Draft Version - originally published in: Conference Proceeding of 31st International Conference on Information Technology Interfaces (ITI 2009) D.28 (2009), S. 323 - 328; 2009

The Modeling of Harmonious Color Combinations for improved Usability and UX

Christian Stickel
Social Learning
Graz Univ. of Technology
Steyrergasse 30/I,
A-8010 Graz
stickel@tugraz.at

Karin Maier
Social Learning
Graz Univ. of Technology
Steyrergasse 30/I,
A-8010 Graz
karin.maier@student.TUGraz.at

Martin Ebner
Social Learning
Graz Univ. of Technology
Steyrergasse 30/I,
A-8010 Graz
martin.ebner@tugraz.at

Andreas Holzinger
Research Unit HCI4MED,
Institute of Medical Informatics,
Statistics and Documentation
Auenbruggerplatz 2/V, A-8036 Graz
andreas.holzinger@meduni-graz.at

Abstract. *This study compares three different models for the calculation and prediction of harmonious color combinations. Therefore a dataset of user rated color combinations was taken from a large online database. The user rating was compared to the outcome of the three models on this dataset in order to test the performance of the models. The first model based on the idea that color combinations are more pleasing the greater their difference in brightness. The second model is a slightly modified version of Ou & Lou (2006) using chromatic difference, lightness sum, lightness difference and hue effect. The last model was invented by us and is based on an experiment of Polzella & Montgomery (1993). From the outcome of their experiment we generated a lookup table for single color rating. This rating is then used in a formula, which is able to evaluate the color harmony for color combinations up to five colors. This model also performed best in the overall comparison between the three color harmony models.*

Keywords. Color Combination, Harmony, Pleasantness, Model, User Experience, Usability

1. Introduction

Colors can be used according to their physiological and psychological effects in order to support human information processing and thus enhance the usability of software [5]. They can influence effectiveness and efficiency directly while dimensions like satisfaction and pleasantness will have an indirect effect on these aspects [18][2]. Colors can guide the attention of the end-user and at the same time create an atmosphere that affects emotions. The careful selection of appropriate colors can optimize desired interactions and the effectiveness of a task. As colors have symbolic meanings and emotional effects they affect both Usability and User Experience (UX). Guidelines for the use of color in user interface design can be found in EN ISO 9241-8.2 [8]. As interfaces have to consist of at least two different colors it is of interest to optimize color combinations, resp. color themes in order to be pleasant and thereby supporting the end-users cognitive performance [6].

This study is about modeling pleasant, resp. “Harmonious color combinations”. These can be defined as “Two or more colors seen in neighboring areas producing a pleasing effect, are said to produce a color harmony” [9]. This harmony is crucial for usable interfaces, especially if one wants to regard the dimensions of UX, which represent a much wider point of view than the function oriented Usability, looking at individual interactions, thoughts, feelings and perceptions [23].

Consequently when combining colors for an interface both aspects have to be considered. Using Harmonious color combinations can do this. The question occurs, how harmonious color combinations can be defined or even modeled. Therefore several researches have been done, e.g. [20][4][22][19], however these studies focus on the relationship between two colors, respectively foreground and background color. Other studies on mathematical models for color combination cover the influence on legibility and readability [11][3][25]. As interfaces today consist of more than two colors, we chose a dataset with combinations consisting of five colors for this study and tested the performance of three different models for the prediction of Harmonious Color Combinations.

2. Theoretical Background

This chapter covers some important aspects of color perception and modeling color harmony. The theoretical background will be helpful to understand the approaches of the applied models.

2.1. Color perception

The base for every visual experience is the eye, which consists roughly of the lens, the retina and the optical nerve. The lens focuses the incoming light on the retina, which contains two different types of photo receptors. These are rods for monochromatic vision at night and cones, which are color sensitive, require a higher level of light intensity in order to be activated and thus are used at daytime. The cones contain three types of photo pigments to detect blue, green and red at different wavelength. As the curves for excitation of the receptors are overlapping to some extend, light at a single wavelength will more or less activate all three types. The sensitivity of the photoreceptors depends on the overall light level of the surrounding and requires some time for adjustment. There is also a minimum threshold for the receptors to be activated. This minimum is varying per wavelength with the highest sensitivity in the center of the spectrum. Consequently blue and red colors must have a higher intensity than green or yellow colors in order to be perceived.

