Weaving Versus Blending: a quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color

Haleh Hagh-Shenas, Sunghee Kim, Victoria Interrante, Christopher Healey

Abstract— In many applications, it is important to understand the individual values of, and relationships between, multiple related scalar variables defined across a common domain. Several approaches have been proposed for representing data in these situations. In this paper we focus on strategies for the visualization of multivariate data that rely on color mixing. In particular, through a series of controlled observer experiments, we seek to establish a fundamental understanding of the information-carrying capacities of two alternative methods for encoding multivariate information using color: color blending and color weaving. We begin with a baseline experiment in which we assess participants' abilities to accurately read numerical data encoded in six different basic color scales defined in the L*a*b* color space. We then assess participants' abilities to read combinations of 2, 3, 4 and 6 different data values represented in a common region of the domain, encoded using either color blending or color weaving. In color blending a single mixed color is formed via linear combination of the individual values in L*a*b* space, and in color weaving the original individual colors are displayed side-by-side in a high frequency texture that fills the region. A third experiment was conducted to clarify some of the trends regarding the color contrast and its effect on the magnitude of the error that was observed in the second experiment. The results indicate that when the component colors are represented side-by-side in a high frequency texture, most participants' abilities to infer the values of individual components are significantly improved, relative to when the colors are blended. Participants' performance was significantly better with color weaving particularly when more than 2 colors were used, and even when the individual colors subtended only 3 minutes of visual angle in the texture. However, the information-carrying capacity of the color weaving approach has its limits. We found that participants' abilities to accurately interpret each of the individual components in a high frequency color texture typically falls off as the number of components increases from 4 to 6. We found no significant advantages, in either color blending or color weaving, to using color scales based on component hues that are more widely separated in the L*a*b* color space. Furthermore, we found some indications that extra difficulties may arise when opponent hues are employed.

Index Terms—Color, perception, visualization, color weaving, color blending.

1 Introduction

Domains related to analysis and understanding of the turbulent flow, meteorology, geology and astronomy are a few examples of applications in which the field experts regularly look at and make decisions based on the relationship between several continuous or segmented variables. As a result, one of the primary struggles of Multivariate Visualization is managing the difficult task of effectively layering the data. Unfortunately, the cumulative effect of the layers and their interaction result in unwanted colors and patterns that might mislead an observer into perceiving correlations or interactions that do not exist among the original layers of data. Presenting the user with too much information may clutter the display and overwhelm the observer. It is thus imperative to keep the layers as distinct from each other as possible and if they are combined, it is done in a perceptually comprehensible manner. Additionally, the layering of information must provide coherence and precision by remaining faithful to the proper relationship among the layers of information such that different pieces of information can be visually compared to each other [12], [13].

Various strategies have been proposed for representing multivariate datasets with color, textons, icons and other graphical objects. Examples of these methods are using various types of glyph for segmented data, combining color and texture for multivariate visualizations and other integrated methods including the following articles [4], [9], [5], [14], [8], [7], [3].

The idea of "small multiples" [12] presents an efficient perception task in which by decoding the design of one slice of data, observers gain cognitive access to the rest of the data in other slices which are expressed in the same manner. The familiarity and consistency of a design with this characteristic enables the observers to maintain a better focus on the changes in the information.

Other popular strategies for showing color coded data, include assembling several semi-transparent layers on top of each other; showing all layers covering the same space sequentially in time by giving the user the ability of going back and forth through them and finally presenting various variables in separate maps.

This paper focuses on strategies for the visualization of multivariate data that rely on color mixing. Particularly, we aim to provide quantitative assessment of the relative effectiveness of two of these representations: color blending and color weaving. In color blending, the most common approach in data representation using color, a single composite color is used to convey the values of multiple color-encoded quantities. In color weaving, a form of which was recently developed for flow illustration, the individual colors of multiple variables are separately woven to form a fine-grained texture pattern. Part of our future research is to compare the effectiveness of multiple maps each representing one of the variables in a multivariate field, to a representation which combines all variables side by side in one display using different colors.

The main motivation behind color weaving is to facilitate the task of finding unknown patterns among different variables involved in a multivariate dataset. The user studies are designed to provide observers with an estimate of the error observers may expect when utilizing this representation.

In Section 1.1 we present some of the previous studies that are the most related to ours including those in the field of cartography which has devoted a considerable amount of effort to investigating the development and evaluation of two dimensional media for representing

- *Haleh Hagh-Shenas is with Boston Scientific, E-mail: haleh@cs.umn.edu.*
- *Sunghee Kim is with Gettysburg College, E-mail:skim@gettysburg.edu.*
- *Victoria Interrante is with University of Minnesota, E-mail:interran@cs.umn.edu.*
- *Christopher Healey is with North Carolina State University, E-mail:healey@csc.ncsu.edu .*

Manuscript received 31 March 2007; accepted 1 August 2007; posted online 2 November 2007.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

