Eurographics Conference on Visualization (EuroVis) (2014)
N. Elmqvist, M. Hlawitschka, and J. Kennedy (Editors)
Short Papers

Studying Color Blending Perception for Data Visualization

S. Gama[1] and D. Gonçalves[1]

[1]INESC-ID & Instituto Superior Técnico, Universidade de Lisboa

Abstract

Visualization is a powerful way to convey data, showing a great potential for joining and interrelating different data items. Nevertheless, when dealing with large amounts of data, visually merging different classes of information poses several challenges. Color, however, due to its effectiveness for labeling and categorizing information, may be a solution to this shortcoming. Merging items with different colors may suggest mixing their original colors. This approach, while generating an immediately perceivable way to represent merged items, keeps context through the association of the resulting color to its original provenience. We studied to which extent color blending provides users with the means to understand the provenience of data items by conducting a user study with 73 subjects using CIE-LCh blending to ascertain (i) to which extent people are able to, given a particular color, understand its provenience, and (ii) the color model in which to perform color blending so that users find blending intuitive. Results showed that people have difficulties in understanding blending of colors that are farther apart in the color wheel and indicated that the CMYK model may show promise for representing blended colors.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Computer Graphics]: Information interfaces and presentation—User Interfaces

1. Introduction

Besides providing a powerful way to convey information, visualization alleviates cognitive load associated with data interpretation. In fact, a visualization technique is highly effective if it both supports visual queries and extends memory. When, when it comes to coding information, using color to display data categories is usually the best choice. This suggests the representation of complex data entities as visual objects, where colors naturally code their attributes [War12].

It is often desirable to visually represent data entities which verify more than one property. It is thus relevant visualize interrelated data items. That may be done by coding different data properties using color and representing entities that verify more than one property in a color which consists of the blending of the original properties' colors. In fact, color blending has been studied as a visualization technique. Nevertheless, some aspects remain to be studied that have not been particularly addressed by current research. One of such aspects is to which extent people are able to, given a specific color which results of the blending of two other colors, understand the blended color's provenience. Furthermore, several algorithms have been created, either relying on RGB or CIE-LAB color models, not having particularly focused on the influence of the color model over human perception regarding color blending. We have studied both of these aspects, which we will present on this paper.

We investigated to which extent color blending provides users with the means to understand the provenience of data items by conducting a user study with 73 subjects to ascertain (i) to which extent people understand the original colors which mixed into a particular given color and (ii) the color model in which to perform color blending so that users find it intuitive. Results showed that people have difficulties in understanding blending of colors that are farther apart in the CIE-LCh color wheel. Additionally, results indicated that there are no statistically significant differences on user preferences on blending between CIE-LCh and CMYK models. This analysis indicates that, since CMYK blending is potentially closer to people's idea of color mixing, the CMYK model may show promise for representing blended colors.

In section 2 we introduce theoretical principles underlaying our study. We then describe and detail the formulation of our study and its protocol. In section 3 our user study is presented and results are analyzed and discussed, drawing the guidelines this study has provided for further work on color blending studies for information visualization.

2. Background

Color is a sensation produced in the brain [CC04] that, if related to measurable phenomena, allows digital representation. The human eye contains three types of cones sensitive to light with different wavelengths, which resulted in the trichromatic color theory [You02], stating that any color can be specified by the weights of red, green and blue color components. Hardware-oriented color models are RGB (red, green, blue) and color-printing devices are usually CMY(K) (cyan, magenta, yellow, (key)). However, these models fail to provide a color perception description [FvDFH90], unlike the HSV (hue, saturation, value). The latter, however, fails to provide perceptual uniformity similar to human vision. The CIE-LCh model overcomes this limitation by assigning a perceptually uniform scale to lightness [Poy96].

If effectively represented, color has great potential. One interesting technique for multi-variate visualization is color blending, where each variable is assigned a different color and the color of the resulting data is computed as the weighted sum of the original colors. Various authors have studied different color blending techniques and a number of researchers have also considered color blending for visualization and compared it to other techniques. For instance, Chuang et al. [CWM09] studied a perception-guided compositing operator for color in which specific attention is paid to the original colors' hue preservation. As a consequence, the resulting colors are suited for labeling, improving visualization methods, particularly useful for volume visualization. Recently, Liang et al. [LQY*13] proposed an interpolation method based on green components and signal correlation. They rely on the RGB model and interpolate RGB color components using bilinear interpolation.

