Perceptual Steps along Color Scales

Haim Levkowitz

Institute for Visualization and Perception Research, Department of Computer Science, University of Massachusetts Lowell, 1 University Avenue, Lowell, Massachusetts 01854

ABSTRACT

Color scales are commonly used to represent numerical information visually. Most scales are derived from some physical or mathematical behavior; even worse, sometimes they are selected based solely on hardware capabilities. In most known cases, no consideration is made to the perceptual capabilities of the human observer, who is ultimately the “consumer” of the information to be delivered by the scale. This article presents a method and an algorithm for the derivation of color scales, such that their perceptual properties—in particular, the perceptual steps between colors along the scale—can be controlled by the scale designer. This approach has been used for the design of the linearized gray scale and the linearized optimized color scale (LOCS). These scales are demonstrated and are also compared to the linearized heated-object scale. © 1996 John Wiley & Sons, Inc.

I. INTRODUCTION

A color scale is a pictorial representation of a set of distinct categorical or numerical values in which each value is assigned its own color. Color scales have been used for the representation of parameter distributions for quite some time. Indeed, whenever an image of some parameter distribution is displayed, some color scale has been employed. Most often, it is the gray scale (usually not considered a color scale). The main advantages of the gray scale are its simplicity and the natural sense of order it induces. Its main claimed disadvantage is a limited perceived dynamic range (PDR) [only 60–90 just-noticeable-differences (JNDs)]. This PDR may be inadequate whenever a broader range of data values is to be represented. For this reason, as well as for aesthetic reasons, people have been seeking alternative color scales—typically referred to as pseudo-color scales.

Schuchard cited several studies implying that medical images displayed with pseudo-color scales may give more information to the observer, but admitted there is no conclusive evidence that proves this. He hinted, with no evidence though, that the heated-object scale possesses some desirable qualities for use in medical image display [1, 2].

There is no controversy that appropriate color coding can significantly improve detection of targets among a few objects (typically <10). However, despite all the claims of the advantages of pseudo-color scales, to date no study has demonstrated the superiority of any color scale—as can be objectively measured by improved performance—for continuous data, or data with large number of distinct values.

Todd-Pokropek claimed the superiority of the gray scale over the heated-object scale for detection of small abnormalities in a reasonably complicated background [3]. Levkowitz and Herman also demonstrated the superiority of a linearized gray scale over their own linearized optimized color scale (LOCS) and over the heated-object scale in a three-scale comparison study of target detection in a noisy background [4].

As Schuchard pointed out, this may be because no large attention has been given to the development of color scales based on perceptual properties [1].

This article partially addresses this issue by presenting a method and an algorithm for the derivation of color scales, such that their perceptual properties—in particular, the perceptual steps between colors along the scale—can be controlled by the scale designer. This approach has been used for the design of the LOCS. The results are demonstrated on the LOCS and the linearized gray scale; both are also compared to the heated-object scale.

Section II presents the main concepts and issues behind color scales. Section III discusses some issues fundamental to the adjustment of perceptual steps along a color scale. Section IV summarizes the main approaches to the linearization of color scales, and presents our own algorithm. Section V provides a short discussion, and concludes with an outline of some future research that is still needed.

II. COLOR SCALES

A. Properties. Given N distinct values \( v_1, \ldots, v_N \), either numerical or categorical, to be represented by N colors \( c_1, \ldots, c_N \), respectively, several researchers have identified three desired properties: Colors in the scale used to represent the original values should (i) be perceived to preserve the order of the values (if such order exists); (ii) correctly convey uniformity among values they are representing, and representative distances between them; and (iii) create no artificial boundaries that do not exist in the original data. That is, the scale should be able to represent continuous values continuously [4–6].

B. Common Scales. The most commonly used scale is the gray scale, the result of traversing the color solid along the achromatic (lightness) axis. This can be implemented by keeping equal intensities for the three primaries, red, green, and blue (R, G, B) and increasing them monotonically from 0 to M—the maximum value they can assume. Another common scale is the rainbow scale in its various variants, which traverses the color solid along a path.
through all the hues of the rainbow [red, orange, yellow, green, blue, indigo, violet (ROYGBIV)]. The two most common variants are one that maintains a constant lightness, and one beginning at black and ending at white, with colors' lightness values monotonically nondecreasing.

Two other frequently mentioned scales are the heated-object and the magenta scales. These two traverse the color solid from black to white through two different paths originating at the red axis. Both scales are based on the claim that natural color scales seem to be produced when the intensities of the three primary colors, red, green, and blue, rise monotonically and maintain the same order of magnitude throughout the entire scale [5]. Both scales satisfy this property. The heated-object scale is implemented by increasing the primaries, maintaining the order $R \geq G \geq B$. Its path through the color solid is limited to 60° counterclockwise from the red axis. It is based on the human visual system's maximum sensitivity to luminance changes for the orange-yellow hue. For example, in the Munsell color system there are more distinct divisions for high-value highly saturated yellows than there are for other hues [7]. The magenta scale is implemented by increasing the primaries, maintaining the order $R \geq B \geq G$. Its path through the color solid is limited to 60° clockwise from the red axis. It is based on the human visual system's highest sensitivity to hue changes for the magenta hue. For example, experiments performed to measure discrimination of hue showed that the best hue discrimination is achieved for the purples [8].

