.

Appendix 1

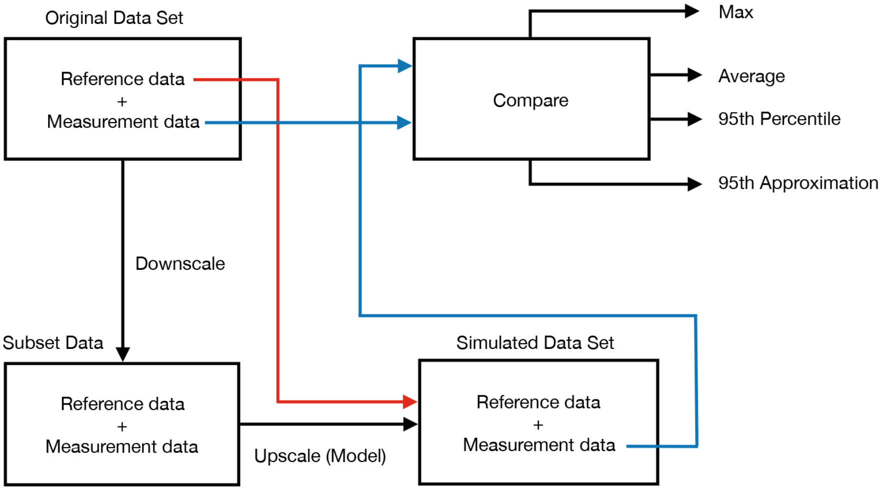

Figure 3: *From the original data set, a downscaled (subset) data set was used, which was then upscaled to the original reference data. The quantitative values of the measurements from the original and upscaled data sets were compared.*