Which Trajectories through which Perceptually Uniform Color Spaces Produce Appropriate Colors Scales for Interval Data?

***Bernice E. Rogowitz*[*], *Alan D. Kalvin*[*], *Adar Pelah*[+], *Aron Cohen*[*]**
[*]*IBM T.J. Watson Research Center, Yorktown Heights, NY USA*
[+]*Cambridge University, Cambridge, United Kingdom*

Abstract

This paper is concerned with the development of perceptual colormaps for visualization. Colormaps are typically designed using algorithmic techniques and do not explicitly take into account the way in which the human visual system processes its underlying components. The typical colormap (the "rainbow" colormap), for example, maps data onto a scale of colors which varies in hue, brightness and perceived saturation over its range. In this paper we experimentally measure how hue, luminance and saturation scales represent magnitude information by constructing colormaps which trace carefully controlled paths through Munsell and L*a*b* color spaces, and comparing these results with colormaps designed using standard computer graphics methods.

Introduction

The purpose of visualization is to create visual representations that effectively communicate the structure of the data without creating visual artifacts. A fundamental characteristic is the magnitude of the data values. If the data are measured, observed, or computed on an interval scale, for example, then the difference between two data values represents the same difference, no matter where on the scale that difference is measured. For example, a change in temperature from 40 to 80 degrees Farenheit is the same magnitude as a change in temperature from 122 to 162 degrees Farenheit; a change in heart rate from 78 bps to 82 bps is the same as a chage from 100 to 104 bps. An important task in visualizing such data is to preserve this magnitude information. When representing the data values, equal steps in data value should correspond to equal perceptual steps. In addition, the sign of the change should also be preserved in the visual representation.

In visualization, a common method for representing interval data is to use a colormap. A colormap is a scale of colors which is mapped onto a range of data values to create a visual representation. Some common examples include the gray-scale colormap used to represent X-rays and the "rainbow" colormap used to represent temperatures in a weather map. Colormaps can be computed and constructed by a number of different methods. They are most commonly expressed in terms of (R,G,B) display values, although they can also be expressed in a space which transforms these values to achieve a more perceptually meaningful set of three dimensions (e.g., double-cone, HLS (hue, lightness, saturation), CIE L*a*b*, or Munsell).

Here we introduce a method to to test the suitability of color sequences for representing interval data. In particular we examine the use of the three perceptual dimensions of color (1) perceived hue, (2) perceived saturation, and (3) perceived brightness in the construction of perceptually uniform colormaps, colormaps in which equal steps in data value correspond to equal perceptual steps.

The perceptual suitability of a given colormap is evaluated psychophysically, using an increment-threshold technique. For 20 steps along each colormap, we measure the amount that colorscale value must be changed in order for a human observer to reliably tell that an increment has occured. That is, we are measureing the just-noticeable perceptual change in magnitude along the colormap, taken at selected values of the colormap. If the necessary increment is constant for all values in the colormap, then that colormap preserves data intervals accurately in perception.

Perceptual Colormaps

The study of perceptually-based colormaps has two roots in the psychophysical literature. Papers by Robertson (1988) and Pizer (1981), for example, base their work on the notion of creating just noticeable differences between the steps in a colormap. Phil Robertson suggests using a blue-yellow trajectory along an isoluminant plane of L*a*b* space because that is the axis which provides the longest span between colors, assuming that L*a*b* space is, itself, perceptually uniform. Pizer recommends a black-body radiator colormap. He creates a spiral trajectory, going from dark to light, beginning at dark red, moving up in angstrom units (down in wavelength) through yellow, orange, and blue, and ending in white. Since the path through the colorspace is not a straight-line, more just noticeable differences can be spanned.

Other papers have based the development of perceptual colormaps on the psychophysical literature on perceptual scaling. This notion is discussed by Green (1992) and has been more fully elaborated and developed by Rogowitz and Treinish (1996, 1998). This work is based on the seminal work of S.S. Stevens (1966), who measured the perceived magnitude of different magnitudes of physical stimuli. He found that for many physical scales, including luminance (cd/m^2) and saturation (the "redness" of a long-wavelength light source), equal ratios in stimulus value produced equal ratios in perceptual magnitude. He interpreted this as indicating that there exists in human cognition a common scale for representing magnitude, and we scale the effects of different physical stimuli to this internal scale. The slope of the scaling function is parameterized by the individual physical stimulus. For example, the brightness function has a much more shallow slope than the slope of the function measuring the perceived magnitude of an electrical shock.

