Choosing Color Palettes for Statistical Graphics

Achim Zeileis and **Kurt Hornik**
Wirtschaftsuniversität Wien, Austria

Abstract

Statistical graphics are often augmented by the use of color coding information contained in some variable. When this involves the shading of areas (and not only points or lines)—e.g., as in bar plots, pie charts, mosaic displays or heatmaps—it is important that the colors are perceptually based and do not introduce optical illusions or systematic bias. Here, we discuss how the perceptually-based Hue-Chroma-Luminance (HCL) color space can be used for deriving suitable color palettes for coding categorical data (qualitative palettes) and numerical variables (sequential and diverging palettes).

Keywords: qualitative palette, sequential palette, diverging palette, HCL colors, HSV colors, perceptually-based color space.

1. Introduction

Color is an integral element of graphical displays in general, and many statistical graphics in particular. Statistical software packages typically provide various color palettes and allow practitioners to employ these for visualizing data in various types of displays. However, more often than not, there is relatively little guidance about how to choose a suitable palette for a certain visualization task. Consequently, sub-optimal color palettes are used for too many color graphics displayed in journal articles or on presentation slides which could easily be enhanced if the underlying information were captured appropriately in the color schemes. Here, we address this problem by suggesting a color selection strategy for visualizing both categorical and numerical information by selecting colors along axes in a color space whose axes can be matched with perceptual axes of the human visual system.

For implementing color palettes in computer programs, typically color spaces are employed that provide a mapping to Red-Green-Blue (RGB) colors. RGB is an additive color model which is used for generating colors on computer screens (by mixing different amounts of the primary colors red, green and blue, see Poynton 2000, for more details). While being eminently useful for generating colors on computer screens, it is hard for humans to select the coordinates in RGB space corresponding to a certain color. Therefore, color picker tools in computer programs are based on different color models—starting from the pioneering introduction of Hue-Saturation-Value (HSV) colors (Smith 1978)—which try to capture the dimensions of human visual perception. HSV space is a simple transformation of RGB space which is also implemented in many statistical software packages (see Wikipedia 2006a, for more details on HSV colors and links to further information about the other color spaces discussed in the following). Although being easy to use, HSV space exhibits a number of disadvantages: It is relatively difficult to select sets of HSV coordinates that yield colors that are "in harmony" (see Munsell 1905). The reason for this is that colors with different hues also differ in saturation and brightness. For statistical graphics, this is important because it can introduce size distortions in the perception of shaded areas and color-caused optical illusions (Cleveland and McGill 1983). Furthermore, the use of HSV colors encourages the use of flashy and highly saturated colors which are good for drawing attention to a plot but hard to look at for a longer time.

Despite these known drawbacks, HSV space is still the predominantly-used color model in color picker tools (Moretti and Lyons 2002; Meier, Spalter, and Karelitz 2004) and the basis for many palettes in statistical software packages. A notable exception is **ColorBrewer.org** (Harrower and Brewer 2003), an online tool for selecting color schemes for maps. It provides a rich collection of prefabricated palettes (with a fixed maximal number of colors) and guides its users how to choose a suitable palette for coding various types of information. Here, we take a similar approach and describe strategies for the choice of color palettes for categorical and numerical data. Following Brewer (1999) and Harrower and Brewer (2003), we distinguish three types of palettes: qualitative, sequential and diverging. The first is tailored for coding categorical information and the latter two are aimed at numerical variables (or ordinal variables if some numerical coding for its levels is used). Unlike **ColorBrewer.org**, we do not only provide fixed sets of colors but suggest a general principle for selecting colors by traversing paths along perceptual axes in a suitable color space. Consequently, the user can decide which path exactly should be taken and how many colors should be selected. The color space employed by us is the Hue-Chroma-Luminance (HCL) color space (see Ihaka 2003) and we show how the general principles suggested can be turned into formulas describing paths of coordinates in this space. Given a mapping from HCL to RGB coordinates, our formulas are extremely easy to implement in software and we provide such an implementation in the R language (R Development Core Team 2006) using the powerful R graphics system (see Murrell 2006) and the HCL color implementation from package **colorspace** (Ihaka 2006). This gives the user both a conceptual and computational tool box for experimenting with color palettes for a particular display.

The remainder of the paper is organized as follows: Section 2 provides several motivating examples, showing how typical HSV-based graphics can be enhanced by using HCL-based palettes. Section 3 gives a brief introduction to the underlying HSV and HCL color spaces before Section 4 suggests strategies for deriving HCL-based qualitative, sequential, and diverging palettes. Section 5 offers some general remarks on the implementation in statistical software packages as well as some details on our implementation in R. Section 6 concludes the paper with a discussion.