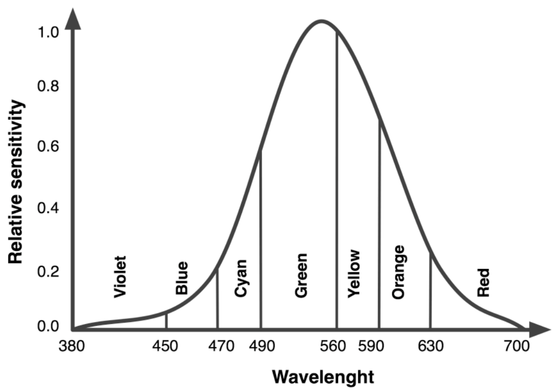

Figure 1. Relative sensitivity of the eye as a function of wavelength (CIE Photopic Luminous Efficiency function)

As mentioned the lens focuses the incoming light on the retina whereby different wavelengths of light are associated with different focal lengths. For all pure hues the lens must change its shape accordingly in order to focus the light correctly, thereby longer wavelengths have a longer focal length, i.e., red accords the longest focal length and blue the shortest. So far the lens curvature must change according to different wavelength. The process of changing is called accommodation. Red spectrum light requires the greatest and blue the least curvature. An effect occurring from this is that the mixture of pure red and blue will cause constant accommodation and thus tiring the eyes. Another effect related to this is chromostereopsis, which describes that some pure colors located at the same distance appear to be at different distances, e.g. in a blue / red pattern, the red color will appear closer. With increasing age the lens absorbs more of the shorter wavelength, consequently people become more sensitive to longer wavelength light e.g. yellow and orange while the sensitivity for shorter wavelength light like blue decreases. The fluid inside the eyeball also absorbs a part of the incoming light. This absorption also increases with age, which so far appears as brightness level decreases [13][14][16]. Consequently for old users one would design interfaces that prefer warm and bright colors [7].

Colors guide the attention and communicate structure and order if used appropriately. Thereby the less frequently a color is used, the better it will grab the user's attention. The combination of too many colors makes the display look cluttered, confuses the end-user, makes tasks more complex, increases errors and reduces productivity [24][17][10]. A reduction to more or less five different colors is suggested [21] and comprehensible regarding the paradigm of cognitive psychology. Thereby the perception of a color is regarded as an activation of a cognitive unit [12]. The conscious mind can be regarded as the pattern of activated units. However the capacity is limited to more or less five activated units, whereby the focus of attention will determine the strength of activation. So the reduction of the number of important items, resp. colors if used as coding for information as well as the redundant and consistent use of color will support the cognitive performance of the end-user. A further optimization of cognitive performance can be achieved taking into account associations between color and distinctive emotions as well as user expectations for certain color coding. Colors are so omnipresent in everyday life that it seems impossible having them not connected to multiple cognitive units, so with the perception of a color a whole network of these units, resp. cognitions, is activated. Red for example will mostly grab the attention as it evokes cognitions like blood, hot, danger, excitement, warning, stop or error, whereby green may produce cognitions like Nature, calmness, hope, good or everything is alright [15]. The specific associations shift slightly on the individual level however there are certain associations for color coding which are developed as a part of culture or within certain contexts.

2.2. Modeling Color Harmony

There have been several approaches to generate models of color harmony, e.g. Helson & Lansford [4] conducted a study wherein they measured the effects of spectral energy distributions of sources and colors of backgrounds on the pleasantness of object colors.

They found that the principle factor determining the pleasantness of color combinations was brightness contrast, whereby the best rated color combinations involved large brightness differences between color and background. The worst rated combinations on the other hand showed little or no difference. This conclusion was later found not to be generally applicable [20]. Another study [25] focused also on the combination of foreground and background color, but in the context of legibility, which describes the visual clarity of text - in short legibility is how well one can see the letters [1]. They found a luminance contrast of 3:1 to be ideal and propose this measure as a designing rule for selecting legible color combinations.

A more comprehensive color harmony model is described by Ou & Lou [19]. It combines four different principles to determine harmonious color combinations, namely chromatic difference, lightness sum, lightness difference and the blue hue effect. Based on the prediction performance of the model they propose four color-harmony principles for two-color combinations. The first principle says that color pairs with equal hues are more harmonious. The second states that lighter component colors correlate with higher color harmony scores. The third principle encourages the use of a moderate lightness difference between the components and the fourth principle says that color pairs consisting of blues tend to harmonize.