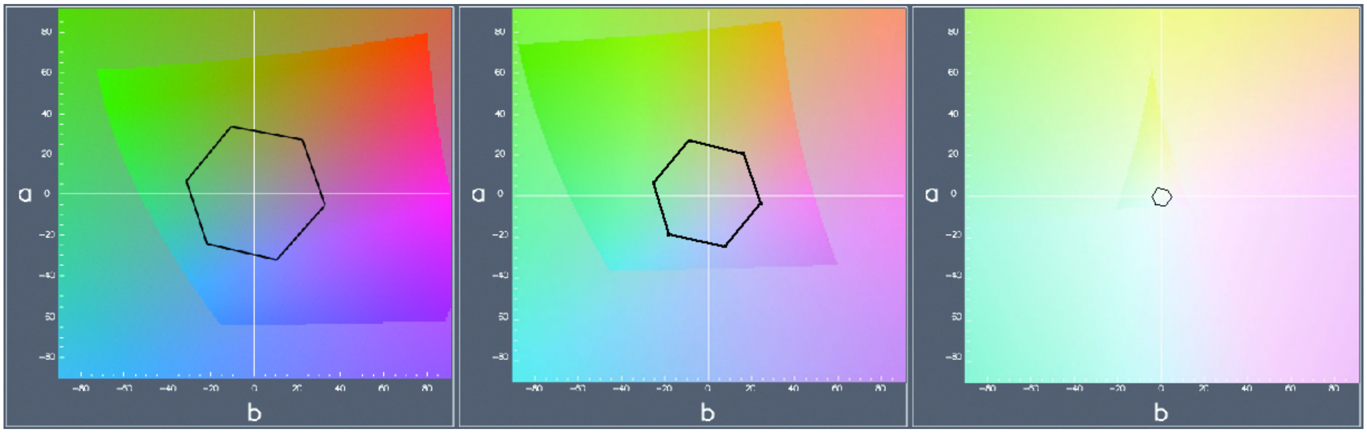

Fig. 1. Three slices of the L*a*b* color space used to construct the color ramps. From left to right: at L=62, L= 79.5, L=97

segmented and continuous data.

We begin with a baseline experiment, in which we assess participants' abilities to accurately read numerical data encoded in six different basic color scales defined in the L*a*b* color space. We then assess participants' abilities to read combinations of 2, 3, 4 and 6 different data values represented in a common region of the domain, encoded using either color blending or color weaving.

Although our present experiments employ census data, in which a single value of a particular variable applies to a moderately sized region of the domain, our ultimate objective is to facilitate the effective representation of multiple co-located continuous distributions, in which each variable typically assumes a slightly different value at every point across the domain; our use of high frequency textures in the present study is intended to lay the foundation for these subsequent investigations.

2 Previous Work

Shelton and Glimartin [6], investigated how quickly and accurately map readers viewing choropleth maps on monitors were able to identify to which class an areal unit on the map belonged to. They provided cartographers with empirical guidelines regarding the level of map reading accuracy that might be expected for choropleth maps. The maps in question had between 4 and 8 classes and were produced in shades of gray, green or magenta. As expected, increasing the number of classes on the map led to a decrease in the accuracy rate and an increase in the reaction times of participants. Hue also affected the accuracy rates and the reaction times.

In an extensive study [10], Mersey investigated the effectiveness of color in symbolizing information on thematic maps. A large number of subjects evaluated the effectiveness of various graded color schemes. These test maps employed six distinct color schemes and four different number of data categories. Subjects were asked to perform ten tasks, which attempted to duplicate realistic map use tasks dealing with both specific and general map information. Some tasks were performed by conferring to the map directly where others were completed from memory. The research emphasized a number of complex interactions among several variables: the choice of color scheme, the number of data classes presented on the map and the nature of the task to be performed with maps.

In one of the earliest studies in cartographic multivariate visualization, Rogers and Groop [11] developed a multi pattern sparse colored dot map to portray multiple distributions in the same map. A user study was conducted to compare the relative effectiveness of the multi pattern dot map to single distribution dot maps by asking participants to do various map reading tasks. Analysis of the participants' responses was based on the consistency with which region boundaries within a map were found and the internal accuracy of identified regions. Results suggest that three variable color dot maps are at least as effective as several single black and white dot maps in portraying the regions.

Urness et. al. [14] introduced the concept of color weaving in the context of flow visualization. They presented a method for simultaneously representing multiple co-located colors through coloring the strokes of LIC textures; these textures were then used to illustrate the relevant flow data. The perceptual effectiveness of this method is still unstudied.

When people view color images, they think in terms of perceptual dimensions such as saturation or darkness. Thus, a successful color design can utilize these perceptual dimensions and map them to the logical structures of the data to allow its organization to be readily perceived [1]. In [1] Brewer describes a set of guidelines for using color for visualization in the field of cartography and presents an efficient method for visualizing two overlapping variables in maps. These guidelines can be applied to choropleth thematic maps, filled isoline maps, and qualitative areal-extent maps. Depending on the nature of the data she suggests four color schemes: binary, qualitative, sequential, and diverging. For example Categorical differences can be shown by a diverging color scheme that emphasizes the meaningful midpoint in the data. However, this method can only be effectively used if the number of data bins is very small.

Some researchers including Byron [2] utilized a triaxial-graph method for the visual portrayal of soil texture in maps and geographic information systems. He displayed the trivariate nature of soil texture by generating a color legend in shape of a triangular graph known as the soil textural triangle. The three sides of this equilateral triangle characterized the composition of the soil in terms of its percentage of sand, silt, or clay. Soil texture was thus spectrally encoded by assigning one of the additive primaries (red, green, blue) to each axis of the triangle. The relative intensity of each color varied with the amounts of sand, silt, or clay. The color value for a specific location was found by adding the amount of the three primaries.

Although a few effective methods exist for color based representation of two and three variables with binned data, in case of continuous data the common method is to blend the colors. The effectiveness of blending the colors in conveying the correct information is not thoroughly investigated.

We hypothesized that despite the complications that were likely to arise when showing the colors side by side, this method had undeniable advantage over blending if the colors were chosen to be perceptually as distant from each other as possible. This paper commences the investigation of capabilities of information carrying ability of color weaving. That is done by asking basic map reading questions involving variables that have a unique value over a segmented two dimensional domain. This method with a minor change has the potential of being extended to continuous data. The effectiveness of this method for representing continuous datasets is being investigated by the authors.

