Colorfulness?—In Search of a Better Metric for Gamut Limit Colors

Chengwu Cui and Steven Weed
Lexmark International, Inc.
740 New Circle Road, Lexington, Kentucky

Abstract

The chromaticness of gamut limit colors of a printer is important because it is often directly compared to that of competitors. The CIELAB C^*_{ab} is often used as the descriptor of the chromatic strength or chromaticness of these colors. However, the lightness of a color may also contribute to the perceived chromaticness. In this study, nine highly chromatic red color samples of the same CIELAB hue but having small variations of CIELAB chroma and lightness were prepared and scaled for colorfulness by 29 subjects using the method of paired-comparison. The results indicate that lightness also significantly contributed to the perceived colorfulness as defined by the subjects according to their everyday color experiences. At the same level of CIELAB chroma, samples of lower lightness were perceived to be more colorful. For this set of samples, colorfulness can be accurately modeled by factoring in the CIELAB L* value. The results also raise some questions regarding the definition and use of colorfulness in color appearance modeling. That is colorfulness, as implied in our everyday color experiences, is a complex parameter.

Introduction

The color reproduction capability of a color printer is limited by its device colors, which define its gamut limit. Subtle differences in gamut limit colors of two competitive printers can give one an important competitive edge when compared to each other. The process of colorant selection as device colors often raises the question as to what parameters to use to accurately describe the chromatic strength of gamut limit colors. We would like first to use the conventional definition for chromatic strength or chromaticness as: "the attribute of a visual sensation according to which the (perceived) color of an area appears to be more or less chromatic". [1]

The CIELAB space has been widely used to describe color gamut in spite of its known weaknesses including its non-uniformity and inaccurate representations of the perceptual equivalents (e.g. hue representation). The CIELAB C^*_{ab} has been used to describe chroma, a quantity largely associated with the Munsell Color Order System. Chroma is often regarded as a quantity relative to the brightness of the perfect white, at least as defined in the Munsell system.

Either due to the deviation of the C^*_{ab} from the true perceptual chroma or to the inherent complex nature of our color perception, using C^*_{ab} alone seems insufficient to describe chromaticness. For example, given the choice of two colorants of the same C^*_{ab} value but one with a higher L*, C^*_{ab} metrics may not be sufficient. Fig. 1 shows three white dots and three black dots on a red background. If one stares at the dots steadily for a short moment and then looks somewhere else on the red background, the afterimages of the dots will appear and they should be of the complementary colors of the dots, in this case, black and white, respectively. The mixture of the red color of the background and that of the afterimages of the dots will produce mixed sensations of red, which will be lighter or darker than that of the background red. The mixed colors will be perceived to be different in chromaticness from the background and from each other.

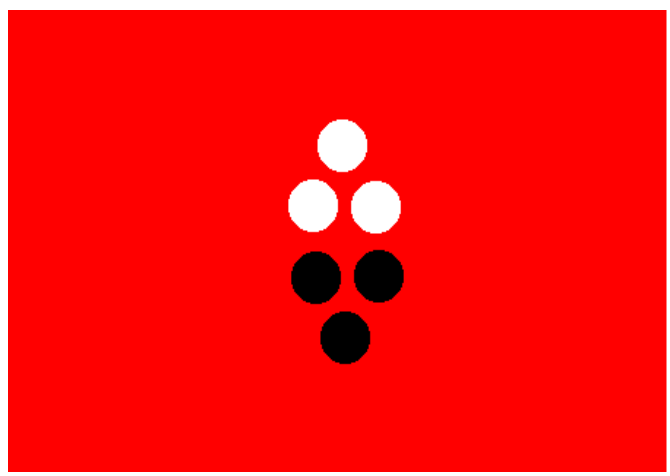

Figure 1. Colorfulness of afterimages. By looking steadily at the dots for a short moment and then looking at another part of the red background, afterimages of different chromaticness can be seen.