Concerning color blending for visualization, interesting research has been performed. Gossett and Baoquan [GB04] aim at improving visualization using color to convey data properties. A subtractive color space has been adopted which uses red, yellow and blue as primary colors. Noise patterns are procedurally generated to create subregions of easily identifiable colors within a mixed region as a complement to color blending. On a different note, Hagh-Shenas et al. [HSIHK06] studied information-carrying capacities of color blending and color weaving to encode multivariate information in map-reading. Livingston and Decker [LDA13] have also studied color blending, among other techniques, to represent trends among data layers on a demographic survey. Even though color blending yielded excellent response times, accuracy was not as promising. Nevertheless, even though CIE-Lab is a perceptual-based color-space, linear interpolations within it do not maintain saturation, even though brightness is uniformly blended. Contrarily, in CIE-LCh, saturation and brightness are perceptually uniform [FvDFH90]. As a result, performing color blending in this model may improve color blending outcome. Additionally, previous work, including the aforementioned studies, has focused on recognizing data properties through color in a particular context. Although relevant in such particular contexts, previous research has not specifically studied to which extent users are capable of perceiving data properties of a particular data item, created by the merging of two other data items with particular colors associated to their properties. We intend to do so, by ascertaining how users are capable of choosing the original colors that were mixed into a particular blended color, out of a limited set of original colors.

3. Studying color blending

We took advantage of the potential of color to depict visual information and considered the problem of merging different sets of data. The main objectives of our study are: (i) to ascertain to which extent users are able to understand the provenience of a given color resulting from the blending of two other colors and (ii) to find out the best color model to perform color blending. To do so, we performed a user study. In the following sections we describe our study in detail.

3.1. Designing a color study

Despite color coding being excellent to display category information, only between about five and ten codes [MG80, Her20] can be rapidly perceived. Taking the opponent process theory into account [Her20], we chose as main colors: (R: Red; G: Green; B: Blue; Y: Yellow). Since the CIE-LCh model represents a perceptual uniform color space [MG80], we performed color blending using this model. The first set of blended colors (set A) consists of the pairwise combinations of the main colors: RG, RB, RY, GY and BY. The second set (set B) consists of the triadic combinations of the main colors: RRG, RGG, RBG, RYG, RRB, RBB, RYB, RYR, RYY, GBG, GBB, GYB, GYG, GYY, BYB and BYY.

Our study started with a questionnaire in which users were asked their age range, gender, education, nationality and country of residence. On the second part, we performed a validated simplified 6-plate Ishihara test [DAK92] for color blindness detection. On the third part of our study, users were randomly presented with each color of sets A and B and asked to pick exactly two colors which mix into that color from a palette. Different palettes were assigned to each set, derived from the way each color was obtained. For set A, the color palette consisted of the four main colors, while the palette for set B consisted of both the main colors and colors from set A, as depicted in Figure 1. Colors were displayed over a white background and each blended color was displayed individually, to avoid visual artifacts. On the following step, we performed color blending in three different color spaces for all colors in sets A and B and asked users to pick the most natural transition, as depicted in Figure 2: either HSV (additive), CMYK (subtractive) or CIE-LCh (perceptual uniformity). Lastly, a satisfaction questionnaire was performed where subjects were asked to rate, given a 5-point

Likert scale (1 point = complete disagreement; 5 points = complete agreement), each of the following: (i) I found it easy to decide which pairs of colors resulted in a given color; (ii) I found it easy to decide the most natural blending option between two colors. After designing our study, we created a web page with questionnaire, using HTML, CSS, Javascript and Perl/CGI technology. Colors were converted to web-safe versions in order to avoid unnecessary artifacts. We broadcasted it on social networks and kept the questionnaire open for a week.

