Understanding popular relationships among colors through the network analysis for crowd sourced color data

CONFERENCE PAPER · MAY 2015

EunJin Kim
Korea Advanced Institute of Science and Te…

Hyeon-Jeong Suk
Korea Advanced Institute of Science and Te…

Understanding popular relationships among colors through the network analysis for crowd sourced color data

EunJin KIM,[1] Hyeon-Jeong Suk[1]
[1] Department of Industrial Design, KAIST

ABSTRACT

This study makes a unique attempt to adopt Social Network Analysis(SNA) in order to exploit the potential of the crowd sourced large-scale color data. We anticipate understanding the trend of color combinations and the way of generating harmonious color schemes. Basic concepts in SNA are introduced along with the method of constructing a color network from the existing color database called Adobe Kuler®. In relation to the characteristics of SNA, the implication of adopting SNA is discussed while demonstrating its benefits through findings from the network analysis of a color database. Various ways of visualization of the color data are suggested to illustrate the difference among them. In addition, quantitative metrics are calculated which identify the relative importance of individual colors and the entire structure of a network. The result reveals key players of a color network and subgroups of colors which are harmonized well to each other.

1. INTRODUCTION

Utilizing a harmonious color scheme is often essential to make an aesthetically pleasing art work and design. In order to support a proper color choice, there have been studies that are focused on developing algorithms or tools to generate a set of color (Shen et al. 2000, Wijffelaars et al. 2008, Hu et al. 2014). However, constructing a satisfied color scheme is still challenging due to the limited practicality and weak theoretical background of existing models. In this regard, this research aims to investigate patterns of color combinations through exploring the latent relationships among colors in color schemes. A network analysis was adopted to examine the relationship of colors that are generated by crowd, and to identify key players of the color combinations and the subgroups that consist of closely related colors.

2. APPLICATION OF SNA FOR A LARGE-SCALE COLOR DATA

Social Network Analysis (SNA) has been successfully utilized in various fields to understand the structure of a community, a society and even a biological system (Ennett and Bauman 1993, Böde et al. 2007, Blanchet and James 2012). Compared to traditional research methods, SNA focuses on the relationships and emerging structures formed by relationships (Scott 2012). Due to this characteristics, SNA has been supported the identification of new paradigms and latent patterns beneath the structure. Nowadays, a growing number of researches adopt SNA to investigate the structure, determinants, and impacts of relationships between actors. To our best knowledge, however, it has been yet applied to the large-scaled color data, especially a database of color schemes built by crowd. A color scheme can be easily interpreted as relational data among colors that belongs to a same scheme. In addition, Adobe Kuler®, one of the largest color scheme database, provides not only the color schemes but also the ratings for each scheme that are given by numerous users. It is expected that existing color schemes and their ratings can provide relevant guidelines to identify a harmonious combinations of colors that is applicable to a new art work. Thus we made an attempt to explore the color schemes of Adobe Kuler using SNA.

Table 1. Summary of the charcateristics of SNA and its implications on a color network .

Characteristicss of Social Network Analysis	Related SNA metrics	Implications of applying SNA to a color network	Related examples from SNS studies
Quick and easy visualization of big database	-	An easy and intuitive understanding of popular colors (See 4.1)	-
Identifting the relative importance of a social actor depending on the network position	Closeness centrality	Identifying universal colors which are easy to match with diverse colors (See 4.2)	A user who can reach to all other users with fewer steps using a friendship network
Identifying the subgroups of a network according to the relational data.	Modularity	Identifying a group of colors that are often used together in order to construct a color scheme (See 4.3)	A group of users that have intimate and dense friendships with each other

Application of network analysis to a crowd sourced database has several advantages. First of all, SNA enables a quick and easy visualization of large scaled data. Due to the visual properties of color data, an immediate visualization of data facilitates initial analysis and interpretation of the entire dataset. Secondly, network analysis focuses on the relational data, and this focal point is much more appropriate to investigate the compatibility of a color rather than its popularity that can be easily captured and compared by conventional methods. SNA provides quantitative metrics such as closeness centrality that measures the relative importance of a node in the aspect of connectivity a node has. This can be employed to identify universal colors that have been belonged to many color schemes with diverse combinations. SNA also supports a clustering of nodes using a modularity algorithm that identifies subgroups of intimate colors based on the relational data.

In conclusion, SNA allows rich and informative analysis to investigate latent but practical knowledge that traditional methods are hard to discover. Table 1 summarizes the characteristics of SNA and its implication on the study of color database with the comparison of examples from Social Network Service(SNS) studies (Perry-Smith 2006, Newman 2006) .