Stevens also also wrote that, unlike luminance and "redness" (saturation), where equal steps in physical magnitude correspond to equal steps in perceived magnitude, hue was not such a scale. This suggests that luminance and saturation scales might be appropriate for communicating, or “carrying” (as Fred Brooks would say) magnitude information, while hue-based scales would not.

Since these ideas have never been tested experimentally, we have begun a series of experiments. In the experiments reported here, we measure just noticeable differences of increments in colormaps which have been constructed as trajectories through various colorspaces. We report data for colormaps which represenent paths in luminance, saturation, and hue. We have collected these data for three different color spaces. The “double-cone” HLS colorspace which transforms RGB values algorithmically to produce variations in hue, saturation and value (Foley *et al.* 1991), the CIE L*a*b* space, which is widely considered to be "perceptually uniform”, and the Munsell space which is widely thought to be the most perceptually uniform space.

Psychophysical Increment Threshold Measurements

Consider how increment thresholds might be expected to vary on colormaps contructed independently along the dimensions of hue, saturation and brightness (perceived luminance). Wavelength discrimination experiments measure the variation in wavelength (lambda) which is just discriminably different from a standard and plot this difference (delta lambda) as a function of the reference wavelength. Wavelength discrimination curves for color-normal observers show a wavy pattern of sensitivity (delta lambda), with regions of low sensitivity (large deltas) particularly at low, middle and high wavelengths (Wright & Pitt 1934). Wavelength discrimination curves also show differences between observers, but reliable averages may be computed. Luminance increment detection experiments measure the difference between a target and background which is just detectable. Increment thresholds for luminance, measured below a saturating high background level and above photon and internal noise-limited level (Barlow 1957), will generally obey Weber's Law (deltaI/I is constant). That is, the increment which is reliably discriminable is a constant fraction of the background luminance. Luminnce increment threshold curves tend to be very reliable. Saturation increment threshold curves can be extracted from McAdam ellipses, but we are not aware at this time of any saturation increment threshold studies, possibly because while brightness and hue have perceptual correlates in physically-instantiable visual stimuli (e.g., luminance and wavelength), saturation is a derived measure.

By saying that brightness and hue are correlates of the physical stimulus, we mean that there is not a linear relationship between the value of the stimulus and its perception. Perceived luminance, for example, is a non-linear function of cd/m^2, and wavelength, although related to hue, also varies in brightness and perceived saturation. One goal of perceptual color spaces is to create a space where the effects of hue, luminance and saturation can be represented orthogonally, with equal distances in the space corresponding to equal perceptual differences. Among the most influential of these spaces are the computer-graphics HLS space, the CIE L*a*b* space, and the Munsell colorspace.

Experimental Method

In this paper we measure how hue, luminance and saturation scales represent magnitude information by constructing colormaps which trace carefully-controlled paths through Munsell and L*a*b* color spaces, and compare these results with standard computer graphics colormaps designed to provide hue, saturation and luminance variations.

We conducted increment threshold experiments using a modified method-of-adjustment procedure, described below. On each trial of the experiment, we measured the detectability of a spatially-Gaussian target on a range of background levels selected randomly from a given color map. The target's maximum luminance is at the centerpoint and decreases in a standard Gaussian manner until reaching the background level. The value at each point in the target changes smoothly along the scale defined by the colormap. For each of 20 background levels, the just noticeable amplitude of this Gaussian was measured. This same method is used to construct stimuli from each of the color scales. In all cases, the target also increases and decreases in a Gaussian manner over the scale, and the amplitude producing a just noticeable difference is measured. For the hue-varying maps, it is the hue of the target, as measured in each scale, which varies in a Gaussian manner. A low-amplitude hue target on a blue background, thus, would appear blue at the edges and perhaps somewhat cyan in the center; a higher amplitude target on the same blue background would appear blue at the edges then cyan and perhaps green in the center. In our experiments, we measured the detectability of Gaussian increments in colormaps which varied in luminance, hue, or saturation, although this procedure could be applied to a color scale describing any trajectory through a colorspace.