Given two samples of the same hue and CIELAB chroma, if we are to judge which one is more chromatic, we are in need of a better descriptor than C^*_{ab}. In color appearance modeling, the term "colorfulness" has been used to describe the same quantity or the chromaticness sensation. In color appearance modeling, colorfulness is emphasized for its dependence on luminance level as characterized by the Hunt effect.[2,3] Chroma is considered as a special case of colorfulness, or colorfulness relative to the perfect white under the same illuminant. The concern with the newer definition is its deviation from the meaning of colorfulness in its everyday use sense. The Webster dictionary gives a simple and straightforward definition: "having striking colors".[4] The magic word is "striking" because it inevitably gives a single "striking" color the sense of being colorful although, to most people, the plural form of color seems to be the condition. It seems that the use of colorfulness in color appearance modeling deals with the "striking" aspect of color sensation for a single color while in everyday use, colorfulness often implies the presence of multiple hues. Despite the variant definition of the redefined "colorfulness" in color appearance modeling, it seems colorfulness has been treated according to its everyday use meaning for its measurement in color appearance modeling (e.g. see reference 3).

In image quality evaluation, e.g. in printing quality evaluation, colorfulness is often one of the quality aspects to be evaluated in surveys. Although the quality of colorfulness can be largely associated with the color correction algorithms, the colorfulness of the gamut limit colors can make a significant contribution to the evaluation outcome. Therefore, it seems colorfulness is a good candidate to specify the chromaticness of printer gamut colors.

In this paper, we describe a visual experiment to determine the effect of lightness (L*) on the perceived colorfulness of a series of nine highly chromatic red color samples of the same H* value but with small variations in C^*_{ab} and L*. We found that L* significantly contributed to the perceived colorfulness for this set of samples. We also discuss the complexity of colorfulness quantity.

Experiment

Given a set of color samples of the same hue but with small variations of lightness and chroma, one can measure the perceived colorfulness by a psychophysical test. One of the most powerful visual scaling methods is the paired comparison test based on the law of comparative judgment. The success of the paired comparison test relies on the reliable administration of the test, the selection of confusable color samples and the proper application of the computation model to the data.

Apparatus

Color sensation is a complex process and is affected by virtually every factor in the visual field such as the illumination of the visual field and the observer's physical and emotional conditions. The goal of the test setup is to minimize potential effects by other factors. On the other hand, the setup of the test has to be consistent with how the colors would be seen in the real world.

The test was conducted in a windowless room with walls painted to a neutral gray (N7). Fig. 2 shows the viewing configuration.

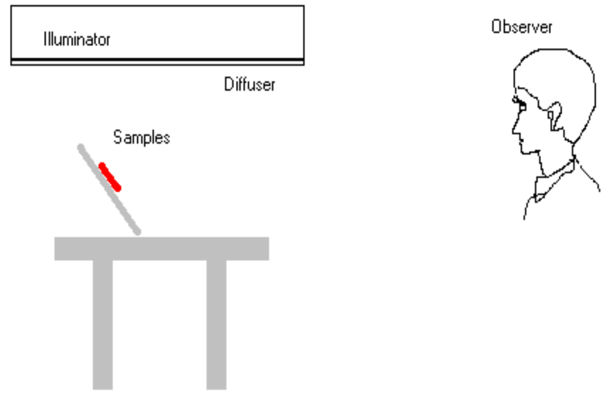

Figure 2. Sample viewing configuration.

The illuminator was a 8-lamp GretagMacbeth overhead fluorescent daylight D50 luminare. On the sample plane, the illumination was about 900lux. Although the luminare was equipped with an optical diffuser, the illumination on the sample plane was considered a mixture of directional illumination and diffuse illumination. The sample holder was adjusted such that the observer could not possibly see specular reflections from the samples. The distance from the observer to the samples was about 35 inches so that the samples (about 0.9" x 0.9") subtended about a 1.5° angle. At such a distance, variations in the surface texture and other printing artifacts on the samples were invisible to the observer.

The administration of the experiment was assisted by a computer program specially designed according to the principle of comparative judgment. The program logged all the information of the observer and randomized the presentation order of the pair combination as well as randomize the left-right presentation order of the pair to be compared. The software also drove a set of speakers to play pre-recorded instructions for each step of the test. As directed by the program, the test administrator presented the samples and recorded the observer's responses. The program then processed these responses into a psychometrically scaled colorfulness report.