Polzella & Montgomery [20] found that the more harmonious combinations involved colors of similar brightness or similar hue, which matches afore mentioned principles. The parameter of saturation seems to have no effect on color harmony. Therefore they determined the pleasantness of 12 base colors and then let the subjects rate combinations of these colors. Surprisingly the combination of pleasant colors did not necessarily result in a harmonious color combination.

3. Methods and Materials

This chapter covers the used dataset, models and results. It describes how these models were modified and applied to the new context of five color combinations and how they performed on the chosen dataset.

3.1. Dataset

The dataset for the present experiment consisted of 20 different color combinations with 5 colors each. These combinations were taken from the Adobe® Kuler™ (http://kuler.adobe.com) website, which is all about creating color combinations online, resp. color themes for user interfaces. Every created theme can be rated, tagged, shared and commented by a huge community and so far there's a pool of nearly 150.000 combinations whereof 7000 are rated. We took the sample of the ten best rated and ten worst rated color combinations. For all single colors of the combinations the HSV and Grayscale values were calculated.

3.2. Methods

Our research question was how the calculated harmony of color combinations would match with the user rating from the Adobe® Kuler™ community. Are afore mentioned principles (see chapter 2.2) able to explain or even reconstruct the sorting of the chosen dataset? Three different models for calculation have been applied to the dataset and the results have been compared. It was expected for all models to find a significant difference between the best and worst rated color combinations.

3.1.1. Grayscale Variance Intersection

This model follows the assumption of Helson & Lansford [4] that large differences in brightness are more pleasant, resp. result in a harmonious color combination. First all colors were converted to a 1-100 % grayscale format, in order to compare the grayscale variance. Then a z-score test has been done on the variances, in order to generate a standardized dataset of both best and worst rated combinations. The intersection and standard error of the variances was then calculated and visualized.

3.1.2. Color Harmony Intersection

Ou & Lou [19] propose a very comprehensive model, which was applied to the dataset of the present study. The metric of color harmony is calculated as the sum of chromatic difference, lightness sum, lightness difference and hue effect. Chromatic difference is determined by hue- and chroma difference. Lightness Sum results from adding up lightness of components. Lightness difference is specified as absolute difference of lightness. The Hue Effect indicates the mean color-harmony score for a specific hue angle combined with all other hue angles.

As the original model covers only the combination of two colors this model was slightly modified to handle the five color combinations of the chosen dataset. The original formula is intended for pairs of colors, so we chose the base color of every color combination and the lightest, resp. darkest of the remaining four colors. The base colors of the combinations were defined by their creators, while the lightest and darkest color was chosen by us using the grayscale value. Then we calculated the intersection of the resulting color harmony of the two pairs.

3.1.3. Pleasantness Intersection

The pleasantness intersection model determines the pleasantness of every color of a combination first, then sums up the outcome and divides by the number of colors to get the color harmony (see (1)). So this model defines the combined pleasantness as color harmony value. Polzella & Montgomery [20] had their subjects rating 12 base colors in order to calculate a pleasantness scale for these colors. We used this scale to build a lookup table (LT) for our pleasantness intersection model. Therefore the original colors have been sorted by their pleasantness value and then got a new rating with one point for the worst color and 12 points for the most pleasing color. White, black and pure grey values get zero points. See Table 1 for detailed information on the colors and scoring. This score (F) was then used in the following formula (see (2)) that determined the pleasantness (P) of each color of a combination. In order to get the score (F) for a certain color the distance to the nearest color from the lookup table was calculated (see (3) & (4)).

(1) $\text{Color Harmony} = \sum P_i$ with i=1 to 5

Where

P: pleasantness of single color

(2) $P = F - ((m/n)*10) - \Delta x$

Where

F: nearest lookup table color score, with

(3) $F = \text{Min}(F1, \ldots, F12)$, with

(4) $F_i = (H_{org} - H_i) + (S_{org} - S_i) + (V_{org} - V_i)$

Where

H_{org} : Hue of color to compare

H_i : Hue of each color from LT

S_{org} : Sat. of color to compare

S_i : Sat. of each color from LT

V_{org} : Value of color to compare

V_i : Value of each color from LT

m : # of colors with same score in the set

n : # of colors in the set

$\Delta x = -1$ if $G >= 50\%$

1 if $G < 50\%$

Where

G : grayscale value

Table 1. Lookup Table for Score (F)

