Fogra Roses - Developing a colour difference dataset for the graphic arts

Andreas Kraushaar; Fogra; Munich, Germany

Abstract

A new colour difference dataset has been developed to address the specific needs of the graphic arts industry. An experiment has been designed with 1288 colour pairs. Concretely 46 colour centres of the Fogra media wedge and 28 randomly intra- and supra threshold test colours around each of them have been observed by 32 observers. The pair comparison constant stimuli method has been used by means of three grey reference pairs comprising colour differences of $\Delta L=1$, 3 and 5. The performance of established and modern colour difference formulae has been evaluated with the new dataset. DIN99o and CIEDE2000 outperformed the other tested metrics.

Introduction - Graphic arts colour appraisal

In the past years colour difference metrics has often been used by means of established pass/fail-tolerances such as PSO conformance protocols [1]. Hence the lack of perceptual uniformity of the used CIE1976 colour difference was of secondary concern. The same applies for packaging applications however using tighter tolerances. A visual assessment has to be conducted separately in order to express the perceived colour difference. In advent of digital printing and process agnostic standardization, the instrumental evaluation of small colour difference that correlates with the human perception is of great importance. It allows for both the establishment of pass/fail-tolerances typically used in the industry (to be extended toward a multi-tolerance schema) and an associated communication of the perceived colour difference to be expected. That is vital for conveying the concept of "Printing the Expected", which is the basic concept of the newly developed ProcessStandard Digital (PSD) by Fogra [2].

The obvious candidate for a perceptual uniform evaluation would be to use modern colour difference formulas by means of either an approach that weights the pertinent colour difference components or a transformation of the CIELAB colour values into a perceptual uniform colour spaces such as DIN99 or its optimized version DIN99o [3]. Colour difference formulas of the first category are mostly based on the CIELAB colour space while weighting the different colour components namely ΔL^*, ΔC^*ab and ΔH^*_{ab} as used in CMC [4], CIE94 [5], "Luebbe" including their corrected version "Luebbe_corr" [6], CIEDE2000 [7] or modern "upgraded" formulas adding a correction term [8].

However all of these metrics are based on colour difference datasets with the most important being RIT-DuPont [9,10], "Witt" [11], "Leeds", [12] and "BFD-P" [13]. All of these datasets however are based on psychophysical experiments based on the CIE recommendation that deviate from the viewing conditions [14] used and standardized in the graphic arts for decades [15, 16].

The most important differences are the reference illuminant of D65 where the graphic arts used D50 simulators, the illumination level of 1000 lx where typical 2000 lx are used and the sample size its preparation and presentation. While prints often comprise slight halftone structures most of the current datasets are based on (glossy) painted samples. For that reason a psychophysical experiment will be presented that was designed to reflect the graphic arts viewing conditions and sample preparations. It was used to derive a new colour difference dataset that mimics the requirements of the graphic arts and can be used to evaluate the performance of the prescribed colour difference metrics.

Experimental Design

In order to represent average graphic arts viewing conditions an ISO 3664:2009 compliant viewing cabinet (Virtual Proof Station from Just Normlicht) was used, see Figure 1.

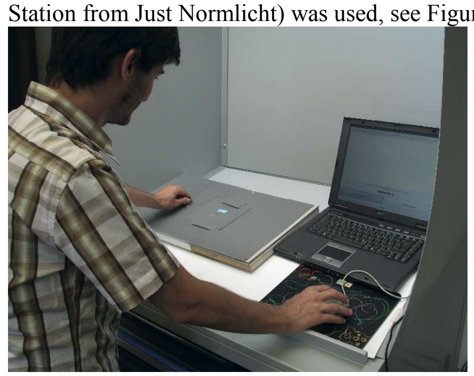

***Figure 1.** The sample colours are presented on a grey background. The computer on the right allowed for making the pair comparison judgments. The left hand rotates the turning wheel beneath the grey cardboard.*