The display screen was viewed at 50 cm, at which distance the background subtended a 12 x 6 degree rectangular patch containing a central fixation mark. The circular target with its Gaussian profile appeared for 100 msec to either the left or the right of fixation, at 3 degrees eccentric to fixation. The subject judged whether the target was visible, and adjusted its amplitude accordingly by pressing the buttons on a standard computer mouse, and by using the space-bar key to indicate that the target was at threshold. This also served to advance to the to the next trial, where the increment threshold was measured on another randomly-selected background. The use of the method-of-adjustment in psychophysical experiments has sometimes been criticized for potentially introducing bias. However, systematic comparisons between method-of-adjustment and so-called forced-choice methods have generally found no important differences in data patterns (e.g. Legge 1981). Although quantitative differences may be introduced by the use of different procedures, the relationships between the measured thresholds is generally found to be quite stable (Kelly & Savoie 1973; Graham 1989). At any rate, experiments of the kind reported here, where a great amount of data needs to be collected from individual observers, are largely impractical using other, more time-consuming methods. The modification of random left/right alternation in target position improved on the standard method of adjustment (where target position would normally remain spatially fixed) by heightening the attention level of the subject, and reducing local retinal adaptation effects.

Color maps were tested which varied in luminance, saturation or hue, constructed in three diffferent colorspaces, 1) uncalibrated RGB, 2) CIE L*a*b* and in 3) Munsell.

Creation of the Colormaps

The Munsell- and L*a*b*-based colormaps were created with a program we developed called ***ColorBilder*** that runs under Microsoft Windows 95/NT. The program was built using the Microsoft Visual Basic programming language, together with the CGSD Color Science Library package.

For both the Munsell and L*a*b* colorspaces, we created three colors maps: (1) a hue-varying colormap, (2) a luminance-varying colormap, and (3) a saturation-varying color map. (See Table 1.)

Table 1. The 8 Color Maps Tested

MUNSELL Hue-varying	
HUE	*0, …, 80 (circular path, counter-clockwise)*
VALUE; CHROMA	*7; 7*
MUNSELL: Luminance-varying	
HUE	*0*
VALUE; CHROMA	*0, …, 10; 0*
MUNSELL Saturation-varying	
HUE	*0*
VALUE; CHROMA	*10; 0, …, 20*
L*a*b* Hue-varying	
*L**	*72*
a*	*13, …, 28 (straight-line path)*
*b**	*-26, …, -78 (straight-line path*
L*a*b* Luminance-varying	
*L**	*0, …, 100 (straight-line path)*
a*	*0*
*b**	*0*
L*a*b* Saturation-varying	
*L**	*100*
a*	*0, …, 103 (straight-line path)*
*b**	*0, …, 27 (straight-line path)*
STANDARD Hue-varying (HLS rainbow)	
H	*0, …, 270 (circular path, clockwise)*
L	*1*
S	*1*
STANDARD Luminance-varying (HLS gray-scale)	
H	*0*
L	*0, …, 1 (straight-line path)*
S	*1*

We used the ColorMap Editor module of the IBM Visualization Data Explorer software package to create standard computer graphics colormaps, a hue (rainbow) colormap, and a luminance (gray-scale) colormap.

We displayed the colormaps on a 21-inch IBM P202 color CRT monitor attached to an IBM Intellistation M Pro. We calibrated the monitor using a Minolta CL-100 Chroma Meter attached to a Minolta CA-100 Color Analyzer.

These colormaps were converted to display-dependent RGB values using the PLVC method (Piecewise Linear interpolation assuming Variable Chromaticity coordinates) (Post and Calhoun, 1989).

Results

In this section, we show preliminary results from three observers. Each observer collected a single increment threshold curve for each of the 8 colormaps tested. Each curve consisted of twenty method-of-adjustment trials, one at each of 20 randomly-presented levels in the colormap. Error bars indicating one standard error are shown. The colormap spanned 100 units, ranging from the lowest value in the colormap (0) to the highest (100). The increment required for detection is also measured in these units. Since the colormaps are calibrated, we can relate these measures back to characteristics of the physical stimulus and of the color spaces in which they are specified for future analysis.