In each of the following three experiments presented in this paper, participants were asked a direct acquisition map reading task and the accuracy of the answers under each condition was statistically analyzed and compared to each other. Specifically these three experiments were designed to give visualization researchers a quantitative measure of the accuracy they can expect in representing their data using color with color weaving or color blending.

3 Experiment 1

Experiment 1 was designed to assess participants' baseline ability to accurately read numerical data encoded via the intensity of a single displayed color. Specifically the focus of this experiment was to reveal any differences in the ability of the observers in performing map reading tasks with any of the chosen colors. The experiment was also used to provide a comparing baseline for the map reading error for experiment one. We observed that essentially the differences among the capability of each colors in delivering the data were negligible. In addition, the difference in the map reading error in different states was not statistically significant.

3.1 Choosing the Colors

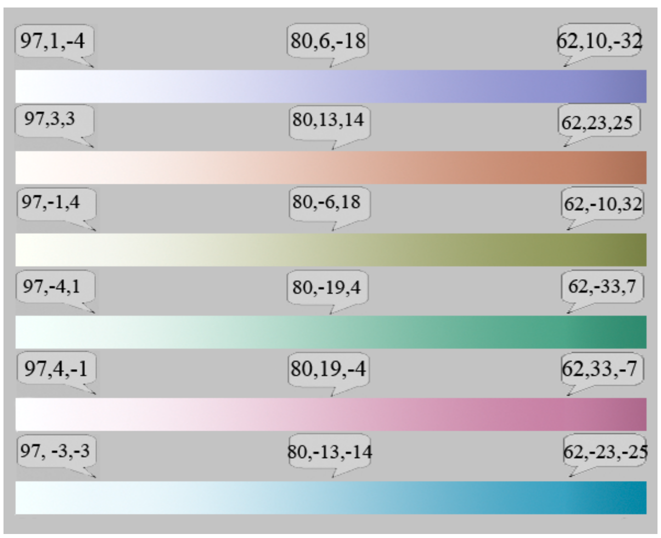

Fig. 2. Color Ramps for all the experiments

We defined the base colors by choosing six evenly spaced points around a circle of constant saturation in a plane of constant luminance in a region of the L*a*b* color space that fit within our monitor gamut. Two circles were chosen at L=62 and L=97. Figure 1 shows the maximum, minimum and the mid point slices in the portion of the color space we chose. From L=62 to L=97, we created six different perceptually linear single-hued color ramps, by continuously increasing the luminance and saturation values of each of the six different base colors. In order to test the observers ability to read the minimum and maximum values accurately, we extended our color ramp on both sides such that the final color ramps included a continuous range of colors between the minimum and maximum values in addition to the right and left extensions which consisted of colors which didn't exist in the maps. Color ramps are presented in Figure 2. The maximum value for each variable was tied to a color at L=62 and the minimum value was mapped to the corresponding color at L=97.

3.2 Method

Apparatus

The stimuli consisted of six maps of the twelve midwestern United States, in which each state was filled with a different constant color from a single color ramp, representing the value of a particular data

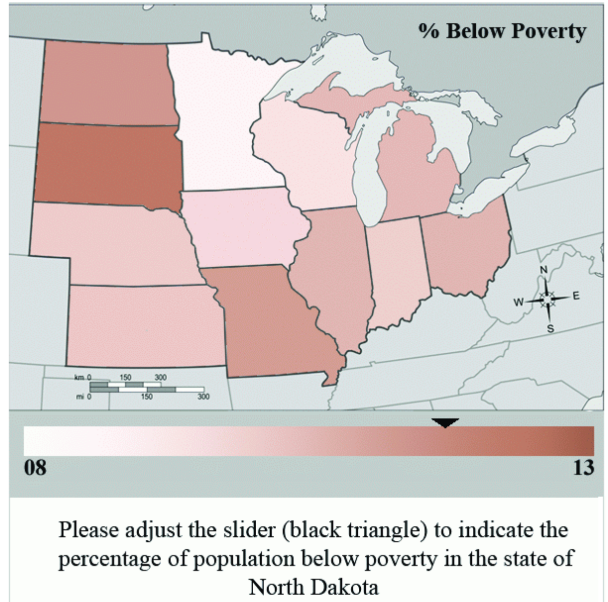

Fig. 3. Experiment 1, example of a display; State in question: North Dakota; Variable: the percentages of population below poverty level

attribute for that state. The actual data values were obtained from the 2000 census data, but the particular assignment of data values to states was randomized in experiments 1 and 2 to prevent people from using domain knowledge about the midwestern US. The six data distributions were: median household income, percentage of the population with a high school degree, percentage of the population with a college degree, percentage of the population living below the poverty line, median cost of a single family dwelling, and home ownership rate. Each color ramp was tied to one specific variable throughout experiment 1 and experiment 2. The map size was chosen to be equal to the map size in experiment 2 which was the maximum possible map size viewable on the 19 inch monitors used for the study, when showing the legends for all the six variables. The background color was chosen to be a neutral gray with the L value halfway between the largest and smallest L values in our color ramps(L=79).

Table 1.

Display Specification	
Distance to monitor	95.2 cm or 37.5 in
Monitor resolution	1280x1024
Width and Height of the maps	17.8x12.2 cm or 7x4.8 in
Width and Height of the screen	38x30.2 cm or 15x12 in

A total number of 216 displays were shown to each of the observers. For each variable a map reading question was asked for each of the states present in the maps. This amounted to 72 (12x6) displays, each of which was repeated 3 times to ensure the accuracy of the results (72*3). These displays were shown randomly to each of the participants.