Aiming for a dataset that comprises intra- and supratreshold colour difference (reflecting small and medium colour differences typically associated with a $\Delta E^*_{ab} \leq 5$) 46 colour centres have been selected by using the Fogra Media Wedge CMYK 2 ($\psi_1, .., \psi_{46}$). The corresponding CIELAB values have been derived by using the characterization dataset FOGRA39 [17] that allows for an unambiguous transformation from CMYK to CIELAB (e.g. by using the published ISOCoatedV2 profile and its abs. colorimetrical transformation). Presenting representative colour pairs for each colour centre with an angle of subtend of approximately 2 degree (and a viewing distance of ca. 50 cm) it was found that 28 colour samples could be placed on a disc-like print. In order to rate the perceived colour differences for each of the j=1,..,1288 (46*28) colour pairs the constant stimuli pair comparison method was chosen. The constant stimuli are represented by three neutral grey colours ($\Delta L^* \approx 1$, 3 and 5) compared against a midgrey $L^*= 50$ with CIEab = 0. The three reference pairs are termed ("50:51", "50:53", "50:55"). A schematic visualization of the experimental design is depicted in Figure 2.

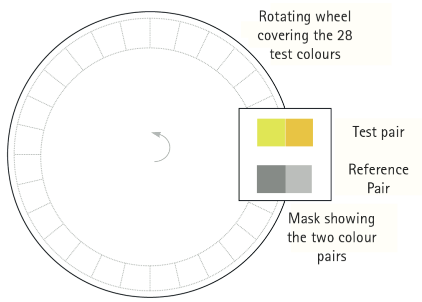

***Figure 2**. Experimental setup. The grey reference below will be shown simultaneously with the test pair. Here for the centre ψ_8=(0,0,70,0) is shown. The 28 test colours are placed on a rotating wheel to allow for only colour to be visible at a time (next to the colour centre of interest).*

Although the disc-like print is not seen by the observer completely (only one patch compared against the associated colour centre) the resulting dataset has been named Fogra roses due to its similarity with a rose as shown in Figure 3.

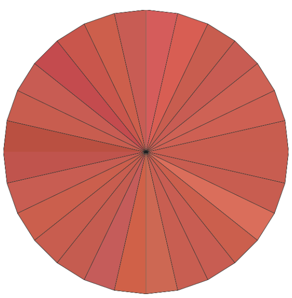

***Figure 3**. Visualisation of the 28 test colours around the (red) colour centre ψ_{28}=(0,70,70,0). Since its similarity with a rose the resulting dataset has been termed Fogra roses. A grey reproduction might not be conveying the information.*

The 28 colours around each of the 46 centres has been selected by randomly chosen colour difference $\Delta E^*_{ab} \leq$ 10 for a systematic representation of the inclination and azimuth in the CIELAB space.

The chosen 32 observers had normal vision and were tested with the FM 100 Hue Test. Details can be found in [18]. The prints have been made with a digital proofing system using an inkjet printer (Epson Stylus Pro 4800) and GMG SM 250 semi-matt proofing stock without optical brightener agents. This allows for colour accurate reproduction. However for the final modelling the actual readings (ISO 13655:2009; M0; white backing) have been used. The question to the observer were as follows:

Please compare the test colour pair with the reference grey pair. If you find the colour difference between the test colour pair greater (worse) than the reference colour pair please select "YES". If it is equal or less please select "No".

Results & Discussion

For each of the j=1288 colour pairs three observations are done by each observer. Repeated observations or multiple observers allow for deriving a so-called psychometrical curve [19] by means of a frequency matrix and hence proportional values p. There are different methods to model a psychometrical curve such as Logit, Log-Log or cumulated Gaussian. In this work the latter method has been proved to work best. Figure 4 is providing more details on how to interpret such a psychometrical curve.