Sample Preparation

The goal of sample selection was to select a set of samples that were easily confusable and representative of highly saturated gamut colors. The color red was chosen for this test. A series of red samples were printed using a Lexmark Optra™ Color 45 inkjet printer with a photo print cartridge on Lexmark™ high resolution premium inkjet paper. The samples were measured with a GretagMacbeth SpectraLino™. The CIELAB L*, C^*_{ab}, h_{ab} values were computed using the CIE 1931 standard observer and the spectral power distribution of the GretagMacbeth™ D50 fluorescent luminare system, measured by a Photo Research PR650™ spectroradiometer. Nine samples of the same CIELAB hue but with small variations of L* and C^*_{ab} were selected. For these samples, the average h_{ab} is about 24.7±0.3°. The $L^*C^*_{ab}$ coordinates of the samples were shown in Fig. 3.

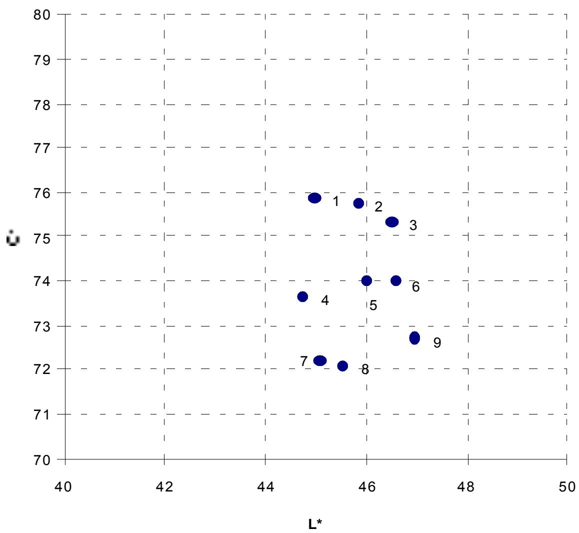

Figure 3. Sample color variation in the C^*_{ab} L plane.*

It can be argued that equal h_{ab} may not warrant the same perceived hue. However, initial visual examination showed that there were no visible variations in perceived hue, which was further confirmed by the observers.

The samples were attached to a 5"x4" neutral gray card as shown in Fig. 4 a). For paired comparison, samples were arranged to have the two color samples abutting each other with a hairline in between as shown in Fig. 4 b).

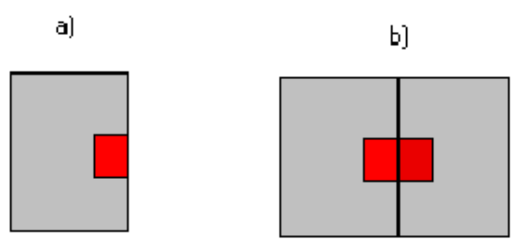

Figure 4. Sample configuration. a) shows a single sample; b) shows the way two samples were compared.

Observers

Twenty-nine observers participated in this test. They were colleagues at Lexmark International, Inc, with ages range from 20 to 50 years old. A few of the observers were skilled color science workers. The majority were engineers and general office workers. They all passed the Munsell-Farnsworth™ 100 hue test immediately prior to this test in the same room and therefore were considered fully adapted to the illumination used in this test.

Procedure

Administration of the test was fully controlled by the computer program. To prepare each observer for the test, the following recorded instructions were played:

"Welcome to this experiment. This experiment is designed to test the colorfulness attribute of various color samples. You will be presented with a pair of samples and asked to make a decision, based on your everyday experience of color, on which color is more colorful, the sample to your left or the sample to your right. Remember you must make a decision even if you are not sure. It is Ok to guess if you have to. Just relax and take your time."

Note that the phrase "based on your everyday experience of color" was meant to be consistent with the instruction from Pointer's colorfulness scaling experiment.[3]

The test administrator then followed the sample presentation order given by the program and recorded the response from the observer by pressing the corresponding left-right button. With the 9 samples, each observer made a total of 36 comparisons. At the end of the test, the observers were specifically asked to report their criteria for choosing the more colorful sample.

