RC24542 (W0804-104) April 29, 2008
Computer Science

IBM Research Report

Self-Corrected Perceptual Colormaps

Donna L. Gresh
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Self-Corrected Perceptual Colormaps

Donna L. Gresh*
IBM T.J. Watson Research Center

Abstract

We describe an algorithm that transforms a given input colormap to one that is more perceptually uniform (*i.e.*equal steps in data value are equally perceivable). We have incorporated the algorithm into a Java™ application. First, the application measures a given user's perception function for a given monitor and colormap. Then, it computes a "normalized" colormap which preserves the color ordering of the original map but modifies the positions of the colors so that the resulting colormap will be much more perceptually uniform. We present results for both commonly used colormaps and personal, "made-up" colormaps, and show the improvement that is achieved in each case. The normalized maps tend to preserve luminance linearity, within the confines of the characteristics of the input colormap. Overall discrimination is best when the original colormap incorporates both hue and luminance variation.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

Keywords: visualization, perception, colormaps

1 Introduction

In many applications of scientific visualization, colormaps are used to represent continuous variables. Applications use both standard, common colormaps (such as the rainbow and grayscale colormaps), as well as colormaps designed for a particular purpose. However, many colormaps currently in popular use do not satisfy properties which would make them appropriate for the accurate interpretation of data, as outlined by Trumbo[11], whose first two principles, for example, are those of Order ("if the levels of a statistical variable are ordered, then the colors chosen to represent them should be perceived as preserving the order") and Separation ("important differences in the levels of a statistical variable should be represented by colors clearly perceived as different"). In particular, [6, 7] discuss the fact that for many maps, equal steps in data value do not produce equal perceptual differences. This can lead to misinterpretation of data; small differences in data value may look larger than they truly are, or large differences may be masked. One solution is to use known perceptually uniform colormaps. For example, luminance maps (grayscale maps and their relatives) tend to be reasonably perceptually uniform [10, 2], and are commonly used in medical imaging, for example, where discrimination of fine detail is extremely important for proper interpretation.

Previous work (e.g. [5], [8]) discuss theoretical frameworks for constructing colormaps which satisfy particular criteria, generally using perceptual color spaces to select colors. However, such approaches require calibrated displays, which are not always easily available. Also, for better or worse, grayscale and other luminance maps tend to be unpopular in scientific visualization. They lack bright colors, and thus are not very visually appealing, and it is difficult to estimate actual data values by eye, even given a colorbar to go with the image. The rainbow colormap is nearly ubiquitous in scientific visualization, and is typically the default map. It is attractive, colorful, with easy to discriminate "regions" of the data which can then be associated with specific data values given a colorbar key (low values with blue, medium-low with cyan, medium with green, etc.). The problem with the rainbow colormap is that the minimum perceivable increment varies widely over the range of the map, as shown by [1], who found that in the green region of the map (between about 40% and 60% of the colormap range), the necessary increment to be perceivable is more than 10 times as large as it is in the red and blue regions of the map. Thus changes in data value in this region are masked by the perceptual non-unifomity of the colormap.

The solution we propose is to allow users their choice of colormap, including the rainbow map, but give them a mechanism for improving it. We do this by measuring the perceptual discrimination function for the given colormap, then adjusting the map to make the perception function flat.

We note that choice of a colormap must depend on the application; for example, [6, 2] showed that the optimal colormap depends on whether large scale trends or small details are more important to detect, and [9] discusses the detection of clusters in data. The method we use is aimed at enabling users to detect small changes in data value (details). The maps we consider are designed to represent continuous variables, thus are smooth.

2 Measurement of the Perception Function: Experimental Setup

Our goal was to design an experiment which would straightforwardly and quickly measure the perception function for a given user, a given monitor, and a given colormap. We were particularly interested in measuring perception akin to looking for small details against a background. To accomplish this, we first created a reference data set of size NxM pixels. The data set has values equal to i/N in the horizontal dimension, and i/N in the vertical dimension, where i is the horizontal pixel number. Thus the function varies from 0 to 1 in x, and for a given x is constant in y (see Figure 1). In practice, we set N equal to 500 pixels and M equal to 100 pixels. To this reference data set is added a small data offset, square in shape, at some location in x and y. The size of the square was 15 pixels.

We created a Java™ program which would color both the reference data set and the test data set using the colormap under consideration. We designed a very simple input format for the colormap, with positions (or data values) specified between 0 and 1, and corresponding color values specified as an RGB triplet. The positions and colors are found in two files, the names of which are specified on the command line to the application. The two images are presented one above the other to the user, with the test set randomly chosen to be either on the top or bottom. The location of the test patch starts at the left side and moves toward the right as each perceptual threshold is measured. Thus the user knows approximately

*gresh@us.ibm.com, IBM T.J. Watson Research Center, P.O. Box 704 Yorktown Heights, NY 10598

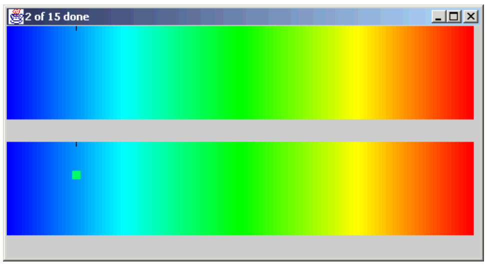

Figure 1: Example of test pattern presented to a user, using the standard rainbow colormap. In this example, the offset patch can be seen in the bottom image, in the blue region near the left.

where to search for the pattern. We found this to result in more reproducible results, as the user was not forced to search over the entire space of the test pattern.