Figure 1: *Color blending options.*

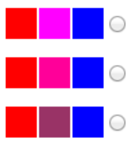

Figure 2: *Color spaces for color blending representation.*

3.2. Study results

Out of 73 subjects, 22 (30.14%) were between 16 and 25, while 49 (67.12%) were between 26 and 35. A person (1.37%) was between 46 and 55 and another was more than 66 (1.37%). Considering gender, 47 (64.38%) were male and 26 (35.62%) female. Regarding academic background, 15 (20.55%) completed high school, 17 (23.29%) have a BsC, 33 (45.21%) hold a MsC and 8 (10.96%) have a PhD. Of our subjects, 68 (93.15%) are european, 2 (2.74%) are asian and 3 (4.1%) are south american. Concerning residence, 70 (95.89%) live in Europe, 2 (2.74%) in Asia and 1 (1.37%) in South America. The Ishihara test showed that all 73 (100%) participants have normal color vision.

Success rates for blending perception tests to set A are as follows: $RG = 31.51\%$, $RB = 32.88\%$, $RY = 84.93\%$, $GB = 45.21\%$, $GY = 80.82\%$, $BY = 4.11\%$. Results ($\overline{X_{success}} = 46.58\%, \sigma = 31.19$) were not particularly encouraging. Nevertheless, scores were higher for both *RY* and *GY*, which correspond to smaller angles in the CIE-LCh color wheel (62.85° and 33.17°, respectively), while lower success rates seem to correspond to wider angles (*BY* (156.56°), *RG* (96.02°) and *RB* (93.71°)). In order to ascertain correlations between angle amplitude and success, we applied Pearson bivariate correlation tests. The one-tailed test showed evidence for correlation ($r = -0.735, sig = 0.048$), indicating higher success for lower with angular distance, a two-tailed Pearson test proved no correlation ($r = -0.735, sig = 0.096$), showing no significant relation. Concerning set B, results are as follows: $RRG = 0.64\%$, $RGG = 7.28\%$, $RBG = 0.00\%$, $RYG = 21.92\%$, $RRB = 27.40\%$, $RBB = 32.88\%$, $RYB = 19.18\%$, $RYR = 38.36\%$, $RYY = 64.38\%$, $GBG = 32.88\%$, $GBB = 67.12\%$, $GYB = 0.00\%$, $GYG = 71.23\%$, $GYY = 68.49\%$, $BYB = 0.00\%$, $BYY = 0.00\%$. Results ($\overline{X_{success}} = 28.23\%, \sigma = 27.01$) are, in general, considerably low, indicating that either choosing two out of ten colors to describe a blended color may be confusing or that subjects were not able to correctly perceive the original colors that mixed into this color. Nevertheless, blending specific colors seems to yield promising success rates, such as *GYG* (71.23%), *GYY* (71.23%) and *RYY* (71.23%). These results are in line with color set A results, showing promise in CIE-LCh blending for green and yellow, as well as red and yellow. Similarly, colors which did not blend well for set A propagated low success rates for set B, resulting in extremely low results: *RRG* (0.64%), *RBG* (0.00%), *BYB* (0.00%), *BYY* (0.00%). Results for these colors were predictable: subjects were not able to recognize the original colors that mix into the corresponding set A colors. Results for color-space blending perception tests for set A are depicted in Table 1. Whenever blending color in two different color spaces yielded a similar visual result, we presented both as a single option. Rates which correspond to merged options are visually merged on the table. We attempted at finding a correlation between success rates obtained at the previous test and color model preferences by performing Pearson correlation tests. One-tailed tests pointed to a correlation between success rates and preference for the LAB-LCh color model ($r_{success,LCh} = 0.996, sig = 0.027$), showing that preference for this model increases with success in perceiving CIE-LCh color blending. However, two-tailed tests failed to confirm a relationship between success and color model preferences ($r_{success,HSV} = -0.936$, $sig = 0.230$, $r_{success,LCh} = 0.996$, $sig = 0.053$, $r_{success,CMYK} = 0.826$, $sig = 0.38$) or inability to decide on a color model ($r_{success,LCh} = -0.682$, $sig = 0.52$). Table 2 presents results obtained on color-space blending perception tests for set B. As expected, results have shown that users prefer the CIE-LCh color space for blended colors whose provenience they correctly identify in the previous test step (*RYY*, *GBB*, *GYG* and *GYY*). We attempted at finding a correlation between success rates and color model preferences on color blending but Pearson correlation tests failed to confirm such correlation, either regarding color models ($r_{success,HSV} = 0.050$, $sig = 0.853$, $r_{success,LCh} = 0.017$, $sig = 0.949$, $r_{success,CMYK} = -0.096$, $sig = 0.723$) or inability to decide on a color model ($r_{success,LCh} = -0.068$, $sig = 0.801$). Nevertheless, there seems to be a preference for the CMYK model, with an average result of 46.30% against the other options (CIE-LCh=34.65%, HSV=20.47%, NA=5.99%). In order to find whether preferences for CMYK are relevant, we performed further analysis. We run a Shapiro-Wilk test, showing evidence against normality $p_{HSV} = 0.01$, $p_{LCh} = 0.775$, $p_{CMYK} = 0.08$, $p_{NA} = 0.092$. We assumed a non-normal distribution and applied pairwise Wilcoxon tests. For set A, since we only tested for four colors, there wasn't sufficient data to find significant differences ($p_{CIE-LCh,HSV}$=1.000, $p_{CIE-LCh,CMYK}$=0.109, while $p_{HSV,CMYK}$=0.593 and also $p_{CIE-LCh,NA}$=0.109, $p_{HSV,NA}$=0.109, $p_{CMYK,NA}$=0.109). As for set B, there are significant differences among CIE-LCh and HSV ($p_{CIE-LCh,HSV}$=0.009), HSV and CMYK ($p_{HSV,CMYK}$=0.005) and choosing a color space is significantly different from not choosing one ($p_{CIE-LCh,NA}$=0.001, $p_{HSV,NA}$=0.01, $p_{CMYK,NA}$=0.001). However, we found no significant differences between the CMYK and CIE-LCh color models ($p_{CIE-LCh,CMYK}$=0.535). Results prove that either CMYK or CIE-LCh models yield better results than the HSV model. However, despite lack of statistical evidence supporting the CMYK model, it may show some promise when blending colors that are more difficult to perceive using the CIE-LCh color space.