A chin rest was employed to fix the distance between the participant's eyes to the monitor at the appropriate level. The display's specifications can be viewed in Table 1.

Procedure

Participants color vision ability was assessed using a collection of Ishihara plates.

A question appeared on each display asking the observers to evaluate the color in a particular state by clicking on and dragging the small slider -a black triangle-.

On each trial in this experiment, the participants task was to identify the value of a particular data attribute for a particular state by: reading the color from the map, setting a slider to the matching color on the provided color scale, and then clicking on the state to indicate that their selection is final. Figure 3 shows a screen shot from one trial. All participants were provided with a printed outline map showing the correspondence of names to states. All trials on which the state was mis-identified were discarded. Clicking on the state was added as an extra assurance that the observer was in fact looking at the correct state when performing the task. We did not enforce any breaks during any of the experiments mentioned in this paper. It was merely suggested that the participants take short breaks looking away from the monitor if they felt too tired.

Participants

There were a total of 9 participants (three females, six males aged 18-38) for this experiment. The participants were graduate, undergraduates and academic staff from University of Minnesota, Gettysburg College and North Carolina State University. Participants were compensated 5 dollars per half an hour for completing the experiment and they took on average an hour and a half to complete the study.

3.3 Result

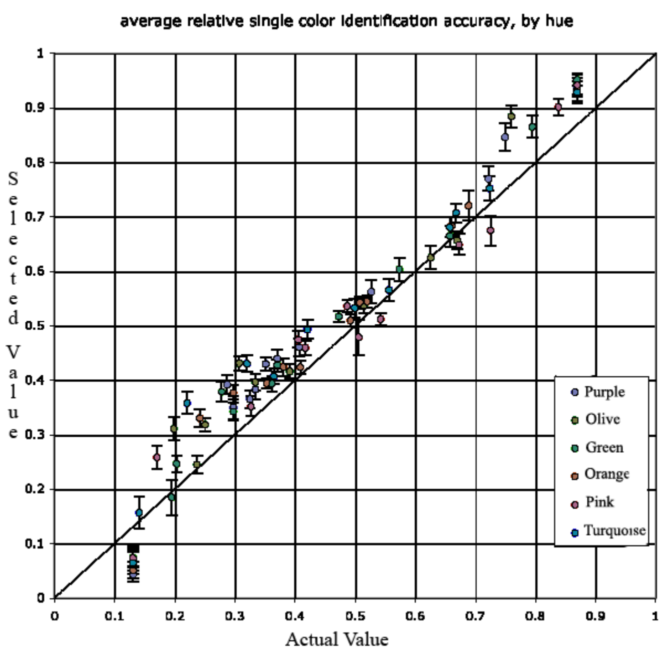

Fig. 4. Scatter plot of actual vs. average user answers in Experiment 1

Participants were uniformly able to perform the task in experiment 1 with fairly good accuracy. We calculated the relative error per display by dividing the pixel difference between the actual value and selected value of the position of the black triangle over the maximum possible error (the size of a color ramp in pixel). The average relative error, computed over all participants and all colors, was 6.02%, with a standard error of 0.57%. Figure 4 shows a scatter plot of the average errors for the displays, computed over the nine participants. The error bars show the extents of the 95% confidence intervals, and each point is color-coded according to the base color of the color scale used for that trial. A single number giving the average of the median relative error per observer, was added into the overall chart of results from experiment 2. The ANOVA analysis showed that the difference among

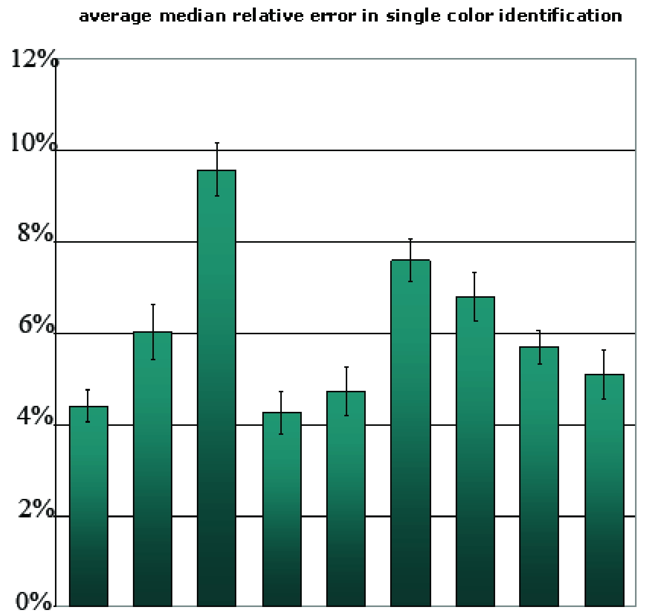

Fig. 5. Experiment 1: Average over the median relative error in the map reading task for nine participants. Each bar corresponds to one of the participants

different colors and among different states were negligible. Figure 5 shows the average over the median relative error in the map reading task for each of the nine participants.

4 Experiment 2

The goal of this experiment was to quantify the differences among the three conditions for visualizing multivariate data when variables are overlapping in a region: blending the colors representing each overlapping variable, representing the overlapping variables by letting the colors coexist in that area with a small texton high frequency pattern and again with a larger texton high frequency pattern.

4.1 Method

Apparatus

The stimuli in experiment 2 consisted of a series of maps of the twelve midwestern United States in which the values of either two, three, four or six different data distributions were simultaneously represented via either color-blending, in which the separate color layers were made semi-transparent and then overlaid to form a single composite representation, or color-weaving, in which the separate color layers were individually sampled at independent pixels defined by a random noise function and then stitched together to form a finely patchworked, unified representation. In case of the blended images, the final color was equivalent to averaging the L*a*b* values of the individual overlapping colors. We tested noise patterns of two different spatial frequencies: small noise, in which each pixel subtended 3 minutes of visual angle, and large noise, in which each pixel subtended 6 minutes of visual angle, and participants viewed all images from a fixed distance enforced by a chin rest. Screen shots of the sample stimuli are shown in Figure 6.