Figures 1 and 2 show increment threshold curves for the two uncalibrated, computer graphics colormaps. The luminance colormap data are shown on the top; the rainbow hue colormap data are shown on the bottom. Consider first the luminance colormap. Over a significant range, a constant increment of 0.4 steps is required for increment detection. At the low luminance end of the range (to the left) however, sensitivity decreases by a factor of five, peaking at 2.0. There is also a small decrease in sensitivity at the high end of the range, suggesting that the requisite increment might be a constant proportion, not a constant increment, of the background luminance.

The shape of the data for the rainbow colormap is significantly different. For several regions, a constant increment is all that is needed for detection. This value varies between 0.4 and 1.0. In the remaining regions, however, most prominently in the central region, the magnitude of the requisite increment can be as much as 11 times the background level. This has a significant impact on how the data will appear. If the background is 40 and the increment threshold is 12, for example, all the data values mapped onto the range between 40 and will be perceived as identical. Depending on the data mapped to this color scale, this could mean tens of degrees in temperature, hundreds of feet in elevation over sea level, or large variations in the radiation dose delivered to oncology patients.

In the CD ROM version of this paper we include picture of a rainbow colormap applied to the topology of Florida compared with a perceptual colormap. This example clearly shows how much these regions where discriminability is reduced can distort the meaning of the data. In this example, the region around sea level is all the same color, distorting the impression of where the boundaries of Florida and the continental shelf are located.

Figures 3 through 8 show the degree to which these departures from uniformity are reduced by using hue, brightness, and saturation scales in the L*a*b* and Munsell spaces. These figures form a matrix of plots, showing the three perceptual dimensions along the rows and the two color spaces in the columns.

The remarkable thing about these graphs is the similarity of the increment threshold curves obtained using hue and luminance scales derived from the L*a*b* and Munsell color spaces. Consider first the hue-based colormaps in the top row. Relative to the standard rainbow colormap, these colormaps show much smaller increment-threshold variations, with values ranging between 1.0 and 5.5. This may be attributable to the much lower saturation of these colormaps, or to the fact that these colormaps are designed to vary in hue only. These colormaps do show factor of 2.5 decreases in hue sensitivity, but tuned broadly in the low- and high-wavelength regions of the colormap. The luminance graphs for colormaps derived from these two color spaces appear even more similar, with increment thresholds steady at approximately 0.4 for all but the lowest luminance ranges, where the increment threshold shoots up a factor of 10 to 4.0. Both these colormaps provide significant improvements over the uncalibrated HLS luminance colormap, and provide a significantly longer range over which deltaL is low and constant.

The saturation curves for L*a*b* and Munsell are also very similar. Over the whole range from achromatic to highligh saturated, linear steps in L*a*b* saturation produce nearly constant increment thresholds.

Discussion

Creating color scales which encode magnitude

These experiments have shown that both luminance and saturation scales are good candidates for representing increments in magnitude. The description of luminance used in both the L*a*b* and Munsell colormaps provided equal-interval perceptual data over a larger luminance range than the uncalibrated HLS colormap. All three of the luminance colormaps tested, however, showed a dramatic decrease in sensitivity at the low-luminance end of the scale.

The saturation scales developed using the L*a*b* definitions of saturation also appear to be good candidates for creating colormaps for encoding data magnitude. For these scales, deltaS was constant across the entire range from achromatic to deeply saturated, in juxtaposition to the luminance color scales, although the size of the increment was greater. This suggests that each perceptual step corresponds to a greater deviation in saturation than in luminance, consistent with the steeper slope of the saturation scaling curve in Stevens' classic paper.

None of the hue-based colormaps tested in this paper provided a good candidate for encoding magnitude information. Large increases in increment threshold were demonstrated both for the uncalibrated rainbow colormap and for the two isoluminant, iso-saturation hue maps produced as trajectories in L*a*b or Munsell space. This supports Stevens' idea that hue is not a "prothetic" dimension and therefore does not produce a constant magnitude scale.

It is interesting to note that the hue increment threshold curve most resembling human wavelength discrimination curves is the graph for rainbow colormap. This curve shows the typical decrease in sensitivity in the middle of the spectrum, with sensitivity losses (high delta lambdas) in the short ("blue") and long ("red") wavelength regions. This may not be surprising since the rainbow map is designed to follow the spectrum locus as much as the color gamut of the display will permit, and produces a colormap which, like the physical wavelength spectrum, varies in luminance and saturation across its range.