C. Optimized Color Scales. Since many colors in a scale may not be perceived as distinct, for an informative representation it is important to maximize the number of JNDs, or distinct perceived steps as equally spaced perceptual steps. Linearization of a color scale (e.g., via interpolation) and to scan it for a subset of colors that are (almost) equally spaced perceptually. The main difference between our approach and Robertson's is that, while we specify our supersampling phase in RGB coordinates as an example, we do not commit ourselves to any source color space, whereas it appears that Robertson is committed to the LHS color space as the source space.

Similarly—and perhaps more important—we do not commit ourselves to any particular uniform color space. The main reason is that we have our doubts about the global uniformity of the CIELUV color space.

B. Algorithm LOS. The output of LOS is $N^n$, the number of colors required of the output scale, and an input color scale with $N^i$ color vectors, $\mathbf{c}_i, \ldots, \mathbf{c}_{N^i}$, whose component (primary) values are (e.g., in RGB space) $(0, 0, 0) \leq (r, g, b) \leq (M^r, M^g, M^b)$, $(N^i = M^i + 1)$ such that the lightness of colors in the scale—computed by the lightness function $\mathit{L}(\mathbf{c}) = \frac{r + g + b}{3}$—is $\mathit{L}(\mathbf{c}_i) = n^i - 1$ for all $1 \leq n^i \leq N^i$.

The output of LOS is a linearized scale with $N^n$ colors, $\mathbf{c}_i', \ldots, \mathbf{c}_{N^n}'$, whose primary values are $(0, 0, 0) \leq (r, g, b) \leq (M^n, M^n, M^n)$, $(N^n = M^n + 1)$ such that the distance $d(c_i, c_i')$ between the first and last colors in the scale is divided (approximately) equally among all the color intervals, where the distance $d(c, c')$ between two colors $c$ and $c'$ is defined as the Euclidean distance between the two colors in the selected uniform color space. For example, despite our reservations about the global uniformity of CIELUV, for lack of a better choice our implementation was actually done in CIELUV. Thus, in our specific implementation, $d(c, c') = [(L^*(c) - L^*(c'))^2 + (a^*(c) - a^*(c'))^2 + (b^*(c) - b^*(c'))^2]^{1/2}$. LOS steps are (Figure 1)
Equal perceptual distance:
\[ \frac{d(c_{n}, c_{i})}{N} \]

1. Super sampling
   (a) "Stretch" the input scale: We compute from the input scale a new, longer scale with \( N' \) colors whose values are (e.g., in RGB space) \((0, 0, 0) \leq (r, g, b) \leq (M', M', M')\), as follows. The maximum primary value \( M' \) is chosen such that the scale features enough super sampling to provide sufficient resolution for the subsequent scanning for equally spaced colors in the scale (perceptual spacing is measured in some uniform color space, e.g., CIELUV). Then, the length \( N' \) of the scale is set to be \( N' = 3 \cdot (M' + 1) \) to enable storing colors in the order of their lightness (including colors with noninteger lightness) by using the value \( 3f \) as an index to the location of a color in the scale. For each color \( c_i \) in the input scale, a color \( c_{i}' = \frac{M'}{M'} c_i \) in the long scale is computed whose location in the scale is \( n' = 3f(c_i) \), three times its lightness.

   (b) Insert missing colors: Stretching the original scale into the long scale leaves "empty color slots" between every two previously adjacent colors (in the input scales). We now generate as many additional colors as necessary via interpolation or other techniques. [For example, for the linearization of our optimized color scale (OCS) to generate the LOCS, we applied our optimized color scales algorithm to generate between each pair of previously-adjacent colors a "mini-optimized scale," i.e., a scale that fulfills the optimization conditions required from the original optimized scale.]

2. Scanning for equally spaced colors in UCS: We scan the complete long scale resulting from steps 1(a) and 1(b) for a sequence of \( N^o \) colors that are equally distant from their adjacent colors and such that the sum of their distances equals the total distance \( d(c_{i}', c_{i}) \) from the first color to the last one in the scale. An alternative approach to minimizing \(|d(c_{i}', c_{i}) - j|\) in Algorithm LOS is to minimize \(|d(c_{i}', c_{i}) - (N^o - n^o) \cdot j|\), where \( N^o - n^o \) is the number of
steps from the last color in the long scale, \( c'_n \), to the current color \( n^\prime \). This has the advantage of avoiding the cumulative error that can be introduced by the method described previously. We keep the algorithm as described for the sake of consistency with the computations we have conducted.

