Studying color blending for visualizing social artifacts

Sandra Gama Daniel Gonçalves
INESC-ID and Instituto Superior Técnico, Universidade de Lisboa
R. Alves Redol, 9, 1000-029 Lisboa, Portugal
sandra.gama@ist.utl.pt, daniel.goncalves@inesc-id.pt

Abstract

Given the tools and applications that are available today, we often communicate at any time and place, generating high volumes of social personal information. Visually representing such data may enable us to navigate and manage it in an effective manner. Color, due to its effectiveness in labeling and categorizing information, may highlight relevant data and further alleviate cognitive load associated with information interpretation. We have studied the use of color blending for representing social artifacts. In this paper we present a user study which we have performed to understand which colors are associated with different social artifacts. It has allowed us to understand which colors humans associate with contacts and conversation topics regarding either send and received messages, as well as volume and variety. Results have enabled us to derive a set of design implications on how to use color blending to represent social facets when visualizing social personal information.

Keywords

Color Blending, Information Visualization, Social Interaction Visualization, Design Principles, Human-Machine Interfaces

1. INTRODUCTION

Not only is visualization a powerful way to represent information, but it also alleviates cognitive load associated with data interpretation [Ware 12]. When it comes to coding information, using color to display data categories is usually the best choice. This suggests representing data entities as visual objects, in which colors naturally code their attributes. When visualizing Social Personal Information (SPI), there are specific aspects that seem particularly relevant to users, namely contacts and keywords [fBR10]. Furthermore, representing trends associated with the ratio between sent and received messages seems relevant as well.

To take advantage of the potential of color to emphasize relevant SPI, we must understand how to use color to represent and highlight social artifacts. Hence, we have conducted a user study with focus on understanding how to use color blending in social information visualization. We have thus studied which colors to use to visualize both contacts and conversation topics, focusing on the representation of the ratio between sent and received messages regarding each facet. We also analyzed to which extent message volume and contacts' and topics' variety may be associated with color features such as brightness and saturation.

Naturally, color sets needed to be chosen to convey either contacts and topics, allowing the representation of sent-received message ratio. Color blending provides a natural choice to code this information, in which different color pairs are used to represent contacts and topics, and the ratio between sent and received messages ranges from one of the original colors (meaning, for instance, *In all my communications with John, he sent all the messages*) to the other original color (meaning *I sent all the messages to John*). The same applies to messages exchanged regarding a particular topic. For instance, one of the original colors represents *In all the messages about computers I have exchanged, all were sent by me*, while the other represents *All messages about computers were sent to me*. So that visual confusion is avoided, we must choose two pairs of colors with different components. To perform such choice, we only took into account the four pairs of colors which produce good results when blended together: *(red, yellow)*, *(green, yellow)*, *(green, blue)* and *(red, blue)* [Gama 14].

This paper is organized as follows: in section 2 we introduce several concepts regarding color representation and present relevant work in the context of our study. Them, in section 3 we describe our study and then present and discuss the results of user tests, deriving a number of design implications. We then draw a set of conclusions regarding color for visualizing social information.

2. BACKGROUND

Color is a sensation produced in the brain [Chapman 04] that, if related to measurable phenomena, allows digital representation. It presents great potential, allowing the visualization of data entities in a way that alleviates cognitive load associated with information interpretation. One interesting use of color in 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. For instance, Gossett and Baoquan [Gossett 04] 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. [HS06] studied information-carrying capacities of color blending and color weaving to encode multivariate information in map-reading. Livingston and Decker [Livingston 13] 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. Although relevant in such particular contexts, previous research has not specifically studied the use of color blending in a social context, particularly regarding which colors are associated with different social facets. We analyzed such aspects, by understanding how human subjects are capable of assigning social artifacts to certain colors in particular.

3. STUDYING COLOR FOR SPI VISUALIZATION

In order to understand the best way to use color blending to represent SPI, we performed a user study with the following goals: (i) to verify which pairs of colors are more easily associated with each conversation facet (*contacts* and *subjects*); (ii) taking these pairs of colors into account, to understand which colors are associated with sent and received messages; (iii) whether other color features, such as brightness and saturation, were associated with measures such as volume and variety of contacts or topics.