Is an equal-increment scale sufficient? The purpose of colormaps is to provide a scale to represent data values. A colormap which just preserves JNDs does not necessarily preserve the order of the increments. The steps in a color sequence might each be discriminably different but, for example, go up and down in perceived magnitude.

Why are delta L and delta S constant across the range? If the steps in L*a*b* and Munsell are already proportional because of a transform performed or because the steps reflect human judgments, then equal steps of this perceptually-proportional scale should produce constant increments.

Summary

In this paper we experimentally measured how hue, luminance and saturation scales represent magnitude information by constructing colormaps which trace carefully controlled paths through Munsell and L*a*b* color spaces, and comparing these results with colormaps designed using standard computer graphics methods.

References

1. H.B. Barlow, *Journal of Physiology*, **136**, 469-488 (1957).
2. James D. Foley, et al., *Computer Graphics: Principals and Practice*, 1991.
3. N.V.S. Graham, *Visual Pattern Analysers*. Oxford: Oxford University Press, New York: Clarendon Press. (1989).
4. D.H. Kelly, and R.E. Savoie, *Perception and Psychophysics*, **14**, 313-318 (1973).
5. G.E. Legge, *Vision Research*, **21**, 457-467 (1981).
6. S. Pizer, *Computer Graphics and Image Processing*, **14**, 262, 1981
7. D. L. Post and Christopher S. Calhoun, *Color Research and Application*, **14**(4), 172 (1989).
8. P. K. Robertson, *IEEE Computer Graphics and Applications*, 50-63, 1998.
9. B. E. Rogowitz and L.A.Treinish, *IEEE Spectrum*, 52-59, 1998.
10. B. E. Rogowitz and L.A. Treinish, *Computers and Physics*, 268-274, 1996.
11. S. S. Stevens, *Perception and Pychophysics*, **1**, 5-8, 1996.
12. W.D. Wright, and F.H.G Pitt, Hue discrimination in normal color vision, *Proceedings of the Physical Society* (London), **46**, 1934.

Acknowledgements

Thanks to Lloyd Treinish for his help in conceptualizing and prototyping the visual stimuli.

