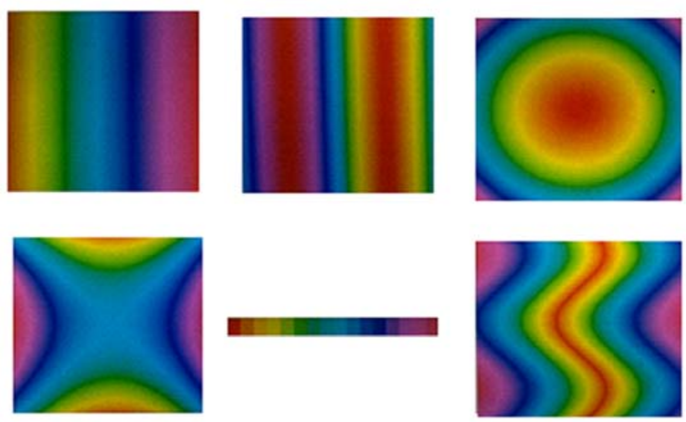
Color Sequences for Univariate Maps: Theory, Experiments, and Principles

Colin Ware
University of New Brunswick



Pseudocoloring is a widely used technique for presenting univariate map information on a graphic display system. This article divides the kinds of information available in maps into two classes. *Metric* (or value) information denotes the quantity stored at each point on the surface, and *form* information denotes the shape and structure of the surface. Theoretical principles are proposed to predict which color sequences will be effective at conveying value and form information respectively. According to this theory a scale that approximates the physical spectrum should be good at conveying metric information because of the reduced effects of simultaneous contrast. It should be poor at conveying form information, however, because the brain prefers from information to come through the lightness processing "luminance" channel. Conversely, a gray scale should be poor at conveying value information and good at conveying form information according to the same theory.

These predictions are tested in a series of psychophysical experiment which test five color sequences. The results show that simultaneous contrast can be a major source of error when reading maps, but only partially confirm the form hypothesis. Guidelines are given, based on the theory, for designing color sequences to be effective in both conveying form and value information. An experimental color sequence is presented to illustrate these guidelines.

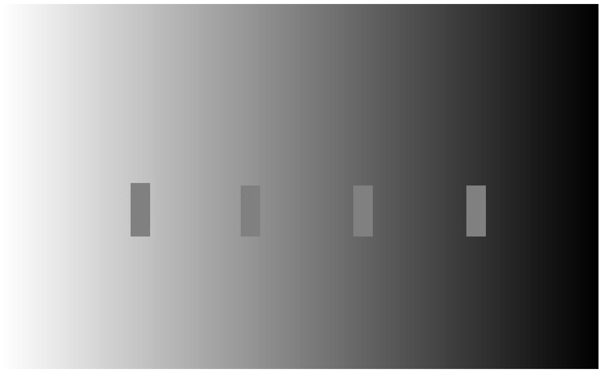


Figure 1. All the gray rectangles are the same lightness. The effect of simultaneous contrast is to make the ones on the light background seem darker than the ones on the dark background.

Univariate maps-such as X-radiographs, astronomical radiation charts, or geographical maps in which data is available over a continuous plane-are often pseudo-colored as a method for revealing aspects of the data. In some cases convention determines the choice of a color scale and the assignment of colors to data values. Geographers have a well-defined scale to display height above sea level; lowlands are always green, which evokes images of lush vegetation, and the scale continues through brown to white at the peaks of mountains. However, in many instances there is no conventional reason for choosing one color sequence over another, as is the case for a map that illustrates the earth's magnetic field. In these situations pseudocoloring is done on an ad hoc basis.

Display systems that have frame buffers and color lookup tables make it easy to pseudocolor maps. Color sequences can be changed merely by changing the entries in the look-up table, which in many systems can be done between frames. This article addresses the problem of how to design a color sequence from a theoretical/experimental perspective.

To evaluate color scales, we must first answer a significant meta-question, which is: What kinds of information do we want to be revealed? Do we want to know if a galaxy has a spiral form or a radiograph has an abrupt change in density denoting a fracture line? Alternatively, do we want to know how much energy is being emitted at a specific point in the galaxy, or the absolute tissue density as measured by X rays through the spleen? There are two distinct and qualitatively different kinds of information implied by these questions.

Metric information consists of measured data quantities for points on the surface. The features of interest are such things as the height of a peak or the energy level at location X. Form information is the shape of the surface. The features of interest are such things as local maxima, valleys, and cusps marking gradient discontinuities.

Theory and predictions

From the point of view of obtaining metric information from a univariate - map, the main issue is the accuracy with which readings can be taken. That large errors can occur in reading map quantities using a key has been well documented.[1-3] The main issue addressed here is the cause of that error.

In a preliminary version of this article' the terms used for this distinction were "value" and "form." Unfortunately "value" was found to be confusing because it is also a technical term in color theory, so the term "metric" has been used here.