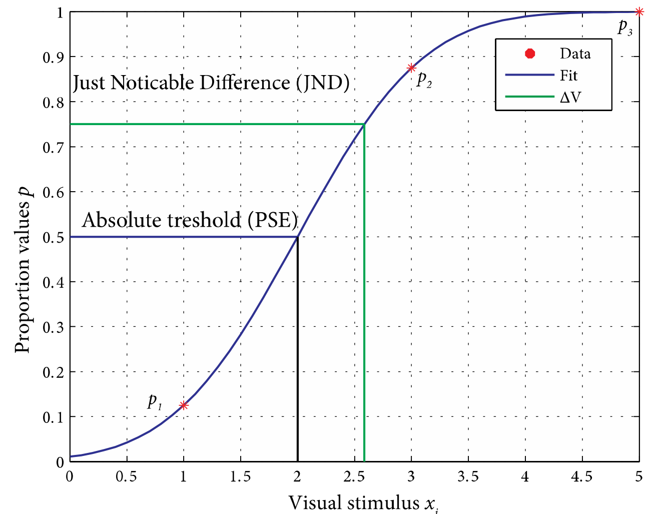

***Figure 4.** Psychometrical curve (blue line) for an arbitrary colour pair. The red points illustrate the visual difference ΔV of the three grey reference pairs with x_1: ΔV=1, x_2: ΔV=3 and x_3: ΔV=5 with the corresponding proportional values p_1, p_2 and p_3. For instance, p_1 of 0.12 corresponds to 12 % of the observations that rate the given test colour pair as equal or smaller then the smallest grey reference pair. The psychometric curve has been modelled by minimizing the quadratic norm of a cumulated Gaussian distribution function. The point of subjective equality is the absolute threshold to found by 50% of the observations. The "true" colour or visual difference can be found by finding the visual stimulus that corresponds to 75% of the observations. This can be though of a just noticeable difference.*

The "true" colour difference ΔV for each of the j=1,..,1288 colour pairs will be determined in two steps. First the psychometrical function F(x) with $F \in \Re \rightarrow [0,1]$ will be fitted by means of the Matlab function "lsqcurvefit" that results in μ_j and σ_j the for each pair by minimizing:

$$\sum_{i=1}^{3}\left(F_{\mu_j,\sigma_j}(x_i)-p_i\right)^2 \tag{1}$$

Here p_i represents the proportional values with i={1,2,3} for each of the 1288 colour pairs. In the second step each psychometrical curve will be inverted to find ΔV namely the corresponding value for p=0.75 for all 28 colour patches around each of the 46 colour centres hence 1288 colour pairs.

$$\Delta V_j = x_j = F^{-1}(p) = \left\{x : F_{\mu_j,\sigma_j}(x) = p\right\} \tag{2}$$

The 75% limit has been chosen as the threshold since the experimental design represents a 2/3-1/3 proportion (“smaller or equal” versus than “greater as”) by means of the defined pair comparison. In addition it is recommended by ISO [20].

Evaluating the outcomes of the psychophysical experiments it was found that not all observations were plausible. In these cases the optimization is not well-defined and subject for noisy results. In order to avoid such behaviour the following 4 restrictions have been established:

- Monotony, i.e. $p_2 \geq p_1$ and $p_3 \geq p_2$
- no double occurrence of p=0 or p=1, i.e. $p_2 = p_1 = 0$ or $p_3 = p_2 = 1$
- Minimal proportion for highest grey p_3, i.e. $p_3 \geq 0.7$
- Maximum proportion for smallest grey pair p_1, i.e. $p_1 \leq 0.5$

Based on these empirical derived optimization constraints 488 colour pairs could be identified as reliable. They constitute the basis of the established dataset called “Fogra-Roses”. Alongside the experimental details and all observer ratings it can be found on the Fogra website.

In order to evaluate the performance of modern colour difference metrics two established methods have been used to verify the variation between ΔV and ΔE, where ΔE represents the general form for a colour difference metric. The performance factor [21] and the SRESS metric [22] have been used. The performance factor PF/3 is not a single figure of merit by itself. It simply combines the measures y-metric, CV-metric and V_{AB}-metric. Both PF/3 and STRESS have dimensionless units whereby a value of e.g. PF/3=40 represent a 40% deviation between the perceived and computed colour difference. In other words the smaller the value the better. A value of 0 represents a perfect correlation. The results of the evaluation are shown in Figure 5. In addition to the previous mentioned colour difference metrics a fairly new method has been used namely the usage of colour appearance models termed $\Delta E_{CIECAM02}$.