This experiment included a total number of 153 displays. The maps were the same size as in experiment 1. However this time, 2 or more variables were overlapping across all the states. All 2 ways (15 displays), 3 ways (20 displays), 4 ways (15 displays) and 6 way(1 display) combination of the 6 variables were considered (51 trials). Three conditions for combining the colors were investigated. Under the first condition, the 51 displays were made by blending the overlapping

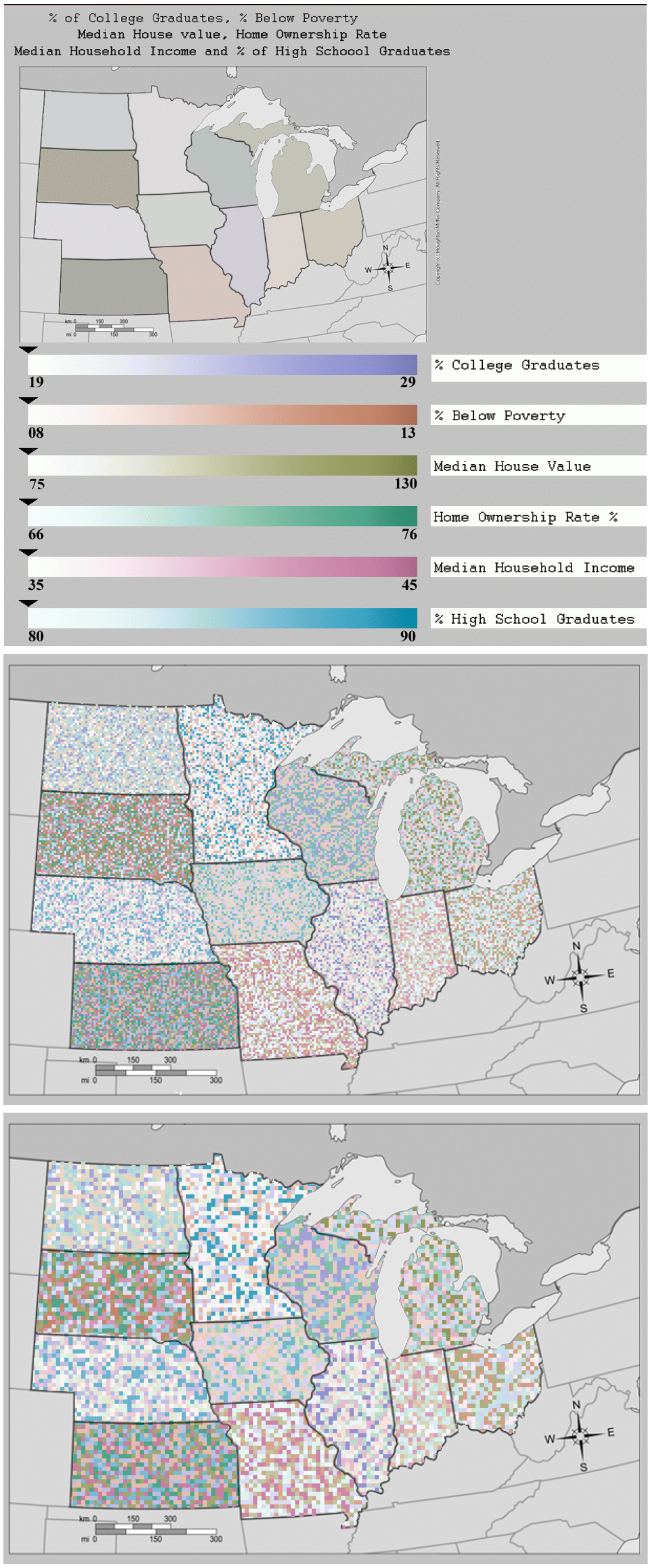

Fig. 6. Experiment 2 stimuli, 6 variables. From top to bottom: Blending 6 colors, 2-pixel size noise, visual angle 3 minutes; 4-pixel size noise, visual angle 6 minutes.

variables in each state. The second condition included 51 displays in which colors coexisted in an area by following a 2 pixel size noise pattern and finally for the third condition the colors coexisted by covering different levels of a 4 pixel size noise pattern. The patterns were made by filling the area of the map by a noise pattern of the appropriate size and posterizing the pattern into 2, 3, 4 and 6 different levels while equalizing the image histogram to get approximately the same number of pixels for each of the gray level values. This procedure ensured that there were approximately the equal number of pixels belonging to each of the coexisting colors(variables) in the states.

In order to remove the effect of any learning, for each of cases of the 2 variables, 3 variables, 4 variables and 6 variables we chose a random swapping of the values belonging to each state while at the same time ensuring that for each of the overlapping conditions, the same maps were used in making the blend, small size noise and big size noise versions.

In order to present a set of cognitively organized tasks to the observers, each observer was shown a random order of the following sets: all the blended images, all the small noise images and all the large size noise images. Within each of the blended, small noise and big noise groups, they randomly saw all the 2 variables, 3 variables, 4 and 6 variables respectively.

Procedure

Due to the large number of displays we were not able to test all the states nor were we able to include repeated measures. The instruction guided the participants to only look at the state of Iowa and make their color evaluations by adjusting up to six sliders each of which corresponding to one of the overlapping or coexisting variables.