Several studies have concentrated on the construction of scales that are equally spaced in terms of just noticeable differences (jnds) and which, in consequence, have uniform resolutions along their lengths.[4,5] However, it seems likely that errors due to resolution problems are minor compared to the much larger and systematic errors likely to arise from effects in the visual system such as simultaneous contrast.

Simultaneous contrast is the effect when the color of a patch is shifted perceptually by the color of adjacent patches. Thus a gray patch on a white ground will be perceived as darker than an identical gray patch on a black ground (see Figure 1). Chromatic contrast also occurs: A gray patch on a red ground will be perceived as greener than an identical gray patch on a green ground. The phenomenon of simultaneous contrast is thought to be produced by changes in the balance of the cone receptors and by changes in cortical opponent processing channels.[6,7]

The existence of these cortical color-processing channels is now widely accepted, although the details are still controversial. According to the canonical theory there are three opponent channels. The achromatic channel outputs the sum of the long- and medium-wavelength sensitive cones (this is sensitive to luminance information): One of the chromatic channels outputs the difference of the long- and medium wavelength sensitive cones (this is sensitive, roughly, to red-green differences). The other chromatic channel outputs the difference of the achromatic channel and the short-wavelength cone signals (this is sensitive, roughly, to yellow blue differences). The magnitude of the simultaneous contrast effects appears to be roughly comparable in each of the three opponent channels.[8]

When the color of a certain point on a map is to be read, the regions surrounding that point will influence the perceived color. If a gray point is surrounded by red, it will be perceived as tinted green. Such contrast effects are strongest where smooth gradients of color are present,[8] exactly the situation found in many pseudocolored continuous tone maps.

If, as hypothesized, errors are introduced by simultaneous contrast in color-coded maps, then these errors should occur for both gray-scale maps and chromatically coded maps. Also, certain color sequences will be more liable to contrast-induced errors than others. A color gradient that increases monotonically with the data and with a color opponent-channel − an example is gray-scale coding − will tend to have large errors due to contrast; because the surrounding regions will exert a concerted influence. Also, a saturation scale, for example, from gray to red, varies monotonically with both the yellow blue and the red-green opponent channels and can therefore be expected to have large errors associated with it. Another scale in which a large contrast effect can be expected is a red-to-green sequence, which will vary monotonically with the red-green opponent channel.

As an example of a color scale that should be resistant to contrast distortion, consider a scale based on an approximation to the physical spectrum, obtained using a color monitor (see Figure 2). This scale does not vary monotonically in any of the three opponent channels. It has weak variation in the achromatic channel and roughly sinusoidal variation in both the red-green and yellow blue opponent channels. Contrast effects should be weaker using this sequence for two reasons:

The surround to a given test patch is less likely to weight the opponent channels systematically in a particular direction.

When a strong contrast effect is registered in one channel, a weak effect will be registered in another channel.

Users might adopt perceptual strategies which enable them to pay attention to the channel that provides more veridical information. This would further reduce the error for this kind of color sequence. There is already some supporting evidence for the above theory. Heath and Flavell[3] reported substantially smaller errors for the spectrum scale than for other scales in an empirical investigation of the errors made while reading colorcoded maps. Here, the possibility that this is due to simultaneous contrast is explicitly tested.

Experiment 1: The effects of simultaneous contrast on reading quantities from maps

The first experiment was designed to measure the susceptibility of five different color sequences to distortion by simultaneous contrast effects. It was also designed to test the specific prediction that a sequence based on the physical spectrum would be less susceptible to simultaneous contrast effects than one based on luminance. To these ends the stimulus pattern was designed to induce large contrast effects.

Stimuli

The stimulus configuration is shown in Figure 3. A parabolic surface was used to simulate a local maximum in a map surface, and a small circular disc was placed at its center as a test patch to measure reading errors due to contrast. A key was placed on the right-hand side of the map consisting of 16 equally spaced samples selected from the color sequence. The key had the alphabetic characters A, B, C...P as labels for each of the sample colors.

The center of the parabolic surface was at the extreme end of whatever color sequence was in use, while the test patch was selected from the set of the middle 12 elements of the 16-step color key.

With this pattern it is possible to make' qualitative predictions about the direction of contrast effects based on opponent process theory. If the test patch and the surrounding colors vary monotonically with respect to an opponent channel, the effects of contrast will be to move the apparent color of the test patch away from the color of the surrounding area in a direction defined by the channel.[8] This should be the case for all of the patterns except for the spectrum approximation.

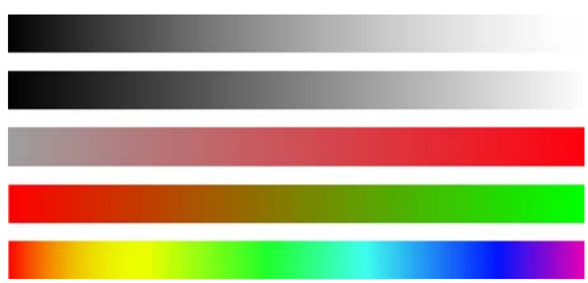


Figure 2. The five color sequences used in experiments 1, 2, and 3.