3. “Shrink” the scanned scale: For each color \( c^\prime_n \) in the long scale that is selected in the scanning process, we compute a color \( c^\prime_n = \frac{M^\prime}{M^\prime + 1} \cdot c^\prime_n \) in the output scale whose \((r, g, b)\) values are \((0, 0, 0) \leq (r, g, b) \leq (M^\prime, M^\prime, M^\prime)\).

See the Appendix for the actual algorithm.

C. LinGray—the Linearized Gray Scale, and LOCS. We ran Algorithm LOS with a goal to produce a LinGray and a LOCS that would be comparable to the optimally generated scale we have used previously. In both cases we started with an input scale with \( N^\prime = 64 \) and used \( M^\prime = 767 \) (the length of the input scale was chosen such that it would be the longest possible to compute within a reasonable amount of time; the longer the input scale, the more accurate it will be). Thus, \( N^\prime = 3 \cdot (M^\prime + 1) = 2304 \). Using Algorithm LOS, we produced linearized scales with \( N^\prime = 256 \). The top of Figure 2 compares the display of a 0–255 wedge, where black is 0, white is 255, using (left to right) the gray scale; the LinGray; the OCS; the LOCS; and the linearized heated-object scale. The bottom of Figure 2 compares brain slices displayed using (left to right) the LinGray, the LOCS, and the linearized heated-object scale.

It is worth noting that the pictures shown here have passed several steps of reproduction, which have degraded their quality accuracy significantly. Thus, they should only be used for qualitative demonstration.

V. DISCUSSION AND CONCLUSIONS

We have described the concept of color scales for representation of data in general, and perceptual properties that can make such representations more accurate, and thus more effective. We have discussed one approach, the linearization of color scales. We have described briefly two previous algorithms for linearization. We have described our own algorithm for linearization, and have shown two examples of its results: LinGray and LOCS.

The main limitation of the approach presented here is the use of the CIELUV uniform color space as a measure of JNDs. The CIELUV space is considered to be reliable as a uniform space and as a measure of JNDs only within small localized neighborhoods in color space. Our approach requires the consideration of JNDs along the entire scale. Such broad range might introduce perceptual differences where equal numerical JND steps have been obtained. This would compromise the linearization process, and thus the uniformity of the resulting scale.

Figure 2. Scale comparison. Top: a 0–255 wedge, where black is 0, white is 255, using (left to right) the gray scale; the linearized version (LinGray); the original optimized color scale (OCS); the linearized optimized color scale (LOCS); and the linearized heated-object scale. Bottom: brain slices displayed using (left to right) the LinGray; the LOCS; and the linearized heated-object scale.
Further work is needed to compare these results with results obtained using other JND measures. In addition, observer performance studies are necessary to verify the uniformity of the obtained scales, and to compare them to their nonlinearized counterparts, as measured by the performance of observers using the two classes of scales.

APPENDIX
Algorithm LOS.
Input:
\[ c'_1, \ldots, c'_{N'}: \] A color scale of \( N' \) colors whose primary values are 
\[ (0, 0, 0) \leq (r, g, b) \leq (M', M', M'), \] \( (N' = M' + 1) \) such that 
\[ \epsilon(c'_{i}) = (n - 1) \] for all \( 1 \leq i \leq N' \).

\( N'' \): The length of the output scale.

Output: \( c''_1, \ldots, c''_{N''} \), a linearized optimal scale of \( N'' \) colors whose primary values are 
\[ (0, 0, 0) \leq (r, g, b) \leq (M'', M'', M''), \] \( (N'' = M'' + 1) \) such that the distance \( d(c''_{i}, c''_{i-1}) \) between the first color and the last color is divided (approximately) equally among all the color intervals.

Auxiliary Variables:
\[ c'_j, n'_b, n'_f: \] Indices to the bottom and top colors of a segment in the long scale that is determined by two adjacent colors \( c'_{i}, c'_{i+1} \) in the input scale (also used in the scanning of colors for the output scale).
\[ j: \] The desired interval in CIELUV space between two adjacent colors in the output scale.

begin
\[ c'_1 := c'_i; \]
\[ n'_b := 1; \]
for \( n'' := 1 \) to \( N'' - 1 \) do begin
\[ c := (M') - c'; \]
\[ n'' := r(c) + g(c) + b(c); \]
\[ c'_{i} := c; \]
INSERT-MISSING-COLORS: (e.g., for optimal scales: OPTIMAL-SCALES(\( c'_{i} \to c'_{i} \)))
\[ n'' := n'' \]
endfor;
\[ j := d(c''_{i}), (c''_{i})/(N'' - 1); \]
\[ c''_{i} := c''_{i} \cdot (M'') \]
end; 

ACKNOWLEDGMENTS
Part of the research reported here was conducted in collaboration with Gabor T. Herman, and was supported by NIH Grant HL 28438, while the author was with the Medical Image Processing Group, Department of Radiology, University of Pennsylvania.

REFERENCES