Participants

Eighteen people (four females fourteen males, aged 21-38) participated in this experiment. Six of the participants were from university of Minnesota, six were recruited from Gettysburg College and six from NCSU.

4.2 Data Analysis and General Results

As shown in Figure 7 the results from experiment 2 indicate that the error rates were significantly lower when the original color information was available via the high frequency texture than when the colors were blended. In the case of the blended representation, error rates steadily rose as the number of components increased (a trend that we found statistically significant in an ANOVA analysis). We observed weak evidence of a similar effect in the case of the woven textures, but it was not statistically significant.

An ANOVA analysis with a standard 95% confidence interval compared two conditions: the type of color mixing (blending, small noise textures, large noise textures) and the number of variables in the map (2, 3, 4, or 6). Median accuracy for the target colors was used as a performance metric (see Section 3.3 describing Experiment 1's results for a full explanation of how color accuracy was calculated). The ANOVA analysis showed that both the type of color mixing and the number of variables had a significant main effect (mixing type: $F = 16.68$ and $p = 0.0035$; number of variables: $F = 52.51$ and $p < 0.001$). Within mixing types, Tukey's HSD analysis found that at the 95% level, blending was significantly different from, and worse than, weaving.

When looking at blended colors, the parwise comparison showed that 3 or 4 variables were significantly less accurate than 2 variables, and 6 variables was significantly less accurate than 3 or 4 variables. For both the small and the large noise textures, 6 variables was significantly less accurate than 2, 3, or 4 variables.

However the performance with 2,3 and 4 were not significantly different for either of the noise sizes. Subdividing results by number of variables (2, 3, 4, and 6), blending was significantly less accurate than noise textures in all four cases, while there were no significant differences in accuracy between the small and large noise textures in any of the four cases.

Interestingly, despite the conviction of the participants that the colors could not be differentiated from each other in the case of blending more than 2 variables, all the participants were still able to perform the task with an accuracy rate of 70% or better. It is possible for different sets of colors to produce the same rendered result in case of the blended images. This may be an important source of the error that blended colors produce. Further experiments are needed to measure how much of the error is due to the ambiguity described above, and how much is due to the fact that blended colors are not as capable as weaved colors in conveying the individual color values.

It is well known that different surroundings can change the appearance of a color. A perceptual process, called assimilation causes small color areas to appear more similar to their surroundings. The process of simultaneous contrast enhances the differences among colors between a larger patch of color and its surrounding color. These perceptual interaction between colors can change the perception of individual colors by shifting the hue and thus making the map reading task difficult. Unfortunately there is no effective method to address these issues and the designers are advised to note the presence of these problems.

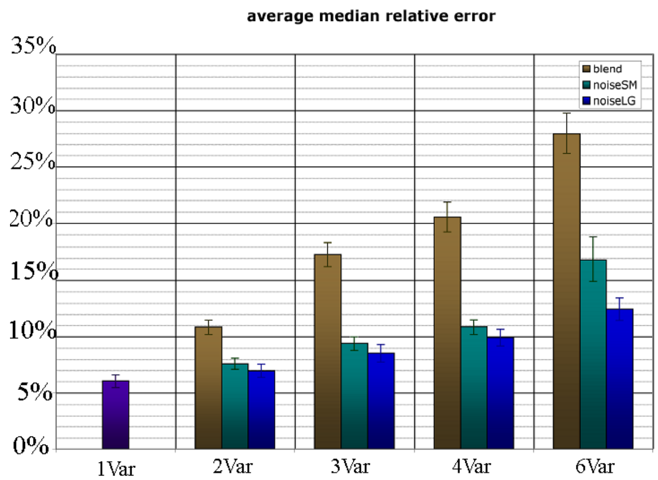

Fig. 7. Experiment 2, Average of Median Relative Error

5 Experiment 3

The color values used in Experiment 2 were defined according to real census data values. This means that the distribution of measurements, although highly representative of what would be encountered in a typical visualization, was not necessarily uniform in its distributions, and therefore may not be sufficient to answer accurately questions we encountered about viewers' ability (or inability) to accurately decompose a composite color into its constituent components.

To address this, we conducted a final experiment designed to explore the effects of two conditions, hue separation and luminance difference, on the error rates in viewers' judgments of the values of individual components in a 2-variable display. Specifically, we studied two questions: 1) will error rates be larger, smaller, or the same for reading color combinations where the hues are separated by 60, 120, or 180-degrees in L*a*b* space? and 2) will error rates be larger, smaller, or the same for reading color combinations where the luminance values of the individual components is nearly equal, moderately close, or more widely separated?

5.1 Method

Apparatus

A total of 180 displays were constructed by combining (via noise or blending) all possible combinations of two colors chosen from the six colors used in the previous experiments. Each of these $\binom{6}{2} = 15$ hue combinations were then matched with three lightness pairs: 1) one color at L=72 and the other at L=76 (nearly equal); 2) one color at L=68 and the other at L=80 (moderately close); and 3) one color at L=64 and the other at L=84 (more widely separated). This resulted in 45 displays representing the 45 different color-luminance pairs. Displaying the colors using either blending or noise textures produced a total of 90 displays. Finally, we mirrored the L-values in each color-luminance pair (e.g. L=68 for the first color and L=80 for the second, versus L=80 for the first color and L=68 for the second). This represents the final set of 180 different displays presented during the experiment. Three examples of the displays can be seen in Figure 8.