Any error in the observer's selections can be attributed either to random error due to an inability of the observer to make fine discriminations or to the influence of simultaneous color or brightness contrast. Systematic errors, always in the same direction, can be attributed to simultaneous contrast.

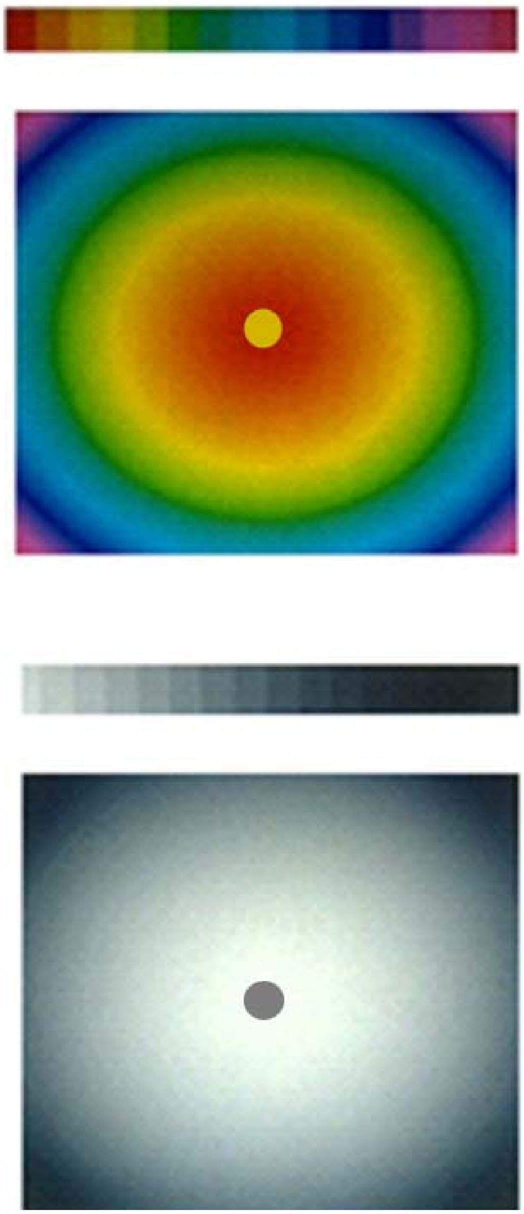


Figure 3. The stimulus pattern used for experiment 1.

Color sequences

Five different color sequences were evaluated. These were chosen either because they are in wide use or because they have theoretical interest. They are shown in Figure 2.

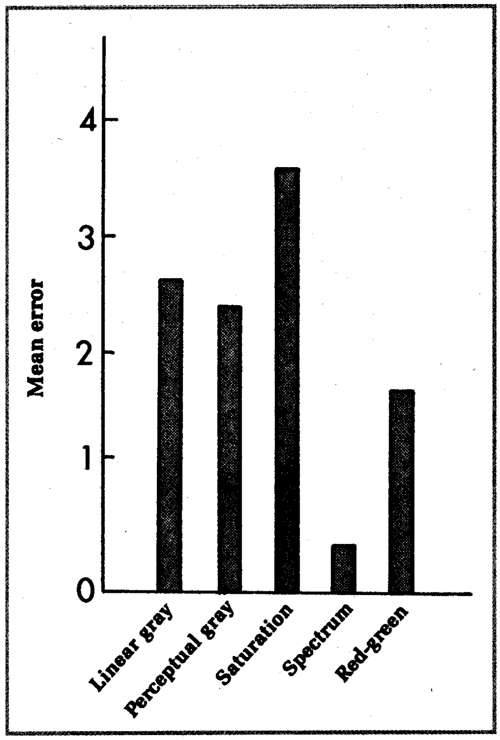


Figure 4. Mean errors obtained for each of the five sequences shown in Figure 2.

Linear gray sequence

A gray sequence was constructed by interpolating equally spaced luminance steps between the brightest white available on the monitor and the darkest black.

Perceptual gray sequence

It is well known that physically equal steps do not produce perceptually equal gray steps,[9] although a recent paper showed that a physically linear scale may be optimal for detection purposes.[10] For a gray scale that was closer to perceptual uniformity, a scale that uses the CIELuv luminance scaling function was created."

$L^* = 116(Y/Y_n)^{1/3} - 16$

L^* is perceived brightness, Y is the luminance, and Y_n is the luminance of a reference white. The value of the reference white chosen for the present study was the maximum monitor white.

