

Evaluation of Color Differences: Use of LCD monitor

Iris Sprow, Tobias Stamm, Peter Zolliker, Laboratory for Media Technology; Swiss Federal Laboratory for Materials Testing and Research (EMPA), Dübendorf, Switzerland

Abstract

The use of LCD displays as a test platform for the evaluation for perceived color differences is examined. The setup and verification of an accurate color reproduction workflow is presented. As a first application we compare a monitor based color difference test with a corresponding already existing test based on printed samples. In view of the results, we regard the digital method as a good candidate for extended color perception studies allowing more flexible test setups compared to tests using surface colors.

Motivation

Colorimetry as it is known nowadays, is largely based on the understanding of human vision, which is studied through psychovisual tests. For quality control and development it is important to include human visual response which still nowadays is an essential part of many projects. This demands visual evaluations which involves judging hundreds of samples by observers. Information technology provides the possibility to embed such tests in a digital environment. This means, with the use of a monitor a more flexible test set up for visual judgments can be designed.

Existing monitor-based studies focused on colorimetric tolerances for real-world images [1] [2], others focused on evaluating color patches in regard to color difference formulae [3] and threshold tolerances for CRT-generated stimuli, see [4].

While former CRT display based studies often evaluated perceived thresholds (JNDs), the aim of the current study is to explore the potential for evaluating perceived color differences as a scaling study in comparison to an actual user study. For this, we re-implement a recent study conducted by the Fogra Graphic Technology Research Association (short 'Fogra') [5] but this time using LCD displays instead of printed samples. The main challenge to accomplish this is to design and control the digital workflow for displaying colors.

The following sections outline the involved steps of the current work: In *Technical Set-up* we describe the general test set-up, focus on the challenges of displaying colors accurately on LCD displays and the employment of a web browser to display colors. Section *Verification* refers to the technical achieved accuracy of LCD displays. The section *Application: Fogra Color Difference Test* describes the employed test set-up which was adopted to enable a comparison of newly gathered LCD based visual data with already existing print based visual response data. The results from this comparison are to be found in *Evaluation & Results*. Section *Discussion* wraps up with ideas for further considerations in the monitor based color difference testing.

Technical Set-up

Within a controlled laboratory environment, following the CIE guidelines for viewing conditions [6], three high-end LCD

displays were used for display and evaluation. All displays had to be characterized and profiled carefully to achieve the best possible accuracy for displaying CIELAB colors on screen. The visual test was realized in HTML and PHP. Color display values were computed directly from CIELAB to RGB within the PHP program by taking each monitor's specific primary values and gamma into account. A database was created which comprised of the CIELAB color reference values.

Calibration

To attain precise color values on the RGB LCD displays, the first step was to calibrate each monitor to six given target settings (see Table 1) and store this information in the form of a profile. The calibration was carried out by utilizing Eizo's own calibration software *Color Navigator* for which it accesses the monitor-stored 10bit or 12bit look-up tables. Measurements for the calibration were carried out with a spectrophotometer, Xrite Eye One.

Setting	Target Value
Gamut	Monitor's native gamut size
WP	D50 0.34567 0.35850
Temp	5000 K
Bright.	120 cd/m^2
Gamma	2.2
Min	Possible minimum

Table 1. Aimed target settings

Three EIZO displays, one of the build CG 220 and two CG 241 were used, denoted by Monitor 1, 2 and 3. The characteristic of each monitor vary from one to the next caused by different model type and age. Due to these characteristics the achieved target values deviated slightly, see 2.

Setting	Monitor 1	Monitor 2	Monitor 3
Gamut	native	native	native
WP	0.3456 0.3585	0.3452 0.3586	0.3453 0.3585
Temp	5004 K	5020 K	5012 K
Bright.	114.8 cd/m^2	121.5 cd/m^2	120.4 cd/m^2
Gamma	2.2	2.2	2.2
Min	0.33 cd/m^2	0.27 cd/m^2	0.16 cd/m^2

Table 2. Achieved target settings

Color accuracy in web browser

We chose a web browser to display colors, since widespread and well defined standards are available as well as many applications exist to verify the behavior of the workflow in many different setups. Implementation of the desired content is sim-

ple (we choose PHP as a programming language) and platform-independent.