3. CONSTRUCTING A COLOR NETWORK

3.1 Data collection and processing

In order to construct a color network, we utilized 44,986 color themes of Adobe Kuler(Adobe Color CC) which were collected and distributed by O'Donovan et al (2011). We filtered out 7,118 popular color themes whose scores are higher than 4.0 out of 5.0. Then we transformed each color theme into nodes and links data. A node indicates a single color, and a link is a connection between a pair of colors which are included in the same color scheme. Since every color scheme consists of five colors, we were able to extract five nodes and ten links (10= (5 x 4)/2) from each color scheme.

In addition, the RGB scores as it was initially crawled were mapped into the CIE1976L*a*b space. The L, a, b scores were rounded off to every 5(for example a color with L:71, a:14, b:13 was transformed to L:70, a:15, b:15). Consequently the color quantization aggregated multiple colors into one, and helped to construct a concise and denser network.

3.2 Visualization in Gephi®

Utilizing the quantized Lab values of Adobe Kuler data, a color network with 4,922 nodes and 26,401 edges was generated. An opensource software Gephi was utilized in order to visualize the network and anlyze it. Figure 1 shows a visualized color network which is composed of colors that are used more than 50 times.

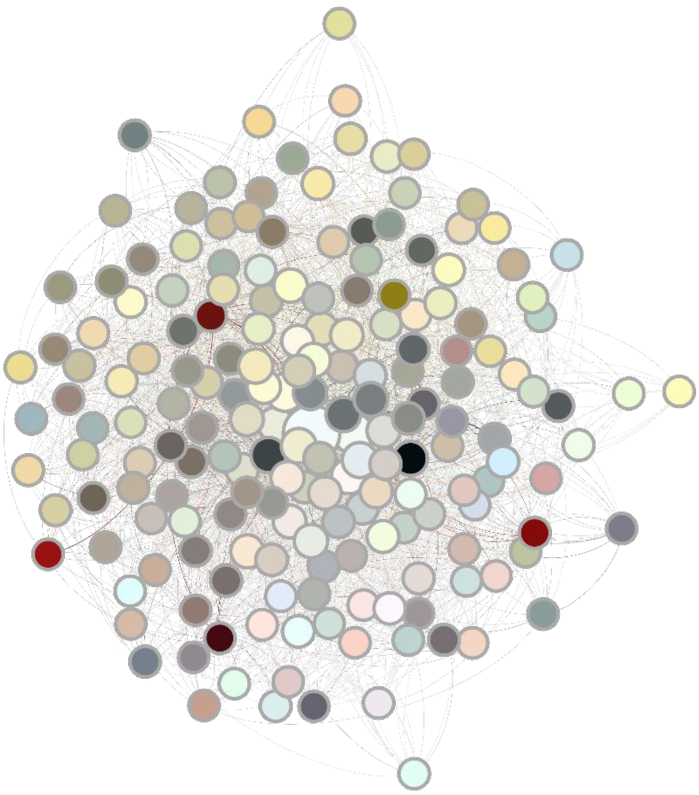


Figure 1: A color network of colors which are utilized more than 50 times

4. FINDINGS FROM THE NETWORK ANALYSIS

4.1 Identifying key players through a visualized network

As previously mentioned, SNA has notable characteristics and advantages that are different from other research methods. First, it provides an effective way of visualization which enables an intuitive understanding of the entire data. Figure 2 shows three networks which utilized the same data but devised different filtering criteria to emphasize the key players of a color network. The left network provides a look of network that are composed of colors utilized more than 20 times (degrees). It covers 37.85 % of the entire links, and shows the general tendency of the color utilized by crowd. The middle one shows popular colors that are utilized more than 70 times. As shown in the graph, most of reddish colors appeared in the low range of L, whereas bluish colors more frequently had higher L values. The right figure was generated by applying the criterion of degrees more than 200. In this case, only a few colors are visualized, and it is really easy to identify the extremely popular colors that are frequently utilized throughout the database. To summarize, SNA provides an easy and effective way of representing the data while investigating the relationships with diverse perspectives. It also provides an intuitive view to compare more than two different color networks. Especially when it combines with time-dependent data, it is expected that the trend change of popular colors can be traced and compared easily using the color network.