Saturation sequence

This sequence consisted of a linear interpolation between a gray and a red. The gray was produced by setting the red, green, and blue phosphors at half their

their maximum values, and the red was the maximum of the red phosphor alone.

Spectrum approximation

A spectrum approximation was achieved by linearly interpolating the following sequence of colors: red +blue, blue, blue+green, green, green+red, red: Blue, green, and red denote the maximum output of the blue, green, and red phosphors respectively. This is a widely used spectrum approximation scheme.

Red-green sequence

The selection of this sequence was motivated by the opponent processing theory of human color vision.[7] Although the red-green gradient produced by red to green phosphor interpolation is not perfectly aligned to affect only the red-green visual channel, it will have its principal effects in stimulating that channel.

All of the above color sequences were encoded as 255 color steps. These steps were sufficiently small that the surfaces appeared as smoothly changing gradients of color. Also, gamma correction was used throughout to correct for the nonlinearities between the quantities stored in the frame buffer and the amount of light produced by each of the monitor phosphors."

Experimental procedure

On an experimental trial the central disc was given a color taken at random from the middle 12 of the set of 16 colors provided in the key. The observer's task was to select the key color that most closely approximated the appearance of that central disc. Each subject was tested using all five color sequences with two trials for each of the 12 test colors. The order of color sequences and of trials was randomized: Ten observers were used, all color normals according to Ishihara pseudoisochromatic plates.

Results from experiment 1

The mean error obtained with each of the color sequences is displayed in Figure 4. Multiple tests show that the spectrum sequence produces significantly more accurate readings than any of the other sequences ($p < 0.01$ for each comparison). In fact the error mean obtained with this sequence is less than a third of that obtained with any other sequence. The next best sequence is red-green, which was significantly better than saturation ($p<0.01$) and the linear gray scale ($p < 0.05$), but not significantly different from the perceptual gray scale. No other statistical differences were found.

The errors that occurred were not random but were in the direction predicted by simultaneous contrast. This is shown in Figure 5, which plots the mean direction' of error for the linear gray scale and for the spectrum scale. It can be seen that in the case of the gray scale the errors are consistently in a direction away from the color of the center of the map. The small errors in the spectrum scale contain a pattern markedly different from that achieved with the other patterns. They are not always "away from" the inducing color when plotted on a straight line continuum. However, opponent color theory predicts that the contrast effects will occur through lateral inhibitory effects in the color opponent channels, and when the results for the spectrum scale are plotted on a chromaticity diagram with the principal axes of the red-green and yellow blue channels, a more consistent picture occurs (see Figure 6). Unfortunately, contrast theory is not yet sufficiently advanced to make detailed predictions for such a complex stimulus,[6,8] and a full treatment is beyond the scope of this article. Nevertheless, the bones of the canonical theory state that the apparent color of the test spot should be shifted in both red-green and yellow-blue channels away from the inducing regions surrounding it, and this is what is seen for the most part.

The data as a whole can therefore be taken as support for the hypothesis that a major source of error in reading color-coded map data is simultaneous contrast. Also, the success of the spectrum scale in reducing error can be attributed to the fact that this color scale does not vary monotonically with any of the opponent channels.

Color and form perception

In general, the form information displayed in a univariate map is far more important than the metric information. Absolute quantities are well represented in a table, whereas maps gain their utility from their ability to display the ridges and valleys, cusps, and other features.[7] The crucial theory, which is relevant to form perception using color sequences, is again the opponent process theory.[9] The luminance channel appears to process shape,[13] motion,[14] and stereoscopic depth[15] far more effectively than the chromatic channels. It seems plausible that the luminance channel has a major function of abstracting the form of a surface in the environment while the chromatic channels are designed primarily to process the material properties of surfaces. Thus we may predict that a luminance gradient will be better at revealing the shape or form of a surface than a gradient that changes only in color.

Experiment 2: Perceiving form

To evaluate color sequences in their capacity to express the form of a surface, we run up against a meta-question once again. What aspects of the form do we wish to be revealed? Form is not a one-dimensional concept. Thus one color sequence might be good at revealing local maxima in a surface, while another might be good for detecting surface discontinuities. Therefore a taxonomy of the elements of surface form is needed to assess color scales.

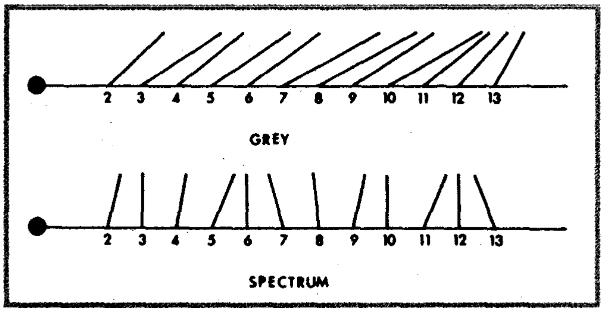


Figure 5. The direction of error obtained for the linear gray sequence and for the spectrum sequence. The colors of the test patches are given by the bases of the diagonal line segments. The horizontal position of the top of each segment indicates the judged color. The circle denotes the color immediately surrounding the test patch. (This work was done by J. Krauskopf, D.R. Williams, and D.W. Heeley, in "Cardinal Directions of Color Space," from Vision Research, Vol. 22, No. 9, pp. 1123-1131.)