Displaying images and colors in a browser is based on the interaction of separate modules: The content, the application, the system, the monitor hardware and possibly other components. At each of these modules, a color conversion may or may not take place. Although many specifications are publicly available, it is hard to determine the precise process of color conversions within such a complex environment. The modules can not be analyzed separately as the resulting color can only be measured after the entire process is finished. Nonetheless it is possible to draw conclusions from the behavior of the whole process when exchanging the modules or changing the behavior of the modules one by one.

By extensive evaluation, it was possible to determine the **standard workflow** for the following modules:

- Colors: Plain CSS background-colors, denoted by integer numbers in the range [0, 255]
- Markup-Language: HTML5 + CSS 2.1
- Browser: Safari 4.0.5
- System: Mac OS X 10.6.2 (Snow Leopard)
- Cable: DVI-D single link
- Monitor: EIZO with integrated graphics card

The standard workflow, as it is intended by ICC color management, requires writing sRGB colors to the webpage. The browser parses the values and assumes sRGB values which are passed to the system. The system allows to set any RGB space as the device space for the connected monitor. The system will convert the sRGB colors into the device space, executed on the CPU or GPU. Finally, the monitor transforms these values into intensities which will result in the color visible on the screen, see Figure 1. The chosen browser considers colorspace definitions in images and passes this information to the system which manages all color conversions. When no colorspace is given, a browser should assume sRGB as the default RGB space [7] which has been verified for the chosen browser. Accordingly, plain CSS colors do not define a color space and therefore are assumed to be in sRGB color space. Since the values are interpreted as sRGB values within the standard workflow, only colors inside the sRGB gamut can be reproduced. Monitors, especially the Eizo models used here, nonetheless are capable of reproducing a larger gamut.

We therefore defined a **tweaked workflow** to take advantage of the larger gamut, see Figure 1. In our tweaked workflow, the monitor hardware contains the uploaded correction values in the same way as in the standard workflow. The system profile nonetheless is set to sRGB. Again, any RGB value written in CSS will be assumed by the browser to be in sRGB space which is passed to the system. As the system now has sRGB as the monitor profile, it will not alter the color values and the system will pass the values directly to the monitor without any conversion. Therefore, device-RGB colors can be written directly in the browser which gives us an additional degree of control and also allows to use the full gamut of the monitor. In return, it is in the responsibility of the implementation to compute the display values correctly.

Computing and displaying the color values

Using the calibration tool of the monitors, the RGB color space definitions shown in Table 2 as well as the three primary

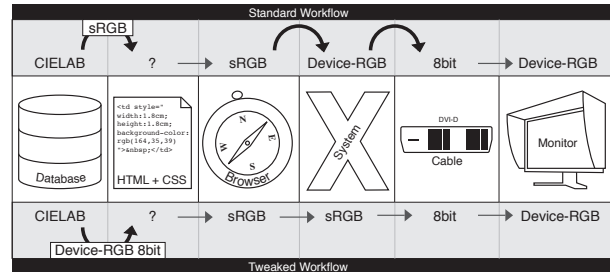


Figure 1. Standard and Tweaked Workflow. The question-mark indicates the missing color space definition of CSS colors which causes the Browser to assume sRGB. The rounded arrows indicate color conversions.

valences (x_r, y_r) , (x_g, y_g) and (x_b, y_b) are gathered. In the implementation, we convert between CIELAB and device RGB values according to standard methods and formulas found in [8] [9] [10]. All computations are done using floating point numbers.

The desired values are available as CIELAB values stored in a database. To display a patch filled with the desired color, the CIELAB color has to be converted to device-RGB space. In current CSS specifications [7], RGB values can be written as 3 integral numbers in the range of [0, 255] or as integral or floating point percentages in the range of [0%, 100%]. As the monitors are connected with a DVI single link cable, only 8 bit precision is possible (see [11]) and therefore, it is favorable to compute integral values in the range of [0, 255] directly in our implementation. Otherwise, somewhere in the color workflow, a quantization (rounding to the next integer value) will occur which can not be controlled.