Results

Scaled Colorfulness

Results from the 29 observers were recalled into the data processing module of the computer program that computed the Z score, or scaled colorfulness, according to the basic comparative judgment computation scheme.[5] The scaled colorfulness for the nine samples is plotted versus C^*_{ab} and shown in Fig. 5.

The average prediction error was computed by using the scaled colorfulness data to back-predict comparison probability data. This was found to be about 8%, which is considered satisfactory. Another approach to estimate the theoretical scaling errors, following Braun and Fairchild,[7] was also used and obtained a standard error of 0.09 (colorfulness scale unit). This standard error is also shown in Fig. 5.

Regression

Fig. 5 shows that C^*_{ab} cannot fully explain the measured colorfulness of these samples. A function of C^*_{ab} with L* involved was used to re-adjust the C^*_{ab} chroma scale according to the scaled colorfulness data and produced the following equation:

$$C^*_{Colorfulness} = 54.38 + 0.1C_{ab} * \left[1 + \left(\frac{C_{ab}^*}{L^*}\right)^{4/3}\right] Eq. \quad (1)$$

The equation was then used to calculate the adjusted CIELAB chroma, plotted against the measured colorfulness, and shown in Fig. 6.

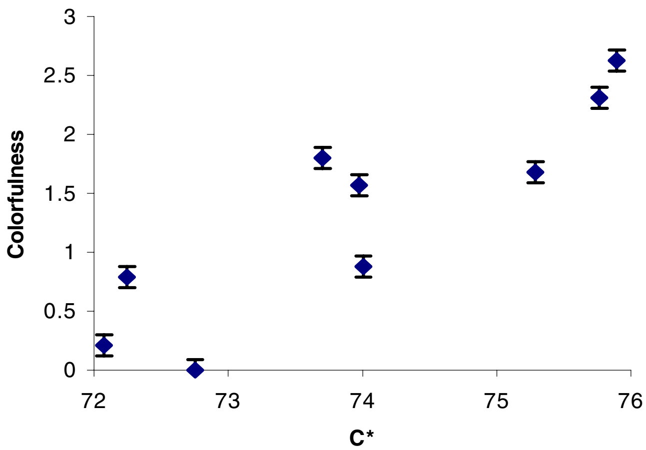

Figure 5. Colorfulness versus the CIE C^*_{ab} chroma. The error bars represent the theoretical standard error of the measured colorfulness.

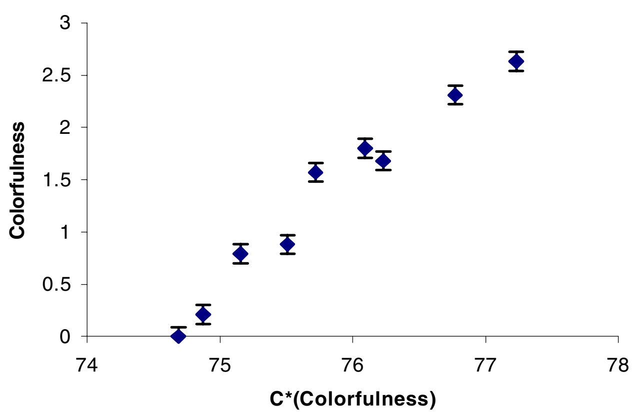

Figure 6. Colorfulness versus the new Chroma $C^*_{Colorfulness}$ scale. The error bars represent the theoretical standard error of the colorfulness.

Observers' Criteria for Colorfulness

Observers' criteria for colorfulness judgment were generally consistent. These criteria included: "looked for the least washed-out color"; "the more appealing one"; "the more vivid color and often the darker but not dirty one"; "the more saturated one". Some reported that they tended to choose the slightly brighter color when the difference was subtle.

No observers reported seeing differences in hue that could make the comparison potentially difficult.

Discussion

Eq. 1 can provides a significantly improved metric for colorfulness (as defined according to the everyday experience of color) for this specific set of samples used in this test under the specific test conditions. However, it is unlikely that it will hold for the entire color space because when L* gets lower, the metric easily over-exaggerate colorfulness. Eq. 1 was only intended to demonstrate that a sound relationship can be derived for colorfulness involving L*. Further data are needed if we are to derive a universal equation.