2. Motivation

To show what can be gained by selecting appropriate color schemes, we present a collection of illustrations with typical color graphics based on HSV palettes (always in the left panels) and more suitable HCL palettes (always in the right panels). All HSV color palettes (and some of the data sets) are taken from recent publications in statistical journals such as *Journal of the American Statistical Association*, *Journal of Computational and Graphical Statistics* or *Computational Statistics & Data Analysis*. The examples have been selected to provide an overview of typically colored statistical graphics and their respective pitfalls (concerning color palette choice).

Our first illustration is a heatmap, a very popular display for visualizing a scalar function of two arguments. Here, we use a bivariate kernel density estimate (for the Old Faithful geyser eruptions data from Azzalini and Bowman 1990)—other typical applications include objective functions with two arguments (as in Gneiting, Ševčíková, Percival, Schlather, and Jiang 2006) or (physical) measurements on a 2-dimensional grid (as in Yang, Buckley, Dudoit, and Speed 2002). Figure 1 shows such a heatmap bringing out the relationship between the duration of an eruption of the Old Faithful geyser in Yellowstone National Park and the waiting time for this eruption. It reveals a multi-modal distribution: short waiting times (around 50 minutes) are typically followed by a long eruption (around 4 minutes) whereas long waiting times (around 80 minutes) can be followed by either a long or short eruption (around 4 minutes).

A simple and very effective palette for such a display is a set of gray colors as in the top right panel of Figure 1. This is often (appropriately) used in printed papers when the journal does not offer color graphics—however, in journals that support color graphics (or on presentation slides and in interactive usage in statistical software packages), many users prefer to have colored displays and most often use HSV palettes (as in the two left two panels). These palettes code the variable of interest by varying hue in an HSV color wheel as done by Yang *et al.* (2002, Figure 4a) or