To provide such a tool, surfaces were classified into five abstract categories: gradient, ridge, convexity/concavity, saddle, and cusp. For each of these abstract categories a specific instance was constructed using simple mathematical functions. These are listed below and illustrated in Figure 7.

Gradient. A plane ramp in the x direction was chosen as a basic gradient.

Ridges. A single cycle of a sinusoid was chosen to represent ridges and valleys.

Convexity/Concavity. A parabolic surface was used to represent convexity/concavity.

Saddle. The saddle surface was parabolic in the x direction and parabolic in the y direction with the curve having an opposite sign for the two directions.

Cusp. The cusp surface was formed using two exponential functions joined so that they formed a gradient discontinuity, or cusp, at the line of intersection. The same five color sequences described for experiment 1 were used in experiment 2, and for all of the surface shapes the surface was scaled to use the entire range of the color sequence.

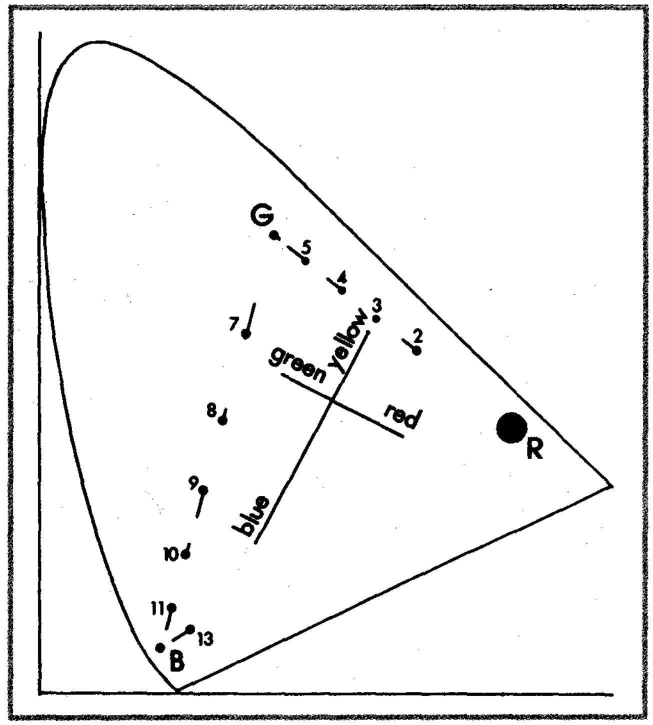


Figure 6. The direction of errors for the spectrum sequence plotted on a CIE chromaticity diagram. The axes of the red-green and yellow-blue opponent channels are also shown from Krauskopf.

Procedure

The 10 color-normal subjects were instructed that the purpose of the experiment was to evaluate various color coding schemes according to how well they revealed the features of surfaces. The experimental protocol consisted of showing each observer one surface at a time, in one of the color-coding schemes, together with a written description of that surface. The observer was then allowed to cycle through the remaining four color-coding schemes. After at least one look at all five schemes, the subject was asked to rate their effectiveness using a five point scale. The specific question the subject was expected to answer was: "flow effective is this color sequence in revealing the important properties of this surface?"

0	1	2	3	4
poor		average		good

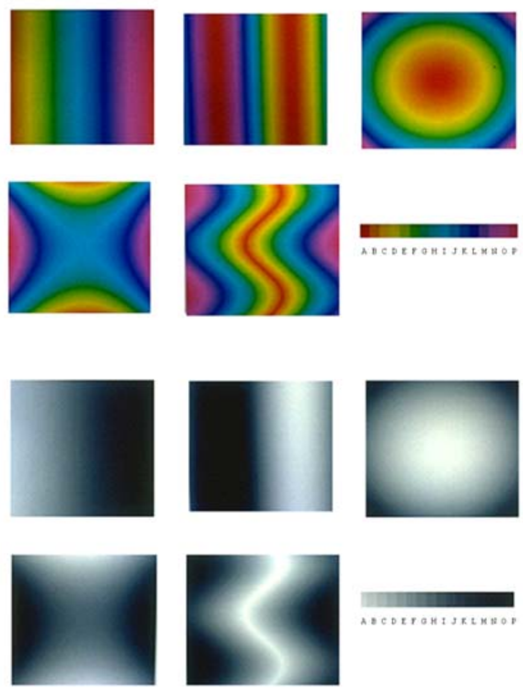


Figure 7. Five surfaces used in experiment 2. These are shown using (a) the linear gray sequence and (b) the spectrum approximation.