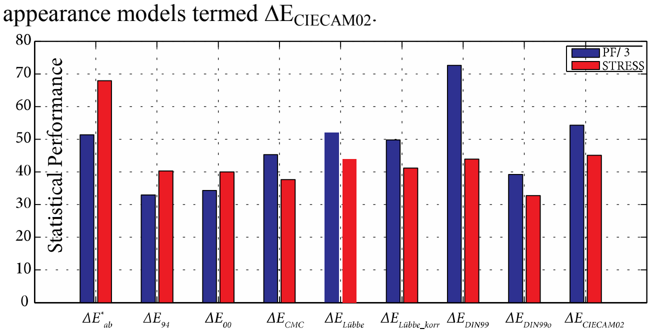

***Figure 5**. Result of the PF73 and STRESS performance of the tested colour difference metrics by using the developed Fogra-roses dataset comprising 488 colour pairs. Information on the statistical significance can be found in [18].*

It can be seen in Figure 5 that DIN99o and CIEDE2000 performs best when focussing on the STRESS metric. As expected the CIELAB1976 colour difference (ΔE^*_{ab}) metric performs worst. However it was interesting to see how to optimization of DIN99 improved the performance for the graphic arts dataset used here. In light of global usage of upcoming process agnostic digital printing standardization (not restricted to ISO standards) the usage of CIEDE2000 is preferred over DIN99o. Since adopting national standards is quite difficult and in light of the present implementation of CIEDE2000 in many programs Fogra recommends the usage of CIEDE2000 for upcoming expert opinions and research projects.

Summary & Conclusion

In this paper a novel colour difference dataset has been presented. It was motivated to reflect exactly the needs of the graphic arts industry when appraisal images (focussing on colour). This was required since the graphic arts viewing condition and sample preparation and presentation differ severely from the CIE recommendation. The dataset was derived by evaluation 28 neighbourhood colours around 46 colour centres hence 1288 colour pairs by using a pair comparison constant stimulus method. Three grey reference pairs with a predefined visual colour difference of ΔL^*=1,3 and 5 served as the stimulus. After applying empirical derived constraints 488 colour pairs have been identified as reliable. These constitute the dataset call “Fogra-Roses”.

It was found that the agreement among the observers was significantly improved when modifying the question from: “Do you perceive the [heterochromatic] colour difference greater then the reference grey pair.” into: “Do you perceive the [heterochromatic] colour difference worse then the reference grey pair.” It can be speculated that the heterochromatic comparison is too ambiguous when asked for a per dimension-based evaluation (triggered by the word “greater”). The word “worse” seems to condense the different dimensions of the colour difference more intuitively.

It was also found that the modern colour difference metrics such as CIEDE2000 perform somehow worse on the dataset developed here. This can be expected since it was tailored toward the prescribed datasets. That means that there is some room for improvement in the optimization of the present colour difference metrics for the usage of printed matter assessed under graphic arts condition.

ACKNOWLEDGMENTS

The project "Evaluation of modern colour difference equations" [Fogra No. 60.054] was supported within the programme of the Federal Ministry of Economics and Technology for the support of “Industrielle Gemeinschaftsforschung [IGF]” via the German Federation of Industrial Research Association [AiF/No. 14893 N].

References

[1] Dolezalek, F.; ProcessStandard Offset printing (PSO); Berlin: bvdm, www.bvdm-online.de
[2] Kraushaar, A.; ProcessStandard Digital (PSD), München: Fogra, www.Fogra.org
[3] Cui, G.; Luo M. R. ; Rigg, B.; Roesler, G.; Witt, K.; Uniform Colour Spaces Based on the DIN99 Colour-Difference Formula; Color Research & Application, 2002, Vol. 27, Nr. 4, p. 282-290
[4] Steen, D.; Dupont, D.; Defining a Practical Method of Ascertaining Textile Color Acceptability; Color Research & Application, 2002, Vol. 27, Nr. 6, p. 391-398
[5] McDonald, R.; Smith, K. J.; CIE94 - A new colour-difference formula; Journal of the Society of Dyers and Colourists, 1995, Nr. 111, S. 376-379
[6] Lübbe, E.; Verbesserung der Farbabstandsformel CIELAB für Druckfarben; Die Farbe, Göttingen: Muster-Schmidt Verlag, 1995, Band 41, p. 25 ff