The quantization of the color values introduces a small deviation of the desired color. As a further advantage of controlling the quantization within the implementation, this error can be computed by converting the quantized values back to CIELAB space. In table 3, measurement values are compared to the CIELAB values computed from the quantized RGB values instead of using the values available in the database. The quantized RGB color is displayed using the CSS `background-color` property applied to a table cell with a width and height of 18mm as specified in [5]. According to [7], a browser must assume a screen resolution of exactly 96 dpi. Using a private tool, it was possible to determine the true resolution of the screens to be about 94.2 dpi or 102.2 dpi respectively. The small resolution deviation is neglected.

Verification of Test Set-up

The colorimetric accuracy of reproduced color values on LCD displays serves as the verification for our developed test design and workflow. 168 reproduced color values were measured on each of the three monitors with the Minolta CS 1000 spectroradiometer. While each reproduced color value was measured by itself, the 168 values really describe 84 color difference pairs. We therefore establish the accuracy of reproduced absolute color values as well as the accuracy between aim- and reproduced color difference pairs. Establishing the accuracy of color differences is chosen in regard to the upcoming visual test which will investigate perceived color differences.

Regularly, also the white point of the monitor was measured to estimate and compensate for small drifts of the whole measuring system (monitor and CS1000) over time. The results are sum-

marized in Table 3. All three monitors showed equal performance thus we present results averaged for all three monitors. Absolute colors show less than 1 dE for the 50% percentile and maximum deviation smaller than 3. Significantly better is the performance for small color differences: if, on top, the quantization to 8 bit data is accounted for, the maximum deviation is even less.

	50%	90%	Max
Absolute Colors			
uncompensated Quantization	0.93	1.83	2.73
compensated Quantization	0.87	2.04	2.66
Small Colors Differences			
uncompensated Quantization	0.44	0.89	1.57
compensated Quantization	0.30	0.59	1.03

Table 3. Average ΔE deviation of color reproduction on the three monitors. Shown are 50% and 90% percentiles and maximum deviations for absolute color measurements and small color difference measurements (i.e. $< 10\Delta E_{76}^*$).

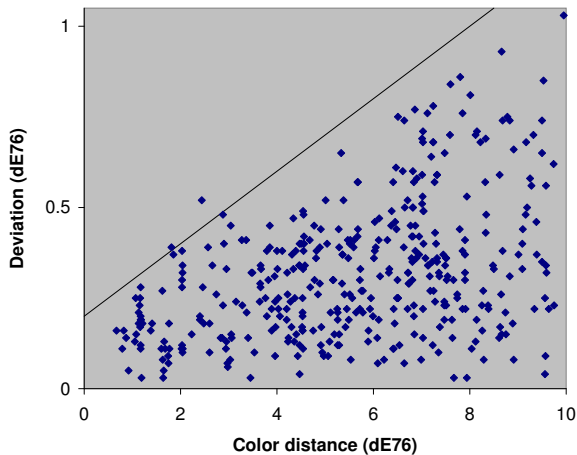


Figure 2. Deviation of measured color distances to aim color distances as a function of the color distance ΔE .

A closer analysis of the displayed color difference accuracy is shown in Fig. 2. For very small color differences (in the order of $dE = 1 - 3$) the accuracy is better than 0.5. The estimated upper limit variance between aim- and measured color difference pairs can be described by a linear 10% increase and an offset of about 0.2. The corresponding line is indicated in Fig 2.

Application: Fogra Color Differences Test

After verifying the digital LCD workflow we can now re-implement an existing print-based study by Fogra [5] on displays. To test the applicability of this LCD test design, a psycho-visual study that was priorly carried out with printed samples, was implemented. In the end, we compare observer results to determine the suitability of conducting a visual study within the newly established LCD based workflow.