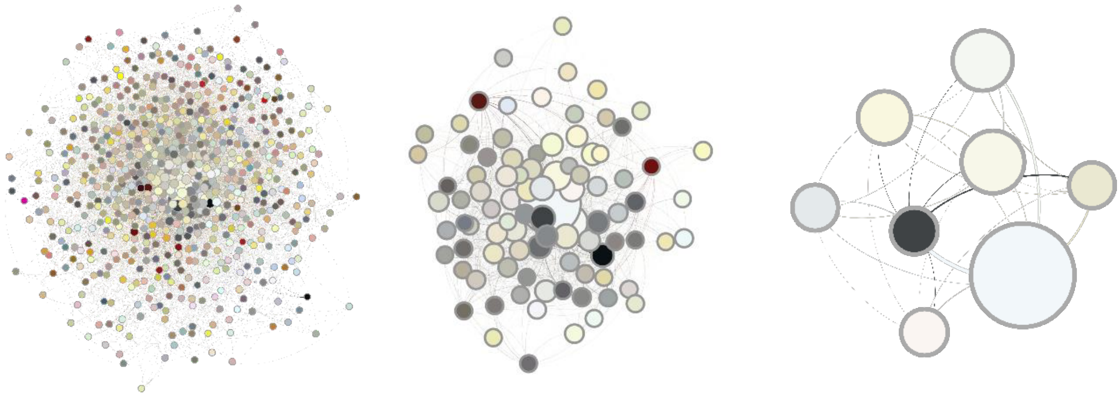


Figure 2: Networks with different filtering criteria - nodes of 20, 70, 200 degrees respectively

4.2 Identifying universal colors using a closeness centrality measure

A color scheme is often constructed with more than 2~3 colors. Hence, it is important to identify colors that can be universally matched with other colors instead of capturing single color with a extreme popularity. In this regard, a color which has a closer relationship with other colors is more meaningful than a color which has been frequently used as a member of a color scheme. The measure of closeness centrality relates to the concept of universal match, and it calculates the average distance from each node to every other nodes in the network. For instance, the value 2 of closeness centrality indicates that the node can reach to other nodes within two degrees in average. The higher closeness centrality implies that a color can be connected to other colors with a shoter distance, and it increases the possiblity that a color belongs to various color schemes. Therefore, a color with a higher closeness centrality can be regarded as a universal one which is compatable to any combinations.

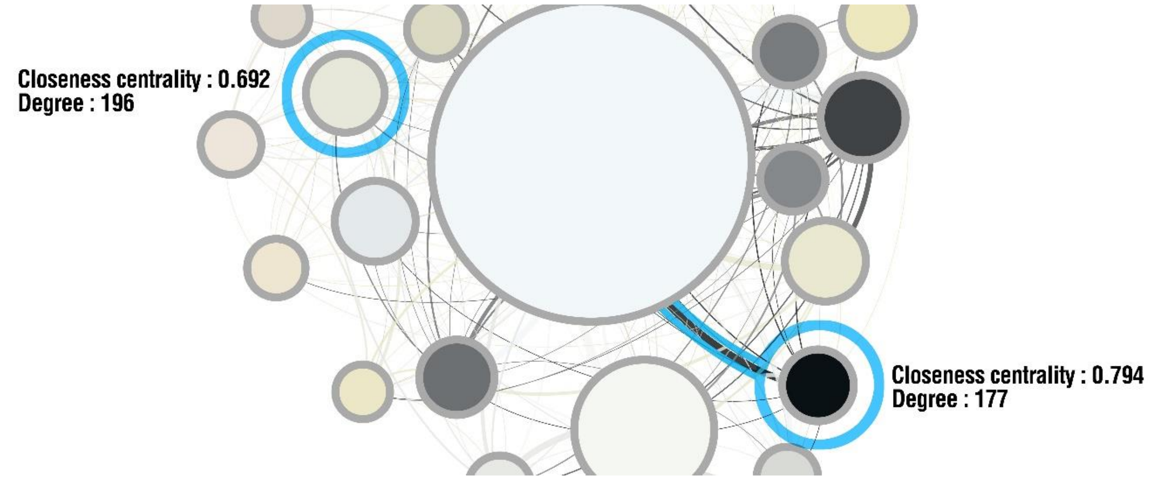
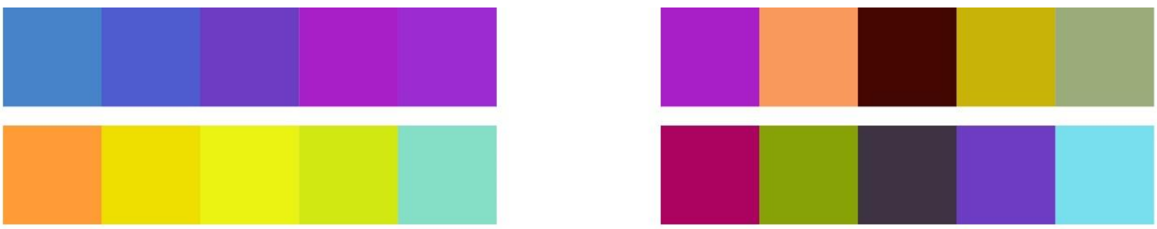