	Munsell			HSV Model			Score
	H	V	C	H	S	V	(F)
Lt. Cyan	5B	7	8	193	77	90	1
Blue	5Pb	4	10	210	82	68	2
Lt. Blue	5Pb	6	10	210	65	90	3
Magenta	5P	4	10	272	48	62	4
Lt. Magenta	5P	5	10	273	43	72	5
Cyan	5B	6	6	195	63	75	6
Lt. Green	5G	7	10	161	75	78	7
Red	5R	4	10	356	65	65	8
Lt. Red	5R	5	12	354	64	82	9
Yellow	5Y	8	10	46	72	93	10
Green	5G	6	10	161	84	67	11
Brown	5Y	5	5	41	65	58	12

So for example a dark red color gets less points than a bright red, while a pink gets less points because there is more magenta in it, which is rated lower.

3.3 Results

The comparison of the intersection of variances of the different colors of the combinations, shown in figure 2, revealed that the color combinations in the best-rated group have a higher variance than the combinations in the worst rated group. Using a z-Test and a 2nd order polynomial trend graph revealed the same. However, the standard error is very high and so far the grayscale variance model is uncertain for predictions.

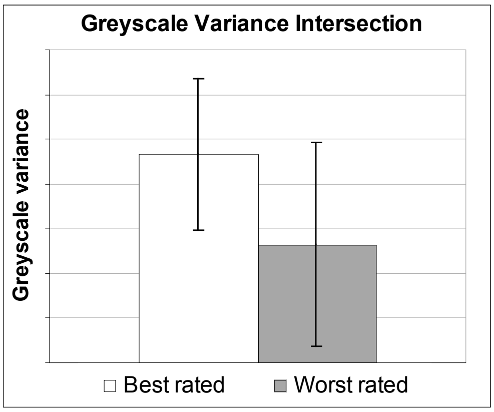

Figure 2. best rated have higher variance

The application of the color harmony model in this context didn't show any significant differences between the two groups. Neither the single comparison of base color vs. lightest and darkest nor the comparison of the base color vs. the average of lightest and darkest did show anything significant. The variance of both groups was equal and the correlation was low. The standard error is very high as can be seen in figure 3.

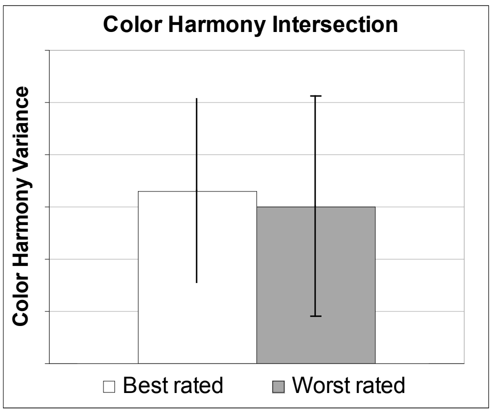

Figure 3. Color Harmony of best vs. worst

The calculation of the intersection of the pleasantness of the color combinations revealed a significant difference between the two groups as shown in figure 4. The standard error is low and so this model would be the preferred one to predict the harmony of a color combination with five colors.

Figure 4. Pleasantness of best vs. worst

4. Discussion

Three different models for the calculation of harmonious color combinations have been applied to a dataset of user rated color combinations. None of the models was able to reconstruct the original sorting of the dataset, however all distinguished more or less between the two groups of best rated and worst rated color combinations.

The grayscale variance model performed well for the two groups. Nevertheless the high standard error shows an uncertainty, which suggests that the examination of grayscales is too simple to explain the preferences of real end-users. So far the focus on brightness difference is not generalizable but may be of use in combination with other factors.

The color harmony intersection model provides a very detailed measure of different factors of harmony; however it fails to perform well with five colors, resp. the way we used it. This might be a problem of the formula, which considers the order of the compared colors to be crucial. Further there were only 3 of 25 possible calculations for every color combination done, so this could be still considered as topic for further experiments.

The pleasantness intersection model was able to distinguish significantly between the two groups. It needs some refinement for handling gray colors and testing with a larger dataset. Nevertheless it performs wells and seems to be able to predict harmonious combinations of multiple color combinations. Further experiments with this model could be conducted by varying the number of colors of the combinations.