We implemented an adaptive staircase method of measuring the minimum discriminable offset from the background. We measured this offset at 15 different values in the colormap. We begin with a relatively large offset of 20% of the entire range added to the background data value. The user has 5 seconds to indicate, using arrow keys, whether the test image, containing the test offset pattern, is on the top or bottom. If the user either does not respond in 5 seconds, or responds incorrectly, that test offset as a function of horizontal location is marked as "incorrect." Two incorrect responses at the same offset result in a "bounce," and the next test offset at that location will be larger. A short beep is sounded when the time limit expires to alert the user that a new pair of images is being presented. If the user responds correctly within the time allotted, the same offset is presented a second time, in a random image. Two correct responses will result in the next offset at that location being smaller. (This mitigates the effect of "lucky guessing.") The locations of three "bounces" (reversals in direction of the tested offset) are averaged to provide an estimate of the threshold for that point in the colormap. An example of the test and reference pattern for the standard rainbow colormap is shown in Figure 1; in this case the test image is on the bottom, and the offset patch can be seen in the blue region near the left of this test image. This image is taken early in the test process, when the offset is still quite easy to detect. A small mark at the top of the image indicates the region in which the mark will be found. This was found to lessen the difficulty that users had in locating the region of the patch, particularly in the central region of the map. Each trial, for an entire colormap, takes approximately 10 minutes to perform.

3 Results

There are three orthogonal issues which we were interested in investigating in these experiments:

- How successful can we be at improving a colormap (especially the defacto default rainbow colormap) for a given user on a given monitor? That is, how flat is the perception function for the *improved* colormap?
- Given different starting colormaps, are there differences in how well the improved maps perform? That is, are the normalized maps significantly better for some starting colormaps than others?
- How much variation is there in the normalized colormaps for different users and different monitors?

The following sections will discuss each of these in turn.

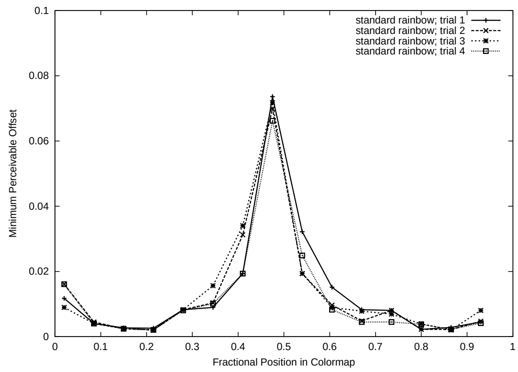

Figure 2: Minimum perceivable increment vs. relative location in colormap as measured for the standard rainbow colormap for a single user using a crt monitor. Four trials were performed; each trial is shown.

3.1 Deriving a normalized colormap

The first step is to derive an normalized colormap from the measurement of the perception function for a given user on a given monitor. Figure 2 shows the measured perception function for a single user. To show the variation for a single user from trial to trial, four trials are shown. This test was conducted on an IBM P201 cathode ray tube monitor for the standard rainbow colormap (*i.e.*,which varies uniformly in hue in a hue-saturation-value colorspace).

Our goal is to modify the colormap so that this perception function is as uniform as possible. We do not want to change the colors in the map, or their ordering relative to one another; we rather want to change the data values to which each color is assigned. We designate the measured function of Figure 2 $f(x)$, where x is the relative position in the colormap, and $f(x)$ is the just-perceivable offset. Thus we know that pairs of colors in the colormap, at x and $x + f(x)$, are "just barely distinguishable." We want to derive a transforming function of x, $g(x)$, such that $g(x)$ and $g(x + f(x))$ are at a constant spacing; if this is the case, then just barely distinguishable colors will be at a constant spacing.
Thus,

$$g(x + f(x)) - g(x) = C$$

where C is some constant. As an approximation, we can write

$$g(x) + f(x)g'(x) - g(x) = C$$

or

$$g'(x) = \frac{C}{f(x)}$$

thus

$$g(y) = \int^y \frac{C}{f(x)}$$

We choose C such that the endpoints of the colormap are 0 and 1, respectively.