Table 1: *Rates for each color space (color set A)*

Color	HSV	CIELCh	CMYK	NA
RG	20.83	**69.44**		9.72
RB	9.72	22.22	**68.06**	0.00
RY	37.50		23.61	**38.89**
GB	13.89	30.56	**52.78**	2.78
GY	**77.78**		19.44	2.78
BY	**68.06**	8.33	18.06	5.56

Table 2: *Rates for each color space (color set B)*

Color	HSV	CIELCh	CMYK	NA
RRG	9.72	**47.22**	41.67	1.39
RGG	2.78	**69.44**	26.39	1.39
RBG	11.11	12.50	**68.06**	8.33
RYG	2.78	13.89	**83.33**	0.00
RRB	6.94	8.33	**80.56**	4.17
RBB	27.78	**43.06**	29.17	0.00
RYB	20.83	**36.11**	33.33	9.72
RYR	27.78	27.78	**40.28**	4.17
RYY	15.28	**34.72**	33.33	16.67
GBG	19.44	**51.39**	27.78	1.39
GBB	36.11	**31.94**	30.56	1.39
GYB	**72.22**	0.00	18.06	9.72
GYG	15.28	**50.00**	26.39	8.33
GYY	16.67	**45.83**	25.00	12.50
BYB	8.33	**47.22**	30.56	13.89
BYY	4.17	**77.78**	5.56	12.50

Regarding user satisfaction, subjects did not find it extremely difficult to understand which pairs of colors resulted in a given blended color (Score=1, 15.07%), but the majority found it slightly difficult (Score=2, 34.25%). In fact, only 23.29% believed this decision to be of medium difficulty and other 23.29% found it relatively easy, while a small number of people (Score=5, 4.11%) found it extremely easy. Despite the higher perceptual quality of the CIE-LCh, users did not find it intuitive to choose original colors that generate a given blended color in such space. The second part of the satisfaction questionnaire aims at understanding the most natural model for color blending. Results reveal that the majority of participants experienced slight decision difficulties (Score=2, 32.88%), while almost as many subjects found it to be neither difficult not easy (Score=3, 30.14%). Even though only 12.33% found it difficult to decide, only a small percentage found it either relatively easy (Score=4, 20.55%), while a minority found it extremely easy (Score=5, 4.11%). In fact, none of the adopted models stands out for color blending representation. Such conclusion is in line with the previous quantitative results.