Results of experiment 2

Mean ratings from 10 observers are given in Figure 8. The most significant effects are as follows: For the "gradient" and the "ridges" patterns the gray scales were judged to be significantly better than the chromatic sequences. This result can be confirmed by the reader by a comparison of Figures 7a and 7b. The spectrum approximation gives the impression of a set of colored stripes without any obvious form, whereas the gray scales convey the ramp and sinusoidal forms unambiguously. For the other three surfaces the main effect was that lie saturation scale was judged to be the least useful. For these scales there were no significant differences between the judged utility of the spectrum sequences and the judged utility of either gray scale: The linear gray scale was judged to be more effective than the perceptual gray scale for both the "convexity/concavity" surface and the "cusp" surface. Apparently in these instances the linear gray scale was better at showing the more significant features of the surfaces.

Designing color sequences

Assume that it is possible to minimize contrast -induced errors by using carefully designed color sequences, and suppose that an achromatic scale is likely to be the best for reading form information. Although the results of experiment 2 have by no means established this unequivocally, let us, for the moment, suppose that it is true. Then we have the possibility of designing a color sequence that is good at conveying both form and metric information.

The idea is to spiral up through color space and thus produce a scale that does vary monotonically on the luminance channel, but which does not produce large systematic distorting effects on the color channels. This assumes that observers can selectively attend to luminance information for the shape of the surface and to the chromatic channels for quantitative readings-an assumption that is plausible but untested.

Experiment 3 was designed to test an experimental color sequence designed to be good at revealing both form and metric information. The color sequence was designed to increase monotonically in luminance while passing through a sequence of colors that do not vary monotonically with either of the two chromatic opponent channels. This sequence is illustrated in Figure 9. The experiment was also designed to test the other five scales using a sample "real" univariate map.

Experiment 3: An experimental color sequence and a map

This experiment used as stimuli the univariate map shown in Figure 9 displayed using six different color sequences. As an aside, this map shows a section of the North Atlantic. The data represent differences between readings of the height of the ocean's surface measured by satellite altimetry and the theoretical smooth surface (geoid) which is used to represent the earth's gravitational field.

This experiment was like experiment 1 in that observers were asked to read map quantities using a key. The 16 locations on the map used for test patches were located in a grid covering a section of the ocean. On a particular trial a cursor was presented at one of these locations and the subject was asked to enter the closest value from the color scale. Twelve observers were tested, five of whom were surveying engineering graduates and undergraduate students, and the remainder were computer science undergraduates. All had normal color vision. The order of the color sequences and the order of the test regions were randomized.

Results of experiment 3

The mean errors for the six color sequences are given in Figure 10. The errors for both the "spectrum" pattern and the "experimental" sequence were significantly lower than all others, but not significantly different from each other. No differences' were found between surveying engineering students and the others. The errors for both of the gray scales; the red-green scale, and the saturation scale did not differ significantly from one another. This is important because it suggests that the errors caused by contrast are fundamental to perception and are not likely to be influenced by experience with maps. It is worth noting that the magnitude of the error obtained for all conditions is amplified by the discrete nature of the key. The test colors were not exactly matched by the key colors (in experiment 1 they were), and the average error simply from discretization to the 16-color scale is 0.25 key steps. This is marked on Figure 10.

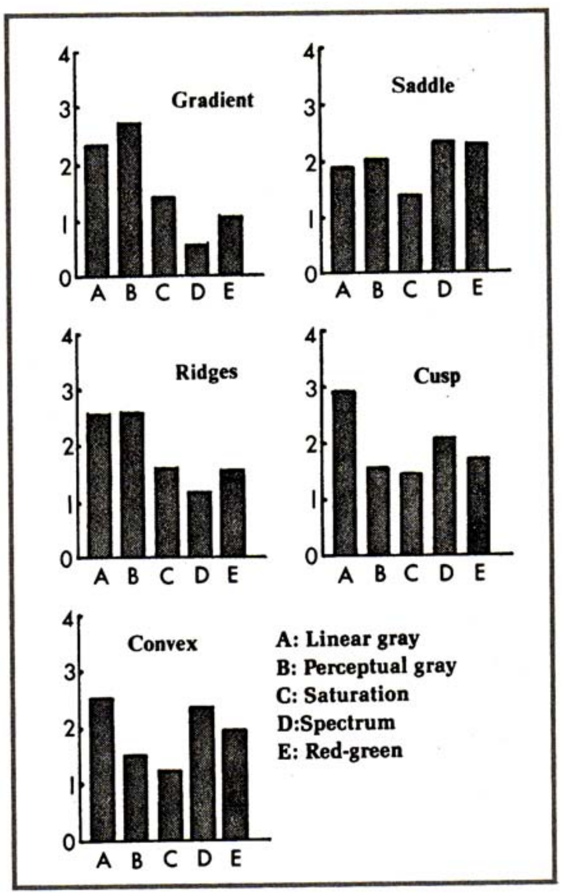


Figure 8. Mean ratings of how well each color sequence conveyed an impression of each surface.