We apply this procedure to the measured perception function to derive a new colormap. We then repeat the experiment using the new map to measure the perception function of the normalized colormap. Given the reproducibility of the perception function measurements shown in Figure 2, we performed this process for only the first trial. The result is shown in Figure 3. We see that the new measured perception function is significantly flatter than the original, which is shown for comparison. We also applied this colormap to a topographic data set, and compared it with the original rainbow colormap. The results are shown in Figures 4 and 5. You can

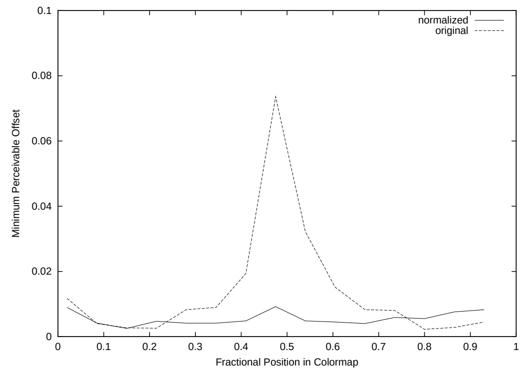

Figure 3: Minimum perceivable increment vs. relative location for the normalized colormap (using the measurements of the first trial in Figure 2). For comparison, the original perception function is shown as well.

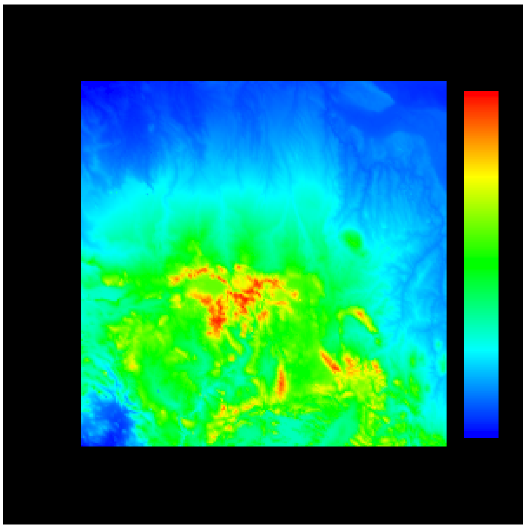

Figure 4: Standard rainbow colormap applied to a topographic data set.

see that the result of the poor discrimination for the original rainbow colormap in the central, green, region has lead to the colormap being significantly compressed in this region. While results for a single data set can only be anectdotal, we see enhanced detail in the low end of the scale, as well as better discrimination in the central, formerly green, region.

One important question is whether the measurement of *positive* data offsets against the background is equivalent to the measurement of *negative* data offsets. That is, whether the discrimination d at a position p when positive offsets are used, would be measured as the same value at position $p+d$ when negative offsets are used. We performed two simple experiments to test this hypothesis. The first was to invert the rainbow colormap, so that it runs from red to blue instead of from blue to red, and then measure the discrimination function using the procedure previously described. To compare the results to those measured for the standard rainbow colormap, it was necessary to adjust the positions (x values) by the function $1-p-d$, where p is the position in the inverted colormap and d is the discrimination value measured at that point. The second experiment was to modify the experimental procedure to produce negative steps in

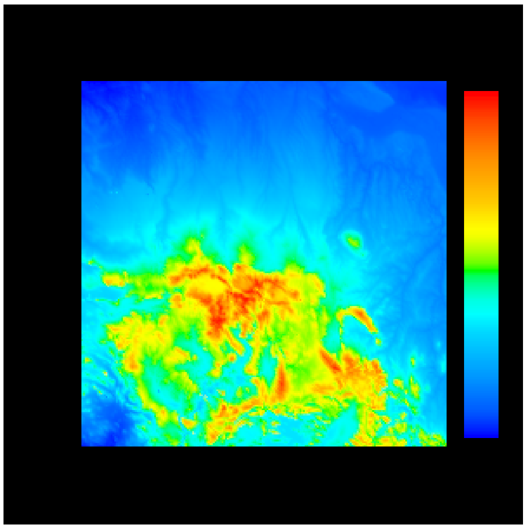

Figure 5: Normalized rainbow colormap computed from the first trial of Figure 2 applied to the same topographic data set as in Figure 4. Note the enhanced detail in the blue region and the reduction in prominence of the green region.

value for the test image, and use the standard rainbow colormap. In this case we again must modify the x values using the function $p-d$ before plotting. The results of these two experiments are shown in Figure 6, where we see excellent agreement between all three experiments for the rainbow colormap. To compare the effect on the resulting colormap, refer to Figure 7, where we compare the eventual position of each color (indexed from 0 to 100) in the normalized colormap.

3.2 Quality of optimized colormaps

The techniques of section 3.1 can be applied to any colormap. As an example, we measured the perception function for the standard grayscale map, which varies linearly in intensity from the minimum to the maximum data value. We then normalized it, with the result shown in Figure 8, which also includes the results of the normalization of the rainbow map (using the first trial shown in Figure 2). Note that the standard grayscale map has relatively poor discrimination in the low end of the scale, but is relatively flat over most of the range. The equalization process brought down the discrimination function in the low end of the scale, as desired. You may note, however, that the level of the discrimination function for the normalized rainbow map is, in general, lower than that of the normalized grayscale.