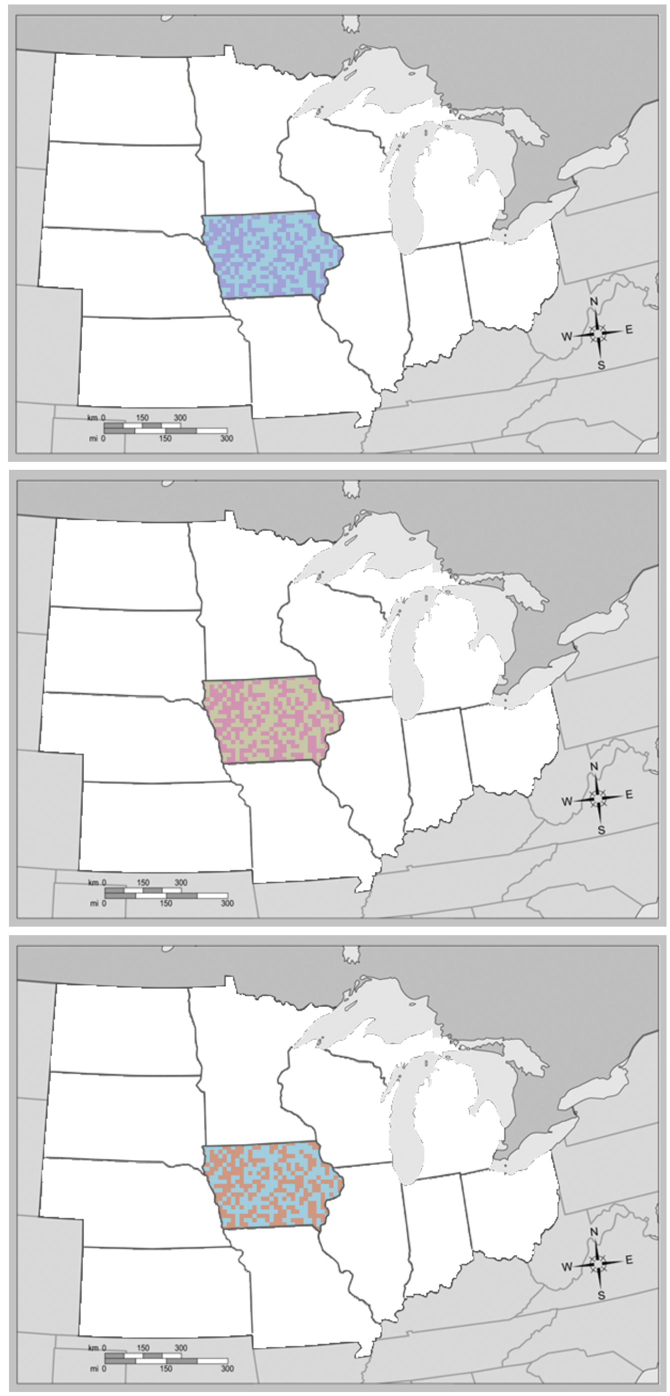

Fig. 8. Bivariate representations where (from top to bottom) two hues are 60 degrees apart, 120 degrees apart and 180 degrees apart in the L*a*b* color space. The lightness gap was 12.

Procedure
The task in Experiment 3 was identical to the task in Experiment 2. Participants were asked to read the variables present in the map in the state of Iowa. In order to remove any possible influence from the surrounding colors, all other states were colored white.

Participants
Four graduate students (two males, two females, aged 30-34), who had passed the Ishihara test participated in the experiment. It took the participants approximately an hour and a half to two hours to complete the study.

5.2 Data Analysis and General Results

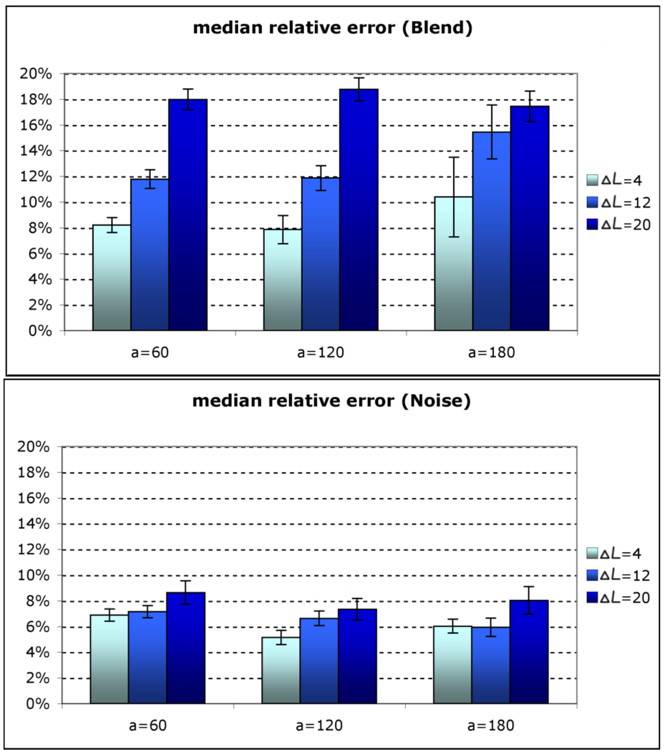

Fig. 9. The average of the median relative error for Experiment 3, the X-axis divides results by the three hue difference angles 60, 120, and 180°, the Y-axis show the median relative error averaged over participants, and the three shaded bars in each group show the results for each luminance difference $\Delta L^* = 4$, 12, and 20.

Table 2. ANOVA Results Experiment 3, looking at the subsets of the data

Type	p value	F value
Blend: Hue difference	0.1351	3.441
Blend: Luminance difference	< 0.001	20.54
Noise: Hue difference	0.9976	0.002359
Noise: Luminance difference	< 0.001	10.01

The 720 trials for the 4 participants were collapsed and averaged over trials with the same mixture type, hue difference, and luminance difference.