Discussion

The success of the "experimental" sequence in reducing contrast effects suggests that an observer can somehow disassociate the chromatic information and the achromatic information in reading quantities using a key. Because the achromatic information varied monotonically with the map data, achromatic distortions should have been as large for this display as for the gray sequence displays. However, the results show much smaller distortions, indicating that the achromatic information was ignored. The usefulness of the "experimental" color sequence in displaying the form of the map data was not empirically evaluated, although a number of observers have judged it better than the other chromatic sequences under informal viewing conditions. Viewing Figure 9 will convince the reader that it shares the advantage of the achromatic scales in that maxima and minima are easy to locate as compared to the spectrum version, for example.

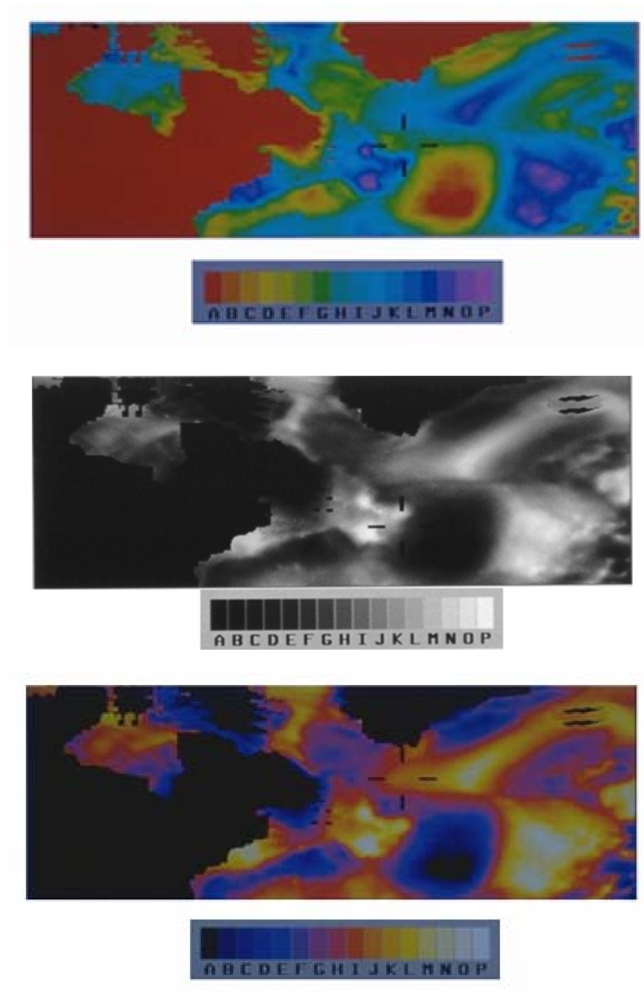


Figure 9. The map used as a stimulus pattern for experiment 3 is shown with three of the six color sequences. Subjects were instructed to attempt to match the color at the exact center of the cross hairs using the key provided.

There is no doubt that the best method for choosing color sequences for a particular example of map data should involve some interactive tool whereby the color scale can be customized to the particular map. The example of the cusp in Figure 7 illustrates this point. To make the cusp salient, it is necessary to create a sharp perceptual gradient for the particular metric range wherein the cusp lies. It so happened that the linear gray sequence made this cusp more salient than the perceptual gray sequence because it emphasized the region of abrupt transition. Nevertheless, before a color sequence is interactively modified, some starting sequence must be selected or constructed.

The work presented in this article is intended to help provide rules of thumb for the construction of stock color sequences, which may be useful in a variety of contexts. Some reasonable rules of thumb follow:

1. If you wish to read metric quantities using a color key, then a sequence that does not vary monotonically with the color opponent channels should be used. A good example is a spectrum approximation.
2. If revealing the shape or form of the surface is important, and features such as parallel ridges are present, then do not use a spectrum approximation or any other purely chromatic sequence. A sequence that varies monotonically with luminance is likely to be effective.
3. To create a color sequence that has good properties for revealing both shape and metric quantities, use a sequence that increases monotonically in luminance, while cycling through a range of hues. The hues provide accurate readings from a key, while the luminance conveys the form of the surface.

Acknowledgments

This research was supported in part by grant #A8202 from the Natural Sciences and Engineering Research Council of Canada. I am indebted to David Wells for suggesting the term "metric" used throughout this article. Thanks are also due to Rosdam Yasdani for the geoid data used in the third study.