To investigate the performance of this algorithm at improving other colormaps, we also applied it to two non-standard, personally created, colormaps. The first of these included both hue variation and luminance variation (see Figure 9), and was intentionally manufactured to have a region of little variation in color. The second map (see Figure 10) was monotonic (but not linear) in luminance, with some hue variation. Our algorithm performed well at improving the perception function for both of these colormaps as well. The results of applying these colormaps to the topographic data set are shown in Figures 11–16.

To quantify the "quality" of a colormap we define two measures: uniformity and precision. Uniformity measures whether equal steps in value are equally perceivable. We define uniformity u as the standard deviation of $f(x)$; we want this value to be small. Precision measures whether small steps in value are perceivable. We define

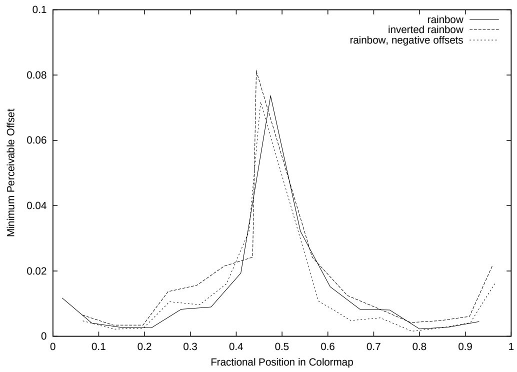

Figure 6: Test of the hypothesis that a negative offset is perceived equivalently to a positive offset. We compare the measured perceptual discrimination function for the standard rainbow colormap, and inverted rainbow colormap, and the standard rainbow colormap measured using negative rather than positive offsets.

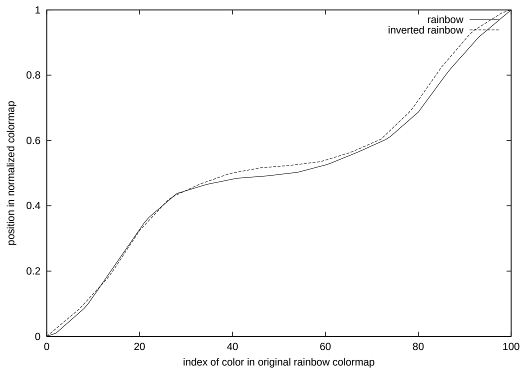

Figure 7: Location of each color in the normalized colormap *vs.*the index of the color in the original colormap, for the rainbow and inverted rainbow colormaps. (The relative positions (y values in this plot) have been subtracted from 1 for the inverted rainbow colormap for a useful comparison.)

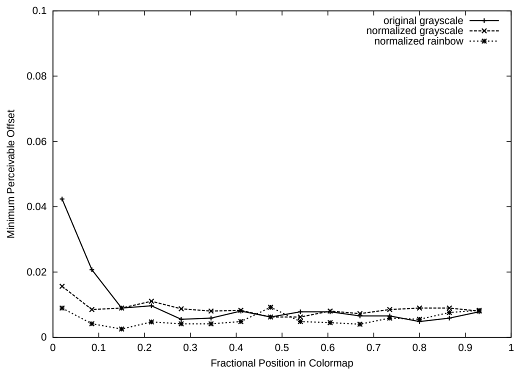

Figure 8: Minimum perceivable increment vs. relative location in colormap as measured for the standard grayscale colormap for a single user using a crt monitor. A normalized colormap was then constructed using these measurements, and the results for this map are shown. For comparison, the results for the normalized rainbow colormap are shown as well.

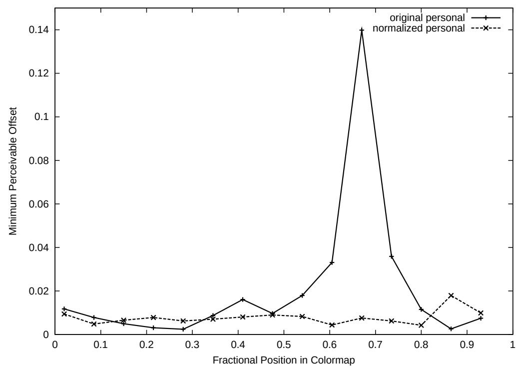

Figure 9: Minimum perceivable increment vs. relative location for a non-standard, personally created colormap. Measurements of the perception function for both the original and the normalized colormaps are shown. The actual colormaps themselves can be seen in Figures 13 and 14.