5. References

[1] Bachfischer, G. (2005). Legibility and readability - a review of literature and research to understand issues referring to typography on screens and device displays. Tech. Rep. ID-WoP.tech.report.05.01, University of Technology Sydney, October.

[2] Christ, R. E. (1975). Review and analysis of color coding research. Human Factors, 17, 542-570.

[3] Gradisar, M.; Humar, I., Turk, T. (2007). The Legibility of Colored Web Page Texts. Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on Volume , Issue , 25-28 June 2007 Page(s): 233-238.

[4] H. Helson and T. Lansford (1970). "The Role of Spectral Energy of Source and Background Color in the Pleasantness of Object Colors," Appl. Opt. 9, 1513-1562

[5] Holzinger, A. & Errath, M. (2007) Mobile Computer Web-Application Design in Medicine: Research Based Guidelines. Springer Universal Access in Information Society International Journal. 6, 1, 31-41.

[6] Holzinger, A., Kickmeier-Rust, M., Wassertheurer, S., Hessinger, M. (2009) Learning Performance with Interactive Simulations in Medical education: Lessons learned from results of learning complex physiological models with the HAEMOdynamics SIMulator. Computers & Education, 52, 1, 292-301

[7] Holzinger, A., Searle, G., Kleinberger, T., Seffah, A., Javahery, H. (2008) Investigating Usability Metrics for the Design and Development of Applications for the Elderly, Lecture Notes in Computer Science (LNCS 5105), Berlin, Heidelberg, New York: Springer, 98–105

[8] ISO/DIS 9241-8.2 (1995). Ergonomic requirements for office work with visual display terminals - Part 8: Requirements for displayed colours (proposed Draft Standard). Genf:International Organization for Standardization.

[9] Judd DB, Wyszecki G. (1975). Color in Business, Science and Industry, Third Edition, John Wiley & Sons, New York

[10] Keister, R. L. (1983). Human factors guidelines for the content, layout, and format of CRT displays in interactive applications systems; for the use of help screens and online documentation (HFP 83-12, version no. 4). Dayton, Ohio: NCR Corporation.

[11] Kiritani Yoshie, Shirai Shogo (2003). Effects of background colors on user's experience in reading websites. In Aoki, H (Ed.), 6th Asian design international conference. Journal of the asian design international conference. Tsukuba: Institute of Art and Design, Univ. of Tsukuba. pp. 1-11.

[12] Martindale, C. (1981). Cognition and consciousness. Homewood: Dorsey Press.

[13] Murch, G. M. (1983). The effective use of color: Physiological principles. Tekniques, 7(4), 13-16.

[14] Murch, G. M. (1984a). The effective use of color: Perceptual principles. Tekniques, 8(1), 4-9.

[15] Murch, G. M. (1984b). The effective use of color: Cognitive principles. Tekniques, 8(2), 25-31.

[16] Murch, G. M. (1984c). Physiological principles for the effective use of color. IEEE Computer Graphics and Applications, 4(11), 49-54.

[17] Narborough-Hall, C. S. (1985). Recommendations for applying colour coding to air traffic control displays. Displays, July, 131-137.

[18] Norman, D. A. (2004). Emotional Design : Why we love or hate everyday things, New York: Basic Books.

[19] Ou, L. and Luo, M. R. (2006). A colour harmony model for two-colour combinations, Color Research and Application 31, 191-204.

[20] Polzella, D. J. & Montgomery, D. A. (1993). Dimensions of color harmony. Bulletin of the Psychonomic Society, 31 (5), 423-425.

[21] Shneiderman, B. (1987). Designing the user interface: Strategies for effective human-computerinteraction (1st edition). Reading, MA: Addison-Wesley.

[22] Smith, S. L (1963). Color Coding and Visual Separability in Information Displays. JI. of Applied Psych., 47, 358-364.

[23] Tullis, T.S. and Albert, B. (2008). Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Burlington, MA, USA, Morgan Kaufmann Publishers

[24] Williges, B. H., & Williges, R. C. (1982). User considerations in computer-based information systems (CSIE-81-2). Blacksburg, Virginia: Virginia Polytechnic Institute and State University, Industrial Engineering and Operations Research.

[25] Zuffi S., Brambilla C., Beretta G., Scala P. (2007). Human Computer Interaction: Legibility and Contrast. ICIAP 2007: 241-246