3.1. Designing the study

We created a four-stage study. The first part consisted of a small number of profiling questions: age, gender, education, nationality and country of residence. Since our focus is color, at the second stage we performed a validated simplified 6-plate Ishihara color blindness test [DA92]. The third part consisted of color code questions, described in section 3.1.1. Lastly, subjects were asked to rate, given a 5-point Likert scale (1=completely disagree; 5=completely agree), the statement: *I found it easy to decide which color to associate with different facets of social information.*

3.1.1. Color code questions

This test stage was designed to understand the best colors for representing social artifacts. We started by asking subjects the aspect they consider more relevant in computer-mediated interactions (*contacts* or *topics*). Since we need two color pairs to represent sent and received messages for each facet, we created two scenarios using distinct color pairs build from different original colors. In Scenario A we considered the *(red, yellow)* and *(green, blue)* pairs and in Scenario B we considered the *(green, yellow)* and *(red, blue)* pairs. For each scenario, we created a weekly calendar, as depicted in Figures 1 and 2.

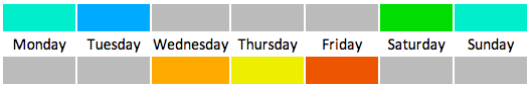

Figure 1. Weekly Calendar View (Scenario A)

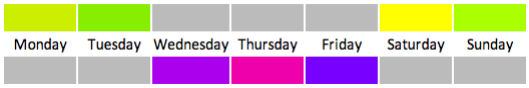

Figure 2. Weekly Calendar View (Scenario B)

On calendar views, colors from each pair were randomly assigned to each of two different sets of days *(Monday, Tuesday, Saturday and Sunday)* and *(Wednesday, Thursday and Friday)*. For both scenarios, subjects were asked on which of the two sets of groups they (i) communicate with more people; (ii) discuss a wider variety of topics.

In order to find out the best colors of each pair to represent either sent or received messages, we created two new calendar views for each scenario, in which we assigned pure color components to each of the aforementioned sets of days, as depicted in Figures 3, 4, 5 and 6. Subjects were asked which set of days corresponded to a higher number of received (over sent) messages.

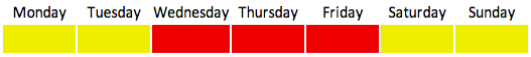

Figure 3. Weekly Calendar View (Scenario A: *(red,yellow)* pair).

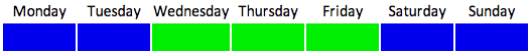

Figure 4. Weekly Calendar View (Scenario A: *(green,blue)* pair).

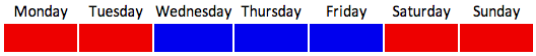

Figure 5. Weekly Calendar View (Scenario B: *(red,blue)* pair).

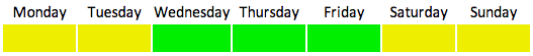

Figure 6. Weekly Calendar View (Scenario B: *(green,yellow)* pair).

In order to study possible variation of the perception of message volume and quantity of contacts or topics with brightness and saturation, we created two additional questions. On the first question, subjects were presented with a calendar view in which a group of days was represented in full brightness while the other was depicted in half the original brightness. Participants were asked to choose the group of days in which they exchanged a higher quantity of messages. On the second question, a calendar view was depicted in which a group of days was represented in full saturation while the other was depicted in half the original saturation. Subjects were asked to choose the group of days in which a wider variety of subjects were discussed or messages were exchanged with a higher number of people.

3.2. Conducting the study

After designing our study, we created a web page to perform our study as a five-step questionnaire using HTML, Javascript and Perl/CGI. We broadcasted it on social networks and kept the questionnaire open for a week.

3.3. Results

We obtained 54 responses of which 14 were incomplete, hence discarded. 40 responses were thus analyzed.