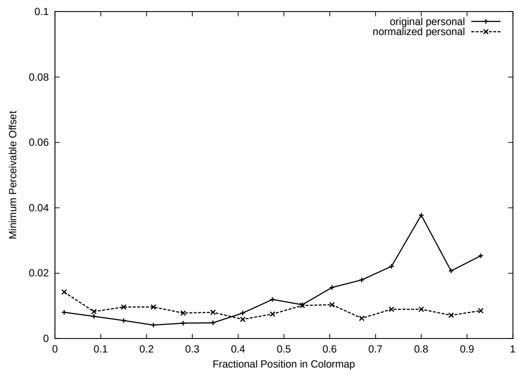

Figure 10: Minimum perceivable increment vs. relative location for a 2nd non-standard, personally created colormap. Measurements of the perception function for both the original and the normalized colormaps are shown. The actual colormaps themselves can be seen in Figures 15 and 16.

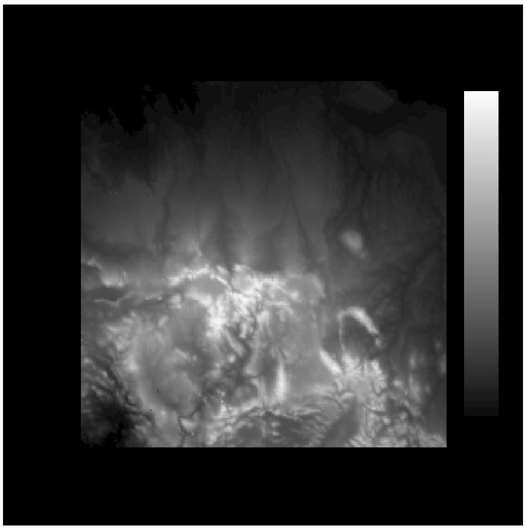

Figure 11: Standard grayscale colormap applied to a topographic data set.

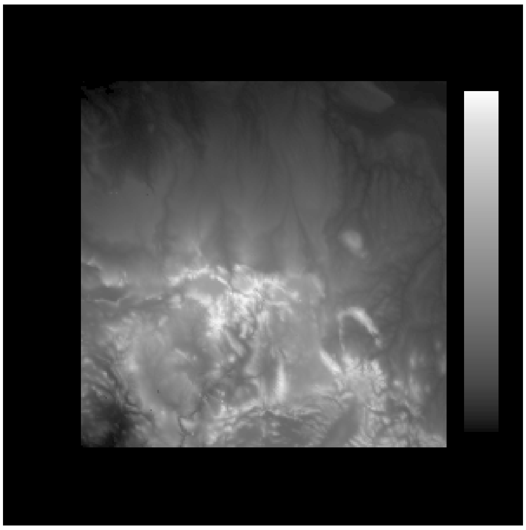

Figure 12: Normalized grayscale colormap applied to the same topographic data set as in Figure 11. Note the enhanced detail in the low data value region.

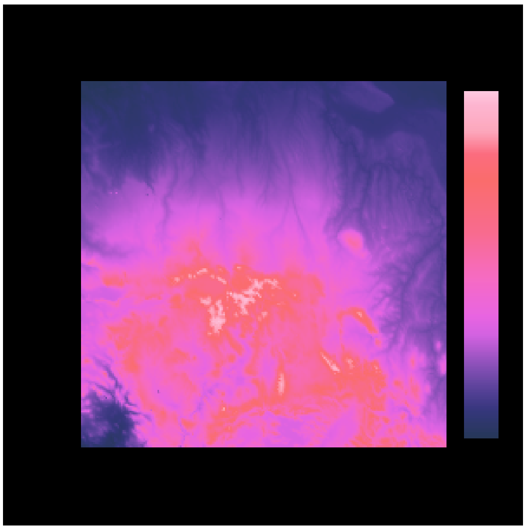

Figure 13: Individually created (non-standard) colormap applied to a topographic data set.

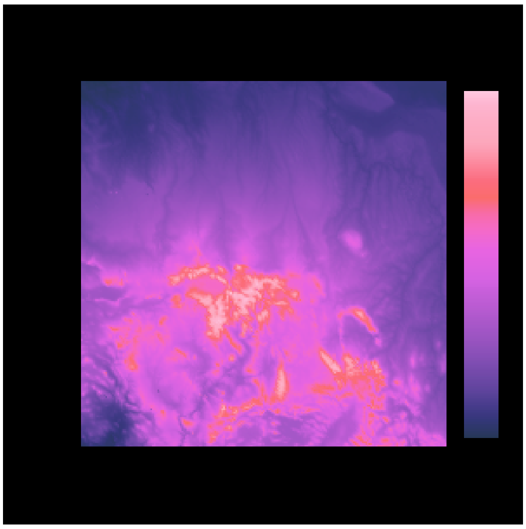

Figure 14: Normalized version of the colormap of Figure 13 applied to the same topographic data set.