References

1. C. Ware, "Choosing a Colour Sequence for Univariate Maps." Proc.IEEE Conf. on Systems, Man and Cybernetics, Vol. 1, New York, Oct. 1986, pp.41-45.
2. W.S. Cleavland and R.A. McGill, ' A Color-Caused Optical Illusion on a Statistical Graph," American Statistician, Vol. 37, No. 2, May 1983, pp. 101-105.
3. Heath and R.B. Flavell, "Colour Coding Scales and Computer Graphics," Proc. Graphics Interface 85, CMCCS, Toronto, pp. 321-328.
4. E.K. Robertson, and J.E O'Callaghan, "The Generation of Color Sequences for Univariate and Bivariate Mapping," CG&A, Vo1.6, No. 2, Feb. 1986; pp. 24-32.
5. J: Tajima, "Uniform Color Scale Applications to Computer Graphics," Computer Vision, Graphics and Image Processing, Vol. 21, No. 3, 1983, pp: 305-325.
6. C. Ware, and W.B. Cowan, "Changes in Perceived Colour due to Chromatic Interactions in Striped Test Fields," Vision Research, Vol. 22, No. 11, 1982, pp. -1353-1363.
7. L.M. Hurvich, Color Vison, Sinauer Associates, Sunderland, Mass., 1981, 328 pp.
8. C. Ware and W.B. Cowan, "Chromatic Mach Bands: Behavioural Evidence for Lateral Inhibition in Color Vision, Perception and Psychophysics, Vol. 41, No. 2, 1987, pp. 173-178.
9. G. Wyszecki and WS. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, John Wiley & Sons, 2nd edition, New York, 1982.
10. F Kingdom and G. Moulden, "Digitized Images: What Type of Gray Scale Should One Use?" Perception, Vol. 15, No.1,1986, pp. 17-25.
11. Commission Internationale de L'Eclairage (CIE), "Recommendations on Uniform Color Spaces, Supplement N2 to CIE Publication N15," Paris, (E-1.3.1),1971.
12. W.B. Cowan, "An Inexpensive Scheme for Calibration of a Colour Monitor in Terms of CIE Standard Coordinates," Computer Graphics (Proc. SIGGRAPH), Vol. 17, No. 3, July 1983, pp. 315-321.
13. K.Y. Mullen, "The Contrast Sensitivity of Human Color Vision to Red-Green and Blue-Yellow Chromatic Gratings," J. Physiology, 359, 1985, pp. 381-400.
14. S.M. Anstis and P Cavanaugh, A Minimum Motion Technique for Judging Equiluminance in Colour Vision: Physiology and Psy-

chophysics, J.D. Mollon and LT Sharpe, eds., Academic Press, London, 1983, pp. 156-166.

15. R.L. Gregory, "Vision with Isoluminant Color Contrast: 1 A Projection Technique and Observations;" Perception, Vol. 6, No. 1, pp. 113119

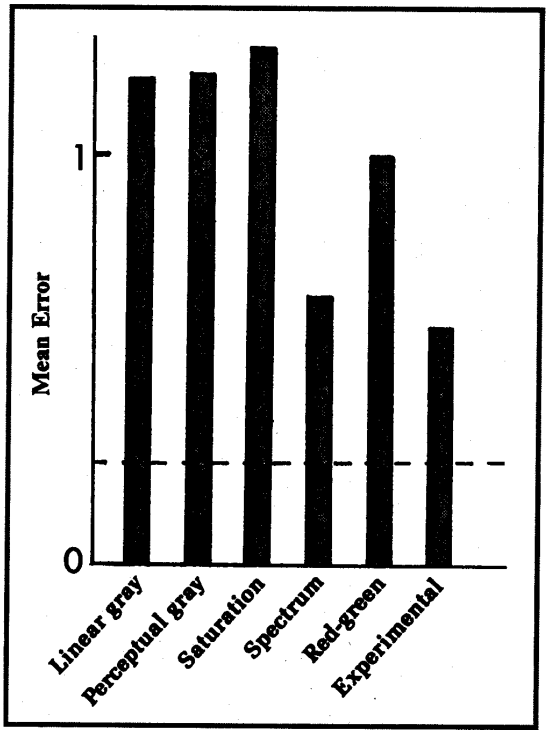
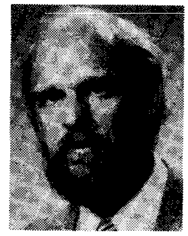


Figure 10. Mean errors obtained with each of the six color sequences used in experiment 3.

Colin Ware is an assistant professor of computer science at the University of New Brunswick, Canada. His research interests include the use of color for information display, techniques for visualizing scientific data, and an investigation of 3D manipulation using computer graphics display systems. Before joining the University of New Brunswick, he was a research associate at the National Research Council of Canada, where he worked on basic color theory and applications in color standards.

Ware received a PhD in experimental psychology from the University of Toronto and an MMath in computer science from the University of Waterloo.



Colin Ware can be reached at the School of Computer Science, University of New Brunswick, Box 440, Fredericton, N.B., Canada E3B 5A3; or at cware@ccom.unh.edu.