A 3-way ANOVA over the conditions mixture type, hue difference, and luminance difference with a standard 95% confidence interval was then performed. We found significant main effects of both mixture type ($F = 26.12$, $p = 0.036$) and luminance difference ($F = 30.22$, $p < 0.001$). Luminance difference is referred to as ΔL^* in Figure 9. Errors were higher when colors were combined by blending than when they were interwoven. Tukey's HSD analysis found that at the 95% level the performance was significantly better when the luminance values of the component colors were closer to each other, which is counter intuitive. Looking separately at individual subsets of the data, divided by mixture type and hue difference, we found a significant main effect of luminance difference in all cases except where the hues were directly complementary. Looking separately at the results involving Blended colors and those involving the interwoven texture, we did not find hue difference to have a significant effect on error rates in any subset of the data (Table 2).

6 Discussion

The results of our three experiments indicate that color weaving is consistently more effective than color blending for conveying the values of individual data distributions in a multivariate visualization. Error rates remain low for woven combinations of 2, 3 and 4 different colors and only begin to rise to a statistically significant extent when the number of component colors increases to six. The advantage of weaving over blending persists even when the area subtended by each patch of continuous color is very small.

Although the problem of inferring the values of the component colors in a blended mixture is inherently ill-posed, observers are able to perform this task fairly accurately, within a moderately constrained domain, when presented with pairs of component colors that have nearly equal luminance values, although errors rise as the luminance values of the component colors begin to differ.

We found no significant advantages, in either color blending or color weaving, to using color scales based on component hues that are more widely separated in L*a*b* color space. On the contrary, we found some indications that extra difficulties may arise when opponent hues are employed. The relative independence of the error from hue separation is probably rooted in the design of the color ramps. Observers only need to make a reasonable estimate of the lightness value for each variable in question and hue is only used for distinguishing the variables from each other. Additionally, observers know that each variable in question has an individual 2D space in the common domain. It is possible that these two facts are sufficient for making a reasonable variable estimation regardless of how close or far the hues in questions are from each other.

7 Conclusion and Future Work

The ultimate goal of this research was to make a contribution to the field of multivariate visualization by presenting a design for which the related perceptual issues are relatively well understood. In order to fully understand the advantages or disadvantages of this design we plan to apply this or similar representations to other methods of color based multivariate visualization.

The success of our design like many other graphical displays relies on the ability of people in perceiving colors. Considering that 8% to 12% of male population and 1% of females have some form of color deficiency, these designs potentially miscommunicate the information to a large percentage of population. The magnitude of the debilitating effect of the errors caused by this condition on the accuracy of the variable estimation must be investigated.

An important venue for our ongoing study is applying similar methods to continuous datasets and examining the ability of the observers in recognizing patterns and exploring relationships among different visualized variables. User studies rigorously investigating the effectiveness of these data representations in delivering recall and recognition-based map reading tasks have to be conducted. Another venue of our current study is applying the same idea of weaving color to other forms of textures.

Acknowledgements

This research was supported by the National Science Foundation through grants ACI-9875368 and CNS-0323471, and by the University of Minnesota through the Digital Technology Center. We are grateful to Dr Gary Meyer, Seth Barrier and Clement Shimizu for helping us in selecting colors and verifying the validity of the representation using their L*a*b* color space exploration tools. We are also very grateful to all of the participants in our experiments for their dedicated and conscientious efforts throughout a long and tedious task. Finally, we thank the reviewers for their thoughtful and constructive comments which helped us improve the quality of this paper.

References

[1] C. Brewer. *Visualization in Modern Cartography*, chapter 7, pages 123–147. Pergamon Press Inc., 1994.

[2] J. R. Byron. Spectral encoding of soil texture: a new visualization method. In *GIS/LIS Proceedings*, pages 125–132, 1994.

[3] E. Chlan and P. Rheingans. Multivariate glyphs for multi-object clusters. In *INFOVIS '05: Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, pages 19–27, 2005.

[4] C. Forsell, S. Seipel, and M. Lind. Simple 3d glyphs for multivariate spatial data. In *INFOVIS '05: Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, pages 119– 124, 2005.

[5] M. Gahegan, D. Guo, A. MacEachren, and B. Zhou. Multivariate analysis and geovisualization with an integrated geographic knowledge. *Cartography and Geographic Information Science*, 32:113–132, 2005.

[6] P. Gilmartin and E. Shelton. Choropleth maps on high-resolution crts the effect of number of classes and hue on communication. *Cartographica*, 26(2):40–52, 1989.

[7] C. Healey and J. Enns. Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 5(2), April-June 1999.

[8] V. Interrante. Harnessing rich natural textures for multivariate visualization. *IEEE Computer Graphics and Applications*, pages 6–11, 2000.

[9] R. Kirby, M. Marmanis, and D. Laidlaw. Visualizing multivalued data from 2d incompressible flows using concepts from painting. In *VIS '99: Proceedings of the conference on Visualization '99*, pages 333–340, 1999.

[10] J. Mersey. Color and thematic map design: the role of colour scheme and map complexity in choropleth map communication. *Cartographica*, 27:1–157, 1990.

[11] J. E. Rogers and R. E. Groop. Regional portrayal with multi-pattern color dot maps. *Cartographica*, 18(4):51–64, 1981.

[12] E. Tufte. *Envisioning Information*. Graphics Press, 1990.

[13] E. Tufte. *Visual Explanations*. Graphics Press, 2002.

[14] T. Urness, V. Interrante, I. M. E. Longmire, and B. Ganapathisubramani. Effectively visualizing multivalued flow data using color and texture. In *VIS '03: Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, pages 151–121, 2003.