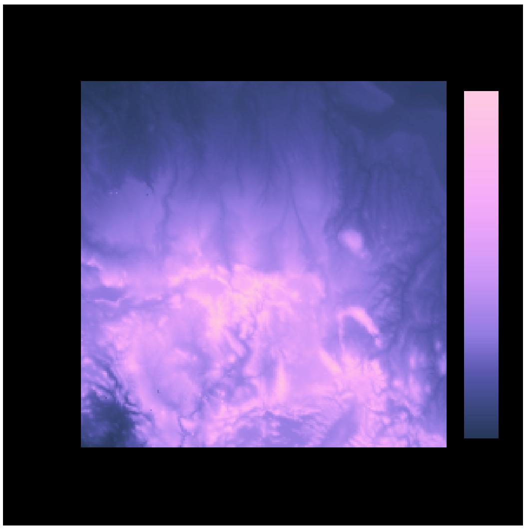

Figure 15: A second individually created (non-standard) colormap applied to a topographic data set.

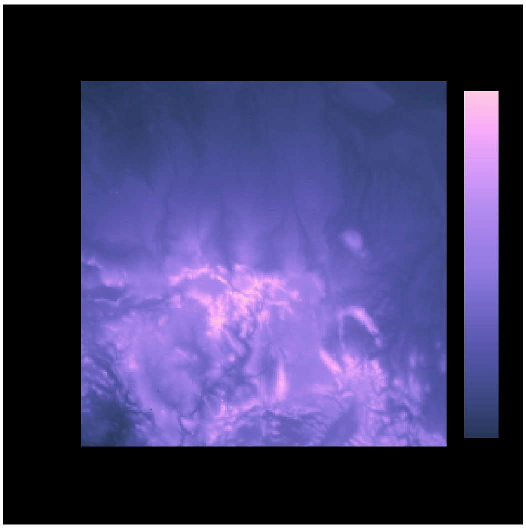

Figure 16: Normalized version of the colormap of Figure 15 applied to the same topographic data set.

colormap	precision p	uniformity u
original rainbow	0.014	0.021
normalized rainbow	0.006	0.006
original grayscale	0.010	0.013
normalized grayscale	0.008	0.008
original “personal”	0.021	0.038
normalized “personal”	0.008	0.008
original “2nd personal”	0.014	0.015
normalized “2nd personal”	0.009	0.008

Table 1: Measurements of the precision p (mean) and uniformity u (standard deviation) of the discrimination function for original and normalized colormaps. Note that the results for the normalized rainbow colormap are better than for any other colormap considered, and that in each case the measurements for the normalized maps are an improvement over the original map.

precision p as the mean of $f(x)$. We also want this value to be small. The results for all of the colormaps investigated here are shown in Table 1. We see that as might be expected, the precision, and particularly the uniformity, of the standard rainbow map are poor; however see also see that the precision and uniformity of the *normalized* rainbow map are better than *all* of the colormaps investigated, both original and normalized. That is, detail may be more easily and accurately seen with the normalized rainbow map than with any of the other maps. We see two possible explanations for this result. One possibility is that the hue differences in the rainbow colormap allow an additional channel for discriminating differences in addition to the luminance variation. The second possibility is that the luminance increases and decreases of the rainbow colormap [1] allow easier detection, as luminance is the primary way in which small details can be seen, and hue allows data with the same luminance but different values to be discriminated.

3.3 Luminance profiles of colormaps

Of interest are the luminance profiles of the normalized colormaps, and how they compare to the luminance profiles of the original colormaps. These may give some guidance to the underlying mechanism of “good” colormaps. Figures 17–20 show the luminance profiles of each of the colormaps discussed thus far. These maps were computed using the formula outlined in [4]. It is interesting to note that in each case, the luminance profile of the normalized map is much closer to linear than that of the original map. In the case of the rainbow colormap, for which a monotonically increasing function is impossible by nature of the colormap, we see that the normalized map has roughly linear increases and decreases of luminance, and that the slopes of the function are roughly constant. This would imply that linear luminance is in fact highly desirable for perceptually uniform colormaps, and that the rainbow colormap is able to depart from monotonicity due to the semantics of hue variation, which allow colors of equal luminance to be distinguished.

Another issue to consider is the eventual consumer of a normalized colormap, and whether any further adjustments to colors are made before images are produced. For example, the Data Explorer software visualization system [3] applies by default a gamma correction of 2 for most computer architectures, to attempt to correct for non-linearities in the monitor output. Thus if a colormap was to be used with Data Explorer or another product which performs gamma correction, the experiment should be done using the same gamma correction. This adjustment is incorporated as an option in the experiment program discussed here.

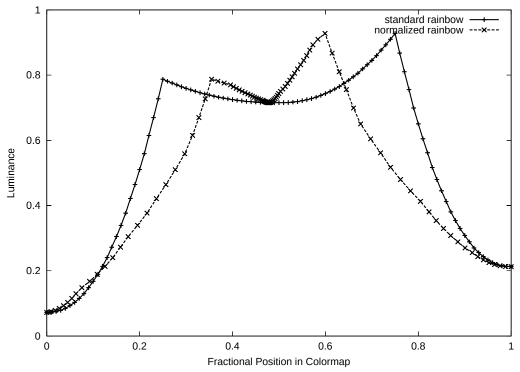

Figure 17: Luminance profiles for the original and the normalized rainbow colormap.