[7] CIE, Improvement to industrial colour-difference evaluation. Vienna: CIE Publication No. 142-2001, Central Bureau of the CIE; 2001
[8] Lissner I; Urban, P.; Upgrading color-difference formulas; Journal of the Optical Society of America, A 27, p. 1620-1629 (2010)
[9] Berns, R. S.; Alman, D. H.; Reniff, L.; Snyder, G. D.; Rosen, M. R.; Visual determination of suprathreshold colour-difference tolerances using probit analysis, Color Research & Application, 1991, Vol. 16, Nr. 5, p. 297-316
[10] Berns, R. S., RIT-DuPont supra-threshold color-tolerance individual color-difference pair dataset, Color Research & Application, 2010, Vol. 35, Nr. 4, p. 274-283
[11] Witt, K., Geometric relations between scales of small colour-differences, Color Research & Application, 1999, Vol. 24, Nr. 2, p. 78-92
[12] Kim, D. H.; Nobbs, J. H., New weighting functions for the weighted CIELAB colour-difference formula, Colour 1997, Kyoto, Tagungsunterlagen, p. 446-449
[13] Luo, M. R.; Rigg, B., Chromaticity-discrimination ellipses for surface colours, Color Research & Application, 1986, Vol. 11, Nr. 1, p. 25-42
[14] Norm ISO 3664 : 2009, Viewing conditions -- Graphic technology and photography, Beuth-Verlag, 10772 Berlin
[15] R. G. Kuehni; Analysis of Five Sets of Color Difference Data; Color Research & Application, 2001, Vol. 26, Nr. 2, p. 141-150
[16] Luo MR, Cui G, Rigg B.; The development of the CIE 2000 colourdifference formula: CIEDE2000. Color Res Appl 2001; 26:340 -350.
[17] N. N.; FOGRA39 – Characterization data for commercial offset printing according to ISO 12647-2:2007 on coated stock (PT 1/2); Screening according 60/cm http://www.fogra.org/fogra-standardisierung/fogra-charakterisierungsdaten/a-icc-de.html
[18] Kraushaar, A.; Beiträge zur qualitativen Beurteilung und Verbesserung des Farbmanagements in der grafischen Industrie; Aachen: RWTH Aachen, Fakultät für Elektrotechnik und Informationstechnik, Dissertation, 2012, ISBN 978-3-8440-1470-9
[19] Engeldrum, P.; Psychometric Scaling - A Toolkit for Imaging System Development; Winchester: Imcotec Press, 2000, ISBN 0-9678706-0-7
[20] Norm ISO 20462-1 : 2005; Photography - Psychophysical experimental methods for estimating image quality - Part 1: Overview of psychophysical elements
Beuth-Verlag, 10772 Berlin
[21] Melgosa, M.; Measurement of the relationship between percieved and computed color differences; Journal of the Optical Society of America, 2007, Vol. 24, Nr. 7, p. 1823-1829
[22] Melgosa, M.; Huertas, R.; Berns, R. S.; Performance if recent advanced color-difference formulas using the standardized residual sum of squares; Journal of the Optical Society of America, 2007, Vol. 25, Nr. 7, p. 1828-1834

Author Biography

Andreas Kraushaar is head of the prepress division at Fogra, Munich, Germany. He joined Fogra in 2001. Within the prepress department he is responsible for research in the fields of colour management and digital printing. He holds the position of the convener of working group 3 (process control and related metrology) within ISO TC 130 (graphic technology) since 2005.

He received a master degree in media technology from the Technical University Ilmenau and his PhD in colour science from the RWTH university Aachen. He is a member of IS&T