Figure 3: A partial network to compare two colors with different closeness centralities

In a network, colors which show higher closeness centralities often overlap with the popular colors which are utilized more frequently than others, but not always the same. Figure 3 shows two colors in a network with a closer look; the highlighted color on the right side of the network represents a color with a higher closeness centrality (0.794) whereas the left one represents a popular color utilized more frequently but less central (0.692). As shown in the graph, the color with higher centrality has strong relationships with other colors that are popular as their size represent. However the left one with a lower centrality value has relatively less links with other popular colors. The strong connectivity with other popular colors increase the possibility that the right color reach to other colors to construct various combinations. In general, colors with remarkable centralities have extreme L values and lower chromaticities throughout the entire graph.

4.3 Identifying groups of harmonious colors using a Modularity measure

The modularity algorithm looks for nodes that are more densely connected together than to the rest of the network (Newman 2006). The nodes that belong to the same modularity group has a strong relationship with each other because they have been frequently utilized in the same color scheme together. Therefore the modular membership of a color can be employed to construct a color scheme which has been proved as harmonious combination by numerous users. As shown in the Figure 4, color schemes that consist of colors from the same modular groups are relatively more harmonious than schemes composed of colors with different modular memberships.



(a) Colors from the same Modularity group (b) Colors from the different Modularity group

Figure 4: Color schemes that are generated using the Modularity of colors

5. CONCLUSION

In this study, Social Network Analysis has been adopted in order to explore the crowd-generated color big data. We have discussed the method to construct a network using the exsiting color database, and have introduced relevant metrics and methods that are applicable to the color data. The constructed network was analyzed and interpreted by visualizing it with different filtering criteria, and by calculating network measurements such as closeness centrality and modularity. The measurement of closeness centrality reveals that colors with extreme L values and low chromaticities are more compatable with other colors. We also generated color schemes by employing the modularity memberships of colors, and the colors from the same modular groups produeced more harmonious color schemes. In conclusion, the network analysis provides a noble way of investigating color data in the perspective of relatinships among colors, and suggests findings that conventional methods are hard to reveal. However this study is intended as a priliminary research to introduce the implication and benefits of applying a network analysis. For further studies, it is required to develop relative analysis methods and verify its significance in order to enhance reliability and practicality of network analysis in color studies.

REFERENCES

Böde, Csaba, István A Kovács, Máté S Szalay, Robin Palotai, Tamás Korcsmáros, and Péter Csermely. 2007. Network analysis of protein dynamics. *Febs Letters* 581 (15):2776-2782.

Blanchet, Karl, and Philip James. 2012. The role of social networks in the governance of health systems: the case of eye care systems in Ghana. *Health policy and planning*:czs031.

Ennett, Susan T, and Karl E Bauman. 1993. Peer group structure and adolescent cigarette smoking: A social network analysis. *Journal of Health and Social Behavior*:226-236.

Hu, Guosheng, Zhigeng Pan, Mingmin Zhang, De Chen, Wenzhen Yang, and Jian Chen. 2014. An interactive method for generating harmonious color schemes. *Color Research & Application* 39 (1):70-78.

Newman, Mark EJ. 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* 103 (23):8577-8582.

O'Donovan, Peter, Aseem Agarwala, and Aaron Hertzmann. 2011. Color compatibility from large datasets. In *ACM SIGGRAPH 2011, Proceedings*, Canada, 1-12.

Perry-Smith, Jill E. 2006. Social yet creative: The role of social relationships in facilitating individual creativity. *Academy of Management Journal* 49 (1):85-101.

Scott, John. 2012. *Social network analysis*. 3rd ed: SAGE Publications Ltd.

Shen, Yu–Chuan, Wu–Hsiung Yuan, Wen–Hsing Hsu, and Yung–Sheng Chen. 2000. Color selection in the consideration of color harmony for interior design. *Color Research & Application* 25 (1):20-31.

Wijffelaars, Martijn, Roel Vliegen, Jarke J Van Wijk, and Erik ‐ Jan Van Der Linden. 2008. Generating color palettes using intuitive parameters. In *Computer Graphics Forum, Proceedings*, 743-750.

Address: Prof. Hyeon-Jeong Suk, Department of Industrial Design, KAIST
Bldg. N25, 291 Daehak-ro, Yuseong-gu, Daejeon, 305-701, KOREA
E-mails: kejkim8@kaist.ac.kr, h.j.suk@kaist.ac.kr