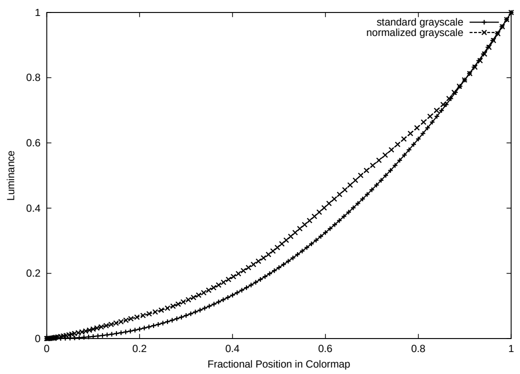

Figure 18: Luminance profiles for the original and the normalized grayscale colormap.

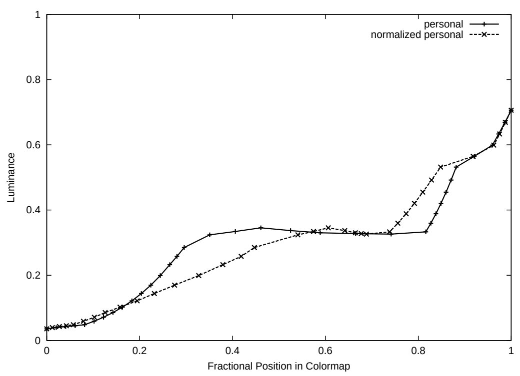

Figure 19: Luminance profiles for the original and the normalized non-standard, personally created colormap.

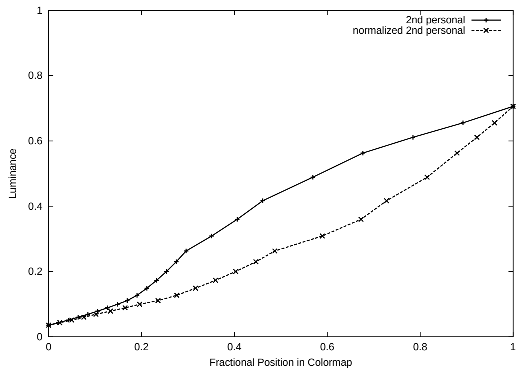

Figure 20: Luminance profiles for the original and the normalized 2nd personally created, non-standard colormap.

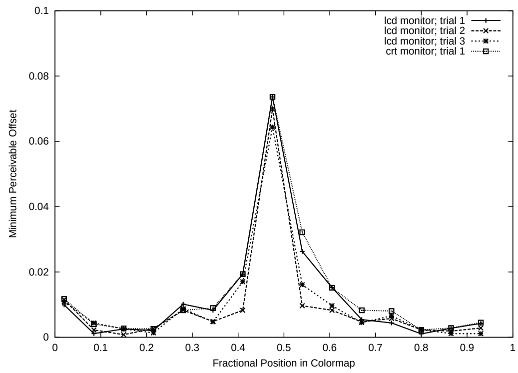

Figure 21: Minimum perceivable increment vs. relative location in colormap as measured for the standard rainbow colormap for a single user using an lcd monitor. Three trials were performed; each trial is shown. In addition, the first trial as measured for the crt monitor is shown for comparison.

3.4 Variation in perception between users and monitors

Another topic of interest is how much variation exists between the perception function as measured for different monitors, and for different users. Figure 21 shows the results for a Thinkpad LCD monitor for the rainbow colormap. The experiment was repeated three times to test the reproducibility for a single user and a single monitor. As Figure 21 shows, reproducibility was excellent.

For comparison, the first trial as measured for the same user on the CRT monitor is shown as well. There is no significant difference between the measurements for the LCD monitor and the CRT monitor, suggesting that a single normalized colormap would be appropriate for either monitor, at least for this user and this pair of monitors. Variation across several CRT monitors is shown in Figure 22. We see that while the overall behavior is consistent, one of the CRT monitors had significantly better discrimination in the central, green region than did the others. It would thus be prudent to measure the perception function for a given monitor to derive the optimal colormap for it.

To investigate the variation between users, we also measured the perception function for the rainbow colormap for several other users on the same CRT monitor as used in Figure 2. The results of this experiment are shown in Figure 23. As can be seen, results are

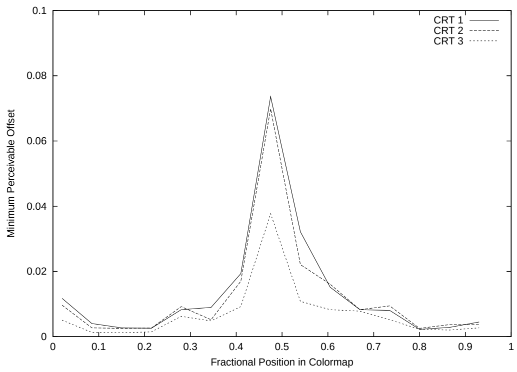

Figure 22: Perception function for three different CRT monitors.