Visual Experiment

Observers were asked to determine whether they perceive color difference pairs to be larger, smaller or equal against grey

anchor pairs. Color difference values are based on the Fogra test and therefore fit within the print gamut. These are derived as follows: for 14 color centers from the Fogra39 print condition, 28 variations, equally varying in chroma or hue angle, were computed. This results in color differences between $\Delta E_{76}^* = 0$ to $\Delta E_{76}^* = 15$ for each color. Color difference pairs were tested against three neutral grey anchor pairs with a nominal color difference of $\Delta L^* = 5, 3, \text{ and } 1$ to (CIELAB: 50, 0, 0).

Here, the method known as constant stimuli comparison was used to compare color difference pairs to an anchor pair. The test question 'Do you perceive the difference between the color pair as larger than the grey pair difference?' was answered by 'yes, color difference is larger' or 'no, smaller or equal' which forced a binary decision from the observer. Observers weren't restrained by a time limit when judging the color differences. The visual test set-up was realized by displaying the four test patches on a homogenous grey background. Each patch had the size of 1.8 cm x 1.8 cm. This equates to a visual field of 2° at a viewing distance of 60 cm. Color difference pairs were displayed on top and the grey anchor pairs beneath. The distance from color pair to anchor pair was also 1.8 cm, see Fig 3.

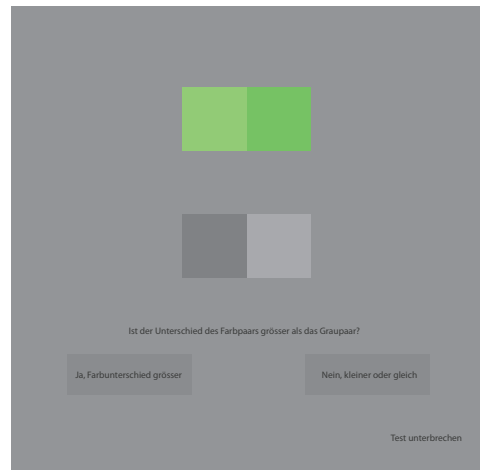


Figure 3. Monitor-based Color Difference Test, Illustration

For the comparison of our LCD based study with the Fogra study we matched test conditions as close as feasible except that color differences were judged on displays instead of printed samples. Each color difference pair was judged 10-20 times on average for the monitor test. The corresponding Fogra data consisted of 5-15 judgments.

About 80 observers took part in the LCD based study. Each observer had normal color vision, according to the Ishihara test. Most observers had previous experience with visual testing and were mainly experts. Answers were given as mouse-clicks and stored into a MySQL database.

Evaluation

Visual judgements of color differences from both studies were evaluated by means of comparing the data sets among themselves as well as comparing perceived distances to color distance measures, ΔE_{76}^* , ΔE_{94}^* and ΔE_{00}^* . Color difference formulas are used to quantify the accuracy and acceptable tolerance limits of a given color in the reproduction workflow. The per-

ceptual non-uniformities of the widely used, underlying CIELAB space which these formulas are based on prevent equal color difference discriminations [12] around each color center. Contrary to ΔE_{76}^* , newer color difference equations such as ΔE_{94}^* [13] and ΔE_{00}^* [14, 15] use location and direction dependent weights to compensate these effects. With the use of these formulas along with probit analysis we determine visual distances between the two data sets from print and monitor visual test.

In a first step the perceived color distance dE_v was determined for each of the 28 color difference of a color center using the available user data for the anchor pairs $k = 1..3$. For that step we used probit analysis [16, 17]. For each anchor pair k percentage values were computed:

$$p_{i,k} = \frac{f_{i,k} + \delta}{n_{i,k} + 2\delta} \quad (1)$$

where $f_{i,k}$ is the frequency that the difference of a color pair i was judged to be larger than that of the anchor pair k and $n_{i,k}$ is the total number of judgments of the color pair i for the anchor pair k ¹. The probit $z_{i,k}$ values were calculated with:

$$z_{i,k} = \Phi^{-1}(p_{i,k}) \quad (2)$$

where Φ is the cumulative distribution function of the standard normal distribution. It is assumed that the probit values $z_{i,k}$ have a linear dependency on the color distance of the anchor pairs. A weighted linear regression was used to estimate this relationship. The weights in the regression were chosen as the inverse of the estimated standard deviation of each probit value z_k . For the determination of the perceived color distance dE_v we used T75 (75% tolerance level). This level (and not T50) was chosen in order to compensate for the asymmetric questioning ('larger' against 'smaller or equal'). For each dE_v a corresponding estimated standard deviation σ_E was computed from the fiducial limits of the probit analysis.

Figure 4 shows a typical sample evaluation of 1 color pair to 3 anchor pairs (1 color difference of 1 color center). 25 % of the observers evaluated this given color difference as larger than ΔE_{76}^* of 3 and 75 % judged the difference as less than or equal to ΔE_{76}^* of 3. The curves of the upper and lower fiducial limits (5%) as well as the upper and lower standard deviation are shown. The determination of the perceived distance dE_v and its standard deviation σ_E is illustrated.

In order to get a quantitative measure how much the two data sets x and y differ we computed the following measures: The average difference

$$d_{1,2} = \frac{1}{n} \sum (dE_{v,1} - dE_{v,2}), \quad (3)$$

the deviation of the differences

$$s_{1,2} = \sqrt{\frac{1}{n-1} \sum (dE_{v,1} - dE_{v,2} - d_{1,2})^2}, \quad (4)$$

and an agreement factor

$$f_{1,2} = \frac{1}{n} \sum \frac{(dE_{v,1} - dE_{v,2})^2}{\sigma_{E,1}^2 + \sigma_{E,2}^2} \quad (5)$$

¹We introduced the bias correction δ in order to eliminate numerical problems for pairs of items, which have zero entries for the frequency. In this paper $\delta = 0.1$. For a discussion of different bias correction formulae see also [18] chapter 9.4.

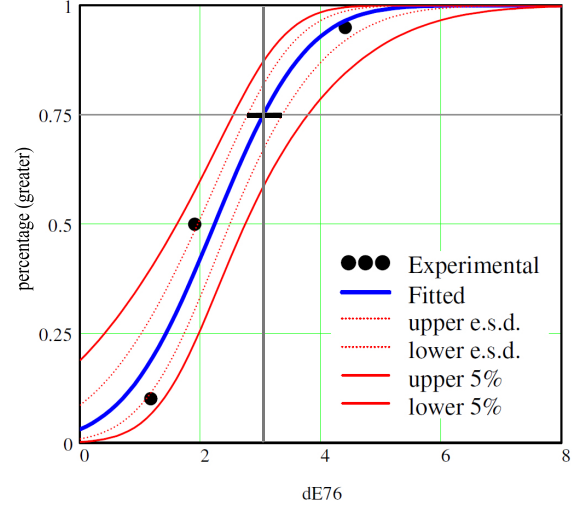


Figure 4. Probit analysis: Fit of a cumulated distribution function to the data of one color difference pair to three anchor pairs. The extracted perceived distance using T-75 level is $dE_v = 3.1$ with an estimated standard deviation of $\sigma_E = 0.3$.

where n is the number of data points compared, $dE_{v,1}$ and $\sigma_{E,1}$ refer to the Fogra test data and $dE_{v,2}$ and $\sigma_{E,2}$ to the monitor test data. If this factor $f_{1,2}$ is close to 1 the deviation is only due to statistical errors. A factor much larger than 1 indicates that there are significant differences between the two data sets. Note that this factor can be derived from a χ^2 -test if the χ^2 -sum is divided by the degrees of freedom. As shown in Table 4 the resulting agreement factor $f_{1,2}$ comparing our two test set was 1.31 which is considered quite small, but still systematically larger than unity.