3.3.1. Subject profile

Concerning age, out of 40 participants, 11 (27.5%) were between 16 and 25, while 29 (72.5%) were between 26 and 35. Considering gender, 28 (70%) were male and the remaining 20 (30%) were female. Regarding academic background, 7 (17.5%) completed high school, 8 (20%) have BsC degrees, 23 (57.5%) hold a MsC and 2 (5%) have completed a PhD. Out of our subjects, all 40 (100%) are european. Concerning current residence, 39 (97.5%) live in Europe and 1 (2.5%) lives in Asia.

3.3.2. Color blindness

The Ishihara test showed that all 40 (100%) participants have normal color vision.

3.3.3. Color coding questions

We started the color coding questions by asking subjects which aspect was considered the most important: contacts or subjects. In fact, 21 participants (52.5%), were not able to decide. Nevertheless, 18 participants (45%) considered contacts to be the most relevant aspect, while only one person (2.5%) thought subjects to be the most important facet.

Regarding color scenario questions, results are summarized in Table 1. In scenario A, results do not establish a preference for the representation of contacts: 20 participants (50%) chose the *(red, yellow)* color pair while the remaining 20 subjects (50%) chose the *(green, blue)* pair. However, there is a stronger preference for conversation subjects. They seem to be associated with colors situated within the *(red, yellow)* color range, as picked by 28 subjects (70%) over the *(green, blue)* color pair, which was selected by only 12 participants (30%). As for scenario B, despite the slight preference for the use of colors within the *(red, blue)* range for representing contacts

Table 1. Colors associated with social facets

Scenario A	Contacts	Subjects
(red, yellow)	20	28
(green, blue)	20	12
Scenario B	Contacts	Subjects
(green, yellow)	18	12
(red, blue)	22	28

(22 subjects (55%)) over the *(green, yellow)* color pair (18 participants (45%)), there is a stronger bias for using the same color range for representing conversation subjects. In fact, 28 (70%) participants selected colors within the *(red, blue)* range for representing subjects, while only 12 subjects (30%) chose the *(green, yellow)* color pair.

Concerning the selection of colors for representing both send and received messages among the color pairs we have considered, user test results are summarized in table 2. Regarding scenario A, concerning the first color pair *(red, yellow)* there seems to be a strong preference for using *red* (28 participants (70%)) over *yellow* (12 participants (30%)) to represent received messages. Considering the second color pair *(green, blue)*, there is a slight preference for using *blue* (18 participants (45%)) for representing received messages over *green* (22 participants (55%)). Regarding scenario B, there is also a strong preference for *red* in the first color set *(red, blue)* (28 participants (70%)) over *blue* (12 participants (30%)). As for the second color pair *(green, yellow)*, there seems to be a slight preference on *yellow* for representing received messages (17 participants (42.5%)) over *green* (23 participants (57.5%)). Interestingly, *red* is strongly associated with received messages, either over another *hot color* such as yellow and a *cold color* such as blue. This fact suggests that participants associate the color *red* with action or attention, which is concomitant with occidental cultures' color conventions, in which *red* is used to draw people's attention to important situations, such as the need for action or imminent danger. On the other hand, *green* has had the opposite results, not being chosen to represent received messages over either yellow (17 (42.5%) against 23 (57.5)) or blue (18 (45%) against 22 (55%)), which is concomitant with the color convention for this color and easily associated with sent messages, in which no further action is required. Since all participants have occidental provenience, there is a strong coherence between cultural color convention and unconscious perception of action associated with received messages.

Table 2. Colors for received messages

Scenario A		
(red, yellow)	29: Red	11: Yellow
(green, blue)	18: Green	22: Blue
Scenario B		
(red, blue)	28: Red	12: Blue
(green, yellow)	17: Green	23: Yellow

Regarding the possible relationship between either message volume and color brightness and variety of contacts or conversation topics and color saturation, results are summarized as follows. The majority of participants (28 (70%)) associated message volume with colors with lower brightness over colors with lower brightness (12 (30%)). Regarding contacts' and subjects' variety in exchanged messages, a very strong interrelation was evident from our test results: 36 subjects (90%) associated highly saturated colors with a higher variety of contacts and subjects while only 4 (10%) associated these less saturated colors. Hence, we may conclude that not only does lower brightness strongly suggest a higher message volume (darker colors represent a larger amount of messages) but also higher saturation is strongly associated with the exchange of messages with a greater variety of people or about a higher number of topics.