3.3. Discussion and design implications

We performed color blending using different models, finding evidence supporting the CIE-LCh model. Despite it taking human color vision perception into account, it does not differ significantly from the CMYK model, regarding user preference on perceived blending, although HSV yields significantly lower results. Although there is lack of statistical evidence supporting preference for the CMYK model, it may show promise when blending colors that are more difficult to perceive using the CIE-LCh model. In fact, satisfaction results have shown that users did not find it intuitive to perceive the provenience of a given color that had been CIE-LCh-blended. However, due to the use of subtractive color models based on pigment mixing in early childhood [GB04], many people may have an idea of color that is closer to the CMYK model. These implications may be verified by studying user performance on blending perception taking both CIE-LCh and CMYK models into account.

4. Conclusions and Future Work

Color is effective for conveying information, providing means for labeling and categorizing. We conducted a user study to investigate whether color blending provides users with the means to understand the provenience of data items. We have also studied the color model in which to perform color blending so that users find it intuitive. Results showed that people have difficulties in understanding blending of colors that are farther apart in the CIE-LCh color wheel and indicated that the CMYK model may show promise for representing blended colors for being closer to people's idea of color. We intend to perform a follow-up study to investigate blending perception taking both CIE-LCh and CMYK models into account.

5. Acknowledgments

This work was supported by national funds through FCT Fundação para a Ciência e a Tecnologia, under project Educare - PTDC/EIA-EIA/110058/2009 and INESC-ID multiannual funding - PEst-OE/EEI/LA0021/2013.

References

[CC04] CHAPMAN N., CHAPMAN J.: *Digital Multimedia*. John Wiley & Sons, 2004. 2

[CWM09] CHUANG J., WEISKOPF D., MOLLER T.: Hue-preserving color blending. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1275–1282. 2

[DAK92] D. ALWIS D., KON C.: A new way to use the ishihara test. *Journal of neurology 239*, 8 (1992), 451–454. 2

[FvDFH90] FOLEY J. D., VAN DAM A., FEINER S. K., HUGHES J. F.: *Computer Graphics: Principles and Practice (2Nd Ed.)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990. 2

[GB04] GOSSETT N., BAOQUAN C.: Paint inspired color mixing and compositing for visualization. In *IEEE Symposium on Information Visualization* (2004), pp. 113–118. 2, 4

[Her20] HERING E.: *Grundzuge der Lehre vom Lichtsinn*. Springer, 1920. 2

[HSIHK06] HAGH-SHENAS H., INTERRANTE V., HEALEY C., KIM S.: Weaving versus blending: A quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization* (2006), pp. 164–164. 2

[LDA13] LIVINGSTON M., DECKER J., AI Z.: Evaluating multivariate visualizations on time-varying data. In *Proceedings of the Conference on Visualization and Data Analysis* (2013). 2

[LQY*13] LIANG X., QIAO W., YANG J., XUE B., QIN J.: Color interpolation algorithm of ccd based on green components and signal correlation. In *Proceedings of the International Symposium on Photoelectronic Detection and Imaging* (2013). 2

[MG80] MEYER G., GREENBERG D.: Perceptual color spaces for computer graphics. In *SIG-GRAPH 78* (1980), pp. 254–261. 2

[Poy96] POYNTON C.: *A Technical Introduction to Digital Video*. John Wiley & Sons, 1996. 2

[War12] WARE C.: *Information Visualization: Perception for Design*. Elsevier, 2012. 1

[You02] YOUNG T.: The bakerian lecture: On the theory of light and colours. *Phil. Trans. R. Soc. Lond 92* (1802), 12–48. 2