The resulting perceived distances dE_v can also be compared with available color distance measures. Here we compare the data with the widely used formulae ΔE_{76}^* and ΔE_{94}^* which both are based on the CIELAB coordinates (L^* , a^* , b^* , C^* and H^*). They are defined as follows[13]:

$$\Delta E_{76}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \quad (6)$$

$$\Delta E_{94}^* = \sqrt{\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C^*}{k_C S_C}\right)^2 + \left(\frac{\Delta H^*}{k_H S_H}\right)^2} \quad (7)$$

with $S_L = 1$, $S_C = 1 + 0.045 C^*$, $S_H = 1 + 0.015 C^*$. k_L , k_C and k_H are parameters which by default are set to 1. The definition of ΔE_{00}^* is taken from Sharma et. al[14]. The implementation of this color difference formula was tested with the color difference test data given there.

For the comparison with known color distance measures ΔE_{xx}^* with a data set j (1 for monitor data, 2 for Fogra data set) the following measures were calculated: an average difference $d_{\Delta E_{xx},j}$, a deviation of the differences $s_{\Delta E_{xx},j}$ and an agreement factor $f_{\Delta E_{xx},j}$ similar to equations (3-5). The corresponding formula for the agreement factor is then

$$f_{\Delta E_{xx},j} = \frac{1}{n-m} \sum \frac{(dE_{v,j} - \Delta E_{xx})^2}{\sigma_{E,j}^2} \quad (8)$$

where n is the number of data points compared, m is the number of free parameters of the color distance measure being optimized using the data set.

Results

The same evaluation was made for both data sets, the monitor data set and the Fogra data set. Only dE_v values with the criteria of σ_E smaller than 2 (for both data sets) were used for the following comparisons. Out of the 392 (14 times 28) possible dE_v -values, 189 values fulfilled this criterion and were used in the further analysis. Most of the discarded values showed dE_v values substantially larger than the largest anchor pair difference. In

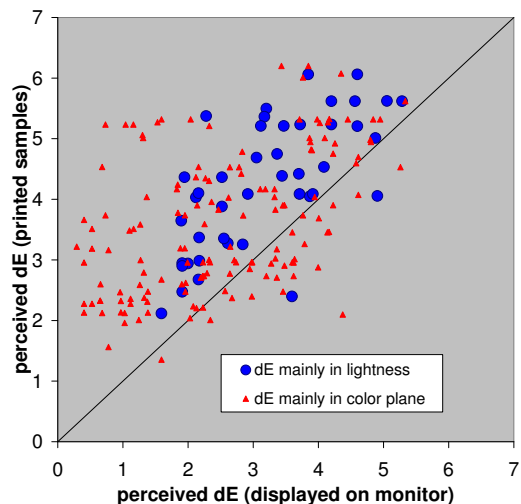


Figure 5. Comparison of perceived color distances from monitor data set with Fogra data set. The same color difference is perceived larger on print than on screen.

Figure 5 we have plotted the dE_v -values of both data sets against each other. If the tests were perceptually identical we would expect a linear relationship with unit slope (shown as a black line). For color differences mainly in the color plane (shown as red triangles) the correlation is less. In particular, small color differences seem to be perceived systematically smaller for the monitor data set as compared to the Fogra data set.

The results of the comparisons for the three color difference formula ΔE_{76}^* , ΔE_{94}^* and ΔE_{00}^* are summarized in Table 4. The correlation with the ΔE_{76}^* color difference formula is very low (high agreement factor) for both the monitor data set and Fogra data set. In general ΔE_{76}^* overestimates the visual color distances and the average deviation is quite large. The ΔE_{94}^* and in particular the ΔE_{00}^* color difference formula shows better results especially for the Fogra data set. In general visual distances were estimated to be larger in the Fogra experiment on printed paper compared to our monitor experiment. Best results were obtained by optimizing the available parameters k_L , k_C and k_H of the ΔE_{00}^* difference measure. The agreement factor between data set and said color distance measure was improved, especially the monitor data set benefited from the adjustment. Parameters k_C and k_H were forced to be equal in the optimization. The resulting agreement of 1.49 for both data sets is close to but larger than the values obtained for the comparison of the two data sets. Interestingly, the