3.3.4. Satisfaction questionnaire

Regarding results for user satisfaction on the association of color with social facets, even though subjects did not find it extremely difficult to associate color with social facets (Score=1, 12.5%), an important part found it slightly difficult (Score=2, 25%). A smaller number of subjects believed this decision to be of medium difficulty (Score=3, 17.5%) and the majority found it relatively easy (Score=4, 35%), while a small number of people (Score=5, 10%) found it extremely easy. Calculating a weighted average of the scores showed users did not find it particularly natural to associate colors with social artifacts ($\overline{X_W} = 3.05$).

3.4. Design Implications

The analysis of our study results has shown that (i) Contacts are the most relevant aspect to take into account; (ii) There are stronger preferences for a particular pair to represent conversation topics, namely *(red, yellow)* (Scenario A) and *(red, blue)* (Scenario B); (iii) In one of the selected scenarios (Scenario B), there are conflicting decisions regarding the representation of either contacts or topics using the pair *(red, blue)*. This suggests that color pairs considered in this scenario should not be used, in order to avoid cognitive confusion. Hence, the best color pairs to use are *(red, yellow)* and *(green, blue)*; (iv) When there is a color pair which includes the color *red*, received messages are associated with that color; (v) When there is a color pair which includes the color *green*, received messages are associated with the other color of the pair (either *yellow* or *blue*); (vi) Colors with lower brightness correspond to higher message volume; (vi) Colors with higher saturation are associated with a higher variety of subjects or conversation topics in exchanged messages.

4. CONCLUSIONS AND FUTURE WORK

We have conducted a user study which aims at providing a set of guidelines on how to take advantage of color in social information visualization. Such considerations suggest using colors within the *(green, blue)* range to represent contacts and colors with components from the *(red, yellow)* pair to depict conversation topics. Regarding the first pair, *green* must be used to represent sent messages and *blue* to depict received messages. Concerning subjects, *yellow* must be used to represent sent messages and *red* to depict received messages. Furthermore, the study suggest that associating brightness and saturation with message volume and subjects' or topics' variety, respectively, will succeed in delivering further information. Using these results will enable our system to meet users' expectations, showing them meaningful social information in a natural, understandable way, while alleviating cognitive load.

5. ACKNOWLEDGEMENTS

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 multianual funding - PEst-OE/EEI/LA0021/2013.

References

[Chapman 04] Nigel Chapman and Jenny Chapman. *Digital Multimedia*. John Wiley & Sons, 2004.

[DA92] Dilogen D. Alwis and C. Kon. A new way to use the ishihara test. *Journal of neurology*, 239(8):451–454, 1992.

[fBR10] Authors Omitted for Blind Review. Title omitted for blind review. In ACM, editor, *ACM 6th Nordic Conference on Human-Computer Interaction*. ACM, 2010.

[Gama 14] Sandra Gama and Daniel Goncalves. Studying color blending perception for data visualization. In *Eurovis 2014 - Eurographics Conference on Visualization*. Eurographics, 2014.

[Gossett 04] N. Gossett and C. Baoquan. Paint inspired color mixing and compositing for visualization. In *IEEE Symposium on Information Visualization*, pages 113–118, 2004.

[HS06] H. Hagh-Shenas, V. Interrante, C. Healey, and S. Kim. Weaving versus blending: A quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color. In *3rd Symposium on Applied Perception in Graphics and Visualization*, pages 164–164, 2006.

[Livingston 13] M. Livingston, J. Decker, and Z. Ai. Evaluating multivariate visualizations on time-varying data. In *Proceedings of the Conference on Visualization and Data Analysis*, 2013.

[Ware 12] Colin Ware. *Information Visualization: Perception for Design*. Elsevier, 2012.