The outcome of this test is that, given two highly chromatic red colors of the same C^*_{ab} value, the darker color will be preferred for being more colorful. This result should be expected if it can be agreed that the majority of the population regard the slightly darker color as the more saturated color and, therefore, more colorful.

Industrial colorists tend to use their own sets of color attributes. They express attributes of color in association with concentration and density of colorants. It seems in everyday life, colorfulness is somewhat related to color material concentration, or "saturation" to its very original meaning.

On the other hand, there were reports by some of the observers that sometimes they chose the slightly brighter one for being more colorful when the difference is subtle, contradictory to the overall results of this test. However, this phenomenon can be related to the well-known Helmholtz-Kohlrausch (H-K) effect,[6,7] which states that the perceived brightness increases with the increase of saturation. In other words, for two colors of the same brightness, the more chromatic color appears to be brighter; or a more chromatic color will appear to be brighter. It seems, then, the converse of the H-K effect should be: for two colors of the same chroma, the lighter one should appear more chromatic. The interesting question is why the overall results do not follow the converse of the H-K effect. We believe this was due to the observers' criteria for colorfulness during the test deviated from chroma. Instead, they might have associated colorfulness with the density of inks. To them, a darker color is more saturated.

It seems that the meaning of colorfulness in our daily life is complicated. As demonstrated by Fig.1, which set of afterimages is considered more colorful probably depends on how they strike the observer's fantasy, as to what they are made of and what they are intended for and so forth.

In recent color appearance modeling, colorfulness is mainly used to reflect the Hunt effect. It can be problematic when appearance models are fitted to experimental data, which often involve experiments employing ordinary people with their own criteria for being colorful.

Conclusions

Nine highly saturated red color samples of the same CIELAB hue but of small variations of CIELAB chroma and lightness were carefully prepared and scaled for colorfulness by 29 subjects using the method of paired-comparison. The results indicate that lightness also significantly contributed to the perceived colorfulness (in the sense of the subjects' everyday experience of color). At the same level of CIELAB chroma, samples of lower lightness were perceived to be more colorful. For this set of samples under the specific conditions, colorfulness can be accurately modeled by factoring in the CIELAB L* value. The results can be used to guide printer device color selection in regard to colorfulness.

The results also raised some concerns regarding the use of colorfulness in contemporary color appearance modeling. The colorfulness attribute of color vision seems to be a complex attribute. It can be traced back to the perception of concentration and density of colorants or colored materials in nature. It can also be related to luminance and intensity of light. It can be related to the presence of a large number of different hues as implied in our everyday color experiences.

Acknowledgement

We would like to thank Mr. Russell Huffman for administering the test, Mr. Cary Ravitz for helpful suggestions and all the observers for volunteering.

References

1. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative data and Formulae*, 2nd ed. John Wiley & Sons, p. 487. (1982)
2. R. W. G. Hunt, "The specification of colour appearance. Concepts and terms", *Col. Res. Appl*. 2, P.55-68. (1977)
3. M. R. Pointer, "The concept of colorfulness and its use for deriving grids for assessing color appearance", *Col. Res. Appl.* **5**, P. 99-107. (1982)
4. *Merriam Webster's Collegiate Dictionary*, 10th Ed. 1994, Merriam-Webster, Inc. Springfield, Massachusetts, USA, P227.
5. W. S. Togerson, Chapter 9: "The law of comparative judgment", *Theory and Methods of Scaling*, Wiley, New York. (1976)
6. D. B Judd, "A new look at the measurement of light and color", *Illum. Eng.* **53**, 61-71. (1958)
7. M. D. Fairchild and E. Pirrotta, " Predicting the lightness of chromatic object colors using CIELAB", *Col. Res. Appl.* **6**, P. 385-393. (1991)
8. J. Davidoff, *Cognition through Color*, The MIT Press, Cambridge, MA. (1991)

Biography

Chengwu(Luke) Cui received his BS degree in optics from Shandong University, MS in color science from Chinese Academy of Science and PhD in vision science from the University of Waterloo. From 1995 to 1999, he worked for GretagMacbeth as a color scientist. He recently joined Lexmark International. His research interests include human vision, ocular optics, image quality, colorimetry, lighting and computer color formulation.