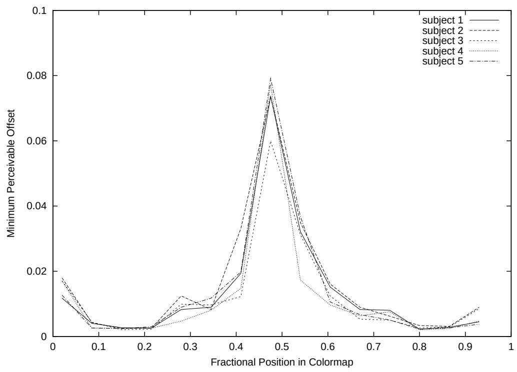

Figure 23: Perception functions for several users measured on a crt for the standard rainbow colormap.

quite consistent. Each user's perception function was then used to construct a normalized rainbow colormap for that user, and the user was tested again. The results are shown in Figure 24. Again, the procedure we describe here did a good job of creating a much more perceptually uniform map. Given the uniform results obtained for a given map on a given monitor, it seems reasonable that one could create a perceptual uniform colormap that would provide good results for most people using it.

4 Conclusions

We have described an algorithm which can be used to derive a normalized colormap given measurements of a user's perception function. The normalized map will exhibit much more uniform behavior, and thus be more perceptually accurate. This algorithm can be used to improve any given input colormap. We have found that for a given monitor and colormap, different users have very similar measured perception functions, indicating that a single normalized map may be used for any user of that monitor. In general, measurements for a given standard colormap (for example, the rainbow colormap) are similar across different monitors, indicating that a "reference" version of the map could be published which would exhibit better behavior than the default; however for optimal results one may wish to create a normalized colormap specifically for a given monitor, as results do vary somewhat.

Our results also show that luminance linearity is desirable for perceptual uniformity, and that the normalized versions of the colormaps investigated here are in general much more linear in luminance than the original maps, within the limits set by the overall luminance behavior of the colormap. We also found that the rainbow colormap, once normalized, allows better discrimination of subtle detail than any other map investigated. We postulate that this is because more inherent luminance variation is achievable using a map which encodes using both hue and luminance.

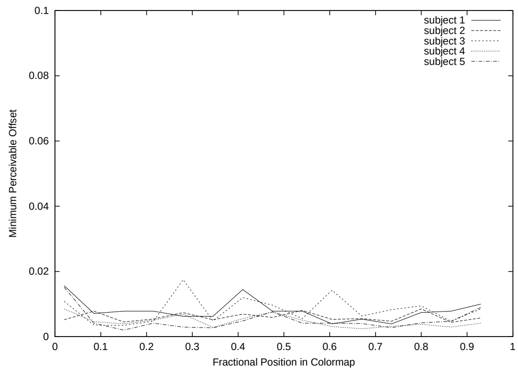

Figure 24: Perception functions for the same users as shown in Figure 23, each using the normalized colormap derived from their measurements shown in that figure.

5 Acknowledgments

The author would like to thank Bernice Rogowitz for her suggested improvements in perception function measurement techniques and for discussions of color map issues in general, and Jeffrey Kephart for help in the development of the equation on which this algorithm is based.

References

[1] A. D. Kalvin, B. E. Rogowitz, A. Pelah, and A. Cohen. Building perceptual color maps for visualizing interval data. In *Human Vision and Electronic Imaging V*, pages 323–335, 2000.

[2] B. R. L.D.Bergman and L. Treinish. A rule-based tool for assisting colormap selection. In *Proceedings of IEEE Visualization*, 1995.

[3] B. Lucas, G. D. Abram, N. S. Collins, D. A. Epstein, D. L. Gresh, and K. P. McAuliffe. An architecture for a scientific visualization system. In *Proceedings of IEEE Visualization '92*, 1992.

[4] C. Poynton. Poynton's color faq, http://www.inforamp.net/˜poynton/colorfaq.html.

[5] P. Robertson and J. O'Callaghan. The generation of color sequences for univariate and bivariate mapping. *IEEE Computer Graphics and Applications*, 6:24–32, 1986.

[6] B. Rogowitz and L. Treinish. How not to lie with visualization. *Computers and Physics*, pages 268–274, 1996.

[7] B. Rogowitz and L. Treinish. Data visualization: the end of the rainbow. *IEEE Spectrum*, Dec. 1998.

[8] S.M.Pizer. Intensity mappings to linearize display devices. *Computer Graphics and Image Processing*, 17:262–268, 1981.

[9] I. Spence and A. Efendov. Target detection in scientific visualization. *Journal of Experimental Psychology; Applied*, 7:13–26, 2001.

[10] S. S. Stevens. Matching functions between loudness and ten other continua. *Perception and Psychophysics*, 1:5–8, 1966.

[11] B. Trumbo. Theory for coloring bivariate statistical maps. *The American Statistician*, 35:220–226, 1981.