	average diff. d	average dev. s	agreement factor f
print - monitor	1.1	1.16	1.3
ΔE_{76}^* - print	2.8	2.8	6.6
ΔE_{94}^* - print	0.1	1.18	1.8
ΔE_{00}^* - print	-0.2	1.11	1.6
ΔE_{00}^* - print $k_L=1.0, k_{C,H}=1.3$	-0.5	1.05	1.49
ΔE_{76}^* - monitor	3.8	2.7	8.9
ΔE_{94}^* - monitor	1.2	1.21	3.1
ΔE_{00}^* - monitor	0.9	1.19	2.5
ΔE_{00}^* - monitor $k_L=1.2, k_{C,H}=1.7$	-0.1	0.98	1.55

Table 4. Comparison of the visual distances from monitor and print data sets among each other and with with the color difference formulae ΔE_{76}^* , ΔE_{94}^* and ΔE_{00}^* . Shown are average difference d , average deviation s and agreement factor f . If the agreement factor is close to 1.0 the data sets agree within the statistics.

optimized parameters for k_C and k_L are significantly larger for the monitor data set compared to the Fogra data set. This difference is in line with the systematical deviations in Fig. 5 of color differences mainly in the color plane: a certain color difference is perceived somewhat larger on print than on a monitor. A similar result was observed in [1] where color differences of real world images had a lower perceived tolerance on print than when displayed on a monitor. Contrary to the optimized parameters for the color patches used in this experiment, [2] showed that real world images need a larger lightness weight than chroma and hue weights. However, the performance for ΔE_{00}^* for monitor color patches benefited from the larger weight in chroma and hue.

Discussion

The results of the Verification section show, that from a technical point of view, LCD displays are a good alternative to tests using printed samples. The accuracy for displaying color differences at least meets the accuracy using printed samples, provided that the color processing and display workflow is under control of the test program. In particular, the compensation of the quantization in the 8 bit bottleneck helped to realize an excellent color accuracy. The reproduction of absolute color values on LCD displays was achieved with $\Delta E_{76}^* 0.87$ and $\Delta E_{76}^* 2.04$ average deviations for the 50th and 90th percentile respectively. While this could be considered a noticeable difference compared to the printed color values, the cross media translation of color values was achieved more accurate for color differences. These differences were reproduced with $\Delta E_{76}^* 0.30$ and $\Delta E_{76}^* 0.59$ deviations for the 50th and 90th percentile respectively. This accurate reproduction provides a technical setting that allows for the application of a color difference evaluation on LCD displays.

The comparison of the two data sets revealed, that there are still differences between the two data sets. One possible reason could be the different white adaption of the human eye in the two test settings: the test layout of the monitor test provided the eye with little indication of what the white point of the monitor was. If the eye had adapted to a darker white, all color distances would

appear somewhat larger. However the fact that also the distance of the anchor pairs appear larger too, the effect of white adaptation is at least partly compensated. Another possible reason is the arrangement of the color samples in the visual test. Since samples are positioned right next to one another, the two test designs produced a) for the physically printed samples a sharp edge with shadow between samples, due to bulging of the paper whereas b) the samples on displays didn't exhibit such an edge. This might have had an influence on observers' judgements on color differences as well. In the process of our study we recognized several points that could be improved in an extended study on color difference:

- Analysis of observer data could benefit from giving observers three answer choices instead of two: 1) color pair difference is larger, 2) color pair difference is smaller AND 3) color differences of both pairs are equal. In this work 2) and 3) were fused into one answer and it is not entirely clear in the later analysis whether observers really couldn't distinguish the difference or if, in fact, the color pair difference was smaller.
- Instead of using a static test procedure (each observer was given the same comparisons), the utilization of an algorithm based test provides the capacity to develop and use a smart, adaptive algorithm. Observers would be given adaptive choices, depending on their former answers. This essentially allows to expedite the process by learning which differences are easily detected and to focus on testing the small color differences quicker.
- Our test set up displayed each color center with according variations in a consecutive manner. A few observers reported slight after effects after a series of vivid colors. We consider a better randomized viewing cycle for further color difference tests on monitors.