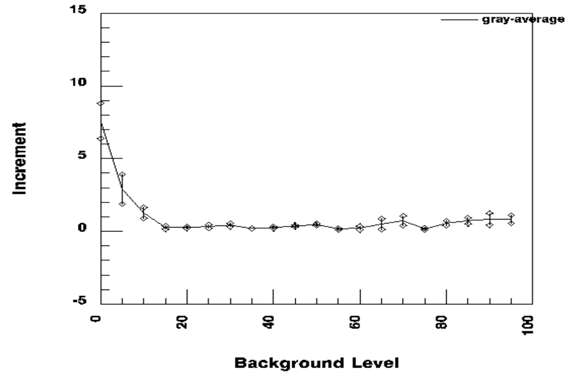

Figure 1 Standard gray-scale

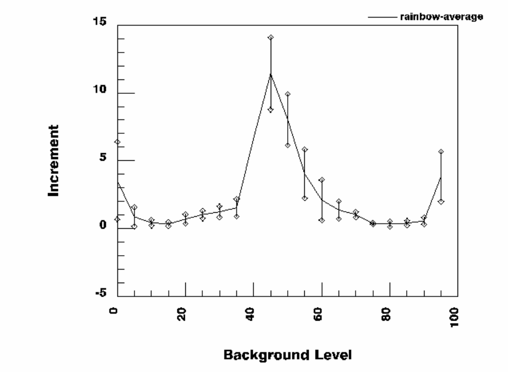

Figure 2 Standard rainbow

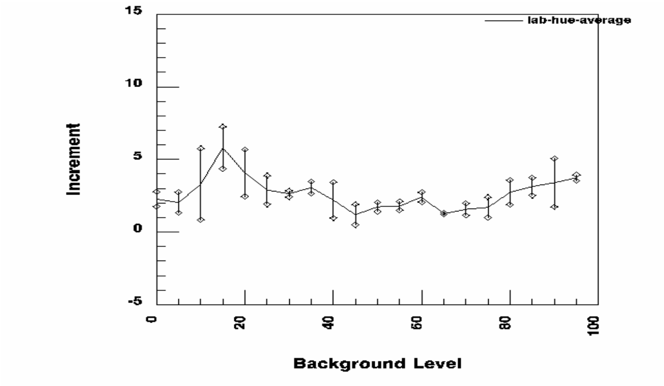

Figure 3 L*a*b* Hue-varying

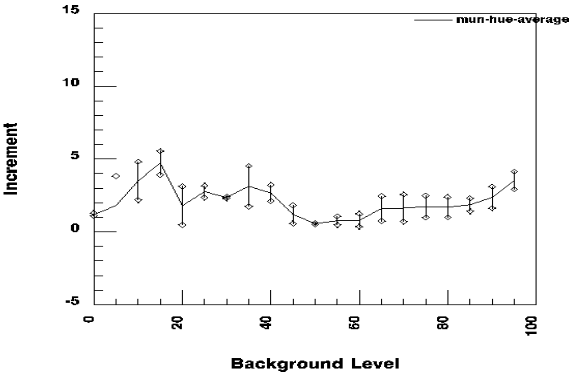

Figure 4 Munsell Hue-varying

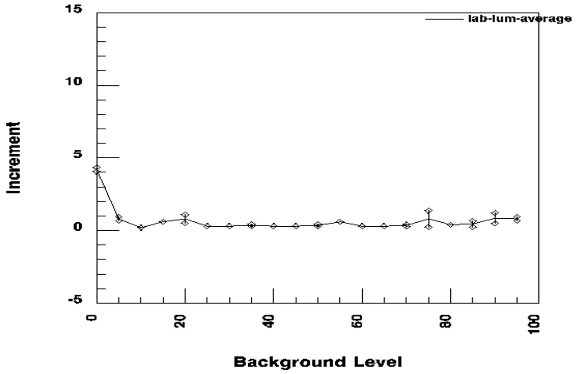

Figure 5 L*a*b* Luminance-varying

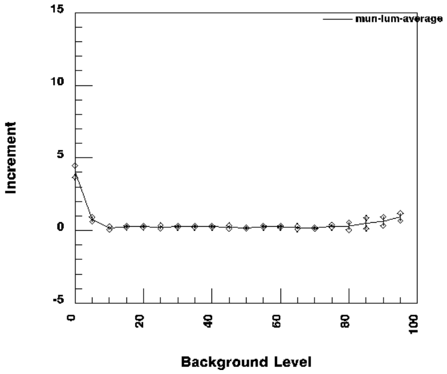

Figure 6 Munsell Luminance-varying

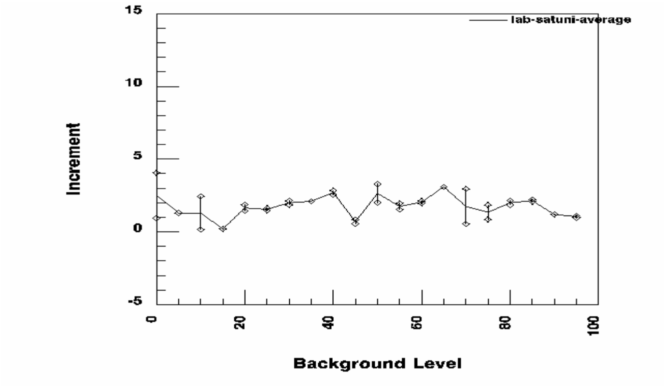

Figure 7 L*a*b* Saturation-varying

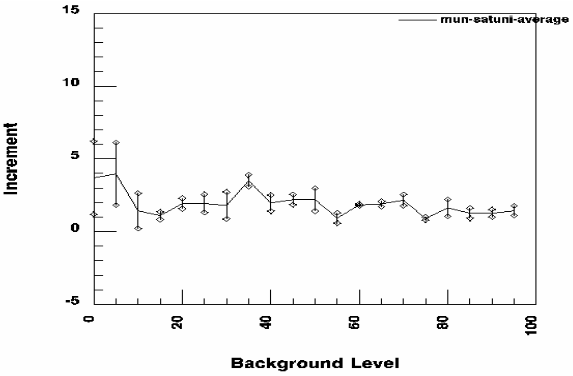

Figure 8 Munsell Saturation-varying