Now that we established a technical accurate workflow for visual testing on LCD displays, we are able to investigate further questions regarding human perception.

Conclusions

A test set-up for evaluating color differences on a LCD display has been developed. The encountered challenges, such as handling the visualization of color values throughout the workflow's employed color spaces, evaluating possible color conversions from different involved modules and finally determining the workflow which best solves the visualization of the desired color values are described here.

We have shown, that LCD displays are a good alternative to tests using surface color samples. An important prerequisite is that the color processing and display workflow is under control. The flexibility of testing using a monitor allows more sophisticated test designs in particular adaptive questioning.

The further investigation should focus on color appearance parameters such as illuminated samples vs. self-luminous samples, ambient light, background and surround colors. The developed workflow for LCD displays described in this paper allows to further research such questions.

Acknowledgments

We gratefully acknowledge, that FOGRA (Graphic Technology Research Association) gave us access to the raw data of their user study.

References

- [1] M. Stokes. Colourimetric tolerance of digital images. *MSc Thesis, RIT, University of Rochester, New York, USA*, 1991.
- [2] T. Song and R. Luo. Testing color difference formulae on complex images using a CRT monitor. *Color Imaging Conference*, pages 44–48, 2000.
- [3] M. Melgosa, A. El Moraghi M. M. Perez, and E. Hita. Color Discrimination Results from a CRT Device: Influence of Luminance. *Color Research and Application*, 24:38–44, 1999.
- [4] E. D. Montag and R. S. Berns. Visual Determination of Hue Suprathreshold Color-Difference Tolerances Using CRT-Generated Stimuli. *Color Research and Application*, 24:164–176, 1999.
- [5] A. Kraushaar, F. Gessner, C. Bickeböller, and P. Karp. Untersuchung moderner Farbabstandsformeln. *Fogra Forschungsbericht Nr. 60.054*, 2008.
- [6] Central Bureau of the CIE, Vienna. *CIE Publication 156: Guidelines for the Evaluation of Gamut Mapping Algorithms*, 2004.
- [7] World Wide Web Consortium (W3C). *CSS 2.1 available on http://www.w3.org/TR/CSS21/*, September 2009.
- [8] M.D. Fairchild. *Color Appearance Models*. Wiley, February 2005.
- [9] R. W. G. Hunt. *Measuring Color*. Fountain Press, 3 edition, 1998.
- [10] G. Sharma and H. J. Trussell. Digital color imaging. *IEEE Transactions on Image Processing*, 6(7):901–932, 1997.
- [11] Digital Display Working Group (DDWG). *Digital Visual Interface DVI available on http://www.ddwg.org/lib/dvi_10.pdf*, April 1999.
- [12] K. Witt. Geometric relations between scales of small color differences. *Color Research & Applications*, 24:78–92, 1999.
- [13] CIE. *Technical Report; Industrial Colour-Difference Evaluation*, 1995. 116-1995.
- [14] G. Sharma, W. Wu, and E. N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Applications*, 30:21–30, 2005.
- [15] M. R. Luo, G. Cui, and B. Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research and Applications*, 26:340–350, 2001.
- [16] R. S. Berns, D. H. Alman, L. Reniff, G. D. Snyder, and M. R. Balonon-Rosen. Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research & Applications*, 16:297–316, 1991.
- [17] D. J. Finney. *Probit Analysis*. Cambridge University Press, 1952.
- [18] P. G. Engeldrum. *Psychometric Scaling, A Toolkit for Imaging Systems Development*. Imcotek Press, Winchester MA, USA, 2000.

Author Biography

Iris Sprow received her BSc in Imaging & Photographic Technology from the Rochester Institute of Technology in 2005 and finished the MSc graduate program in Digital Colour Imaging at the London College of Communication. In 2005 she joined the Media Technology group at EMPA Dübendorf where her work is focused on subjective image evaluation.