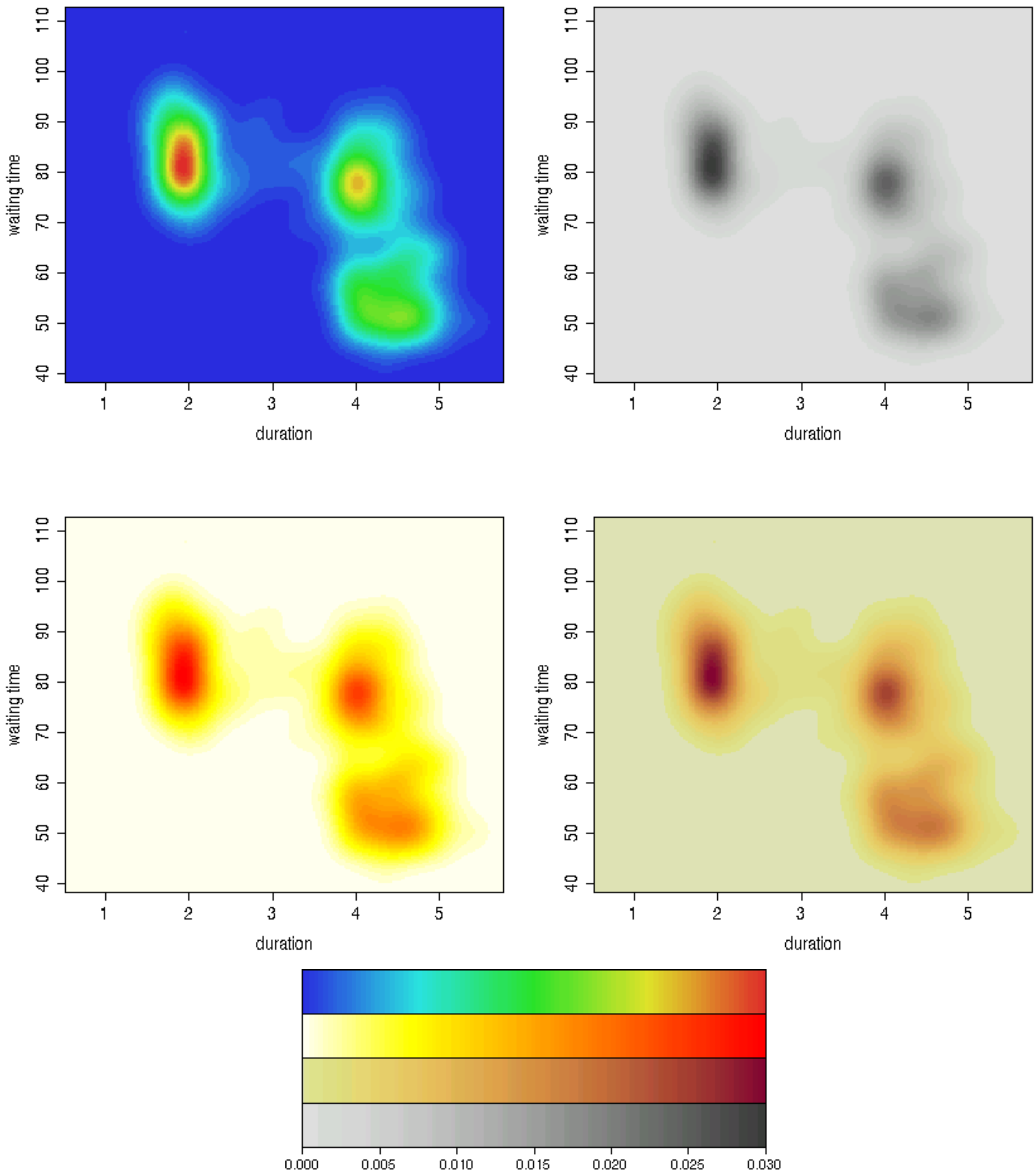

Figure 1: Bivariate density estimation for Old Faithful geyser eruptions.

Gneiting *et al.* (2006, Figures 1–4). The palette in the upper left panel codes increasing density by going from a blue to a red hue (via green and yellow)—a similar strategy are the "heat colors" in the lower left panel that increase from yellow to red. The latter works a bit better than the former, however both palettes exhibit several drawbacks. The modes in the map look much more like "rings" rather than a smoothly increasing/decreasing density. The heatmap looks very flashy which—although good for drawing attention to a plot—makes it hard to hold the attention of the viewer for a longer time because the large areas shaded with saturated colors can be distracting and produce after-image effects (Ihaka 2003).

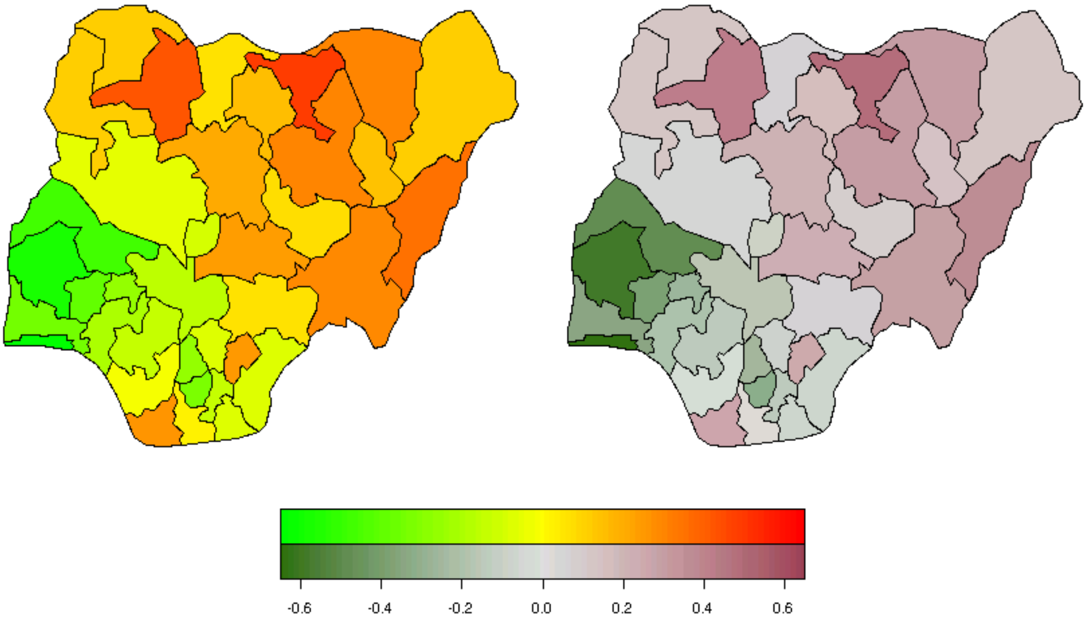

Figure 2: Posterior mode estimates for childhood mortality in Nigeria.

In contrast, the gray colors used in the top right panel do not exhibit the same disadvantages, coding the variable of interest much better and without flashy colors. If, however, the user wants to add some color to the plot, this could be done by using a better balanced version of the heat colors (as shown in the bottom right panel). These colors also increase from a yellow to a red hue while being balanced towards the same gray as in the grayscale palette (i.e., when converted to a grayscale or printed out on a grayscale printer, the upper and lower right panel would essentially look identical). Both palettes have in common that they give increasing perceptual emphasis to regions with increasing density, resulting in a heatmap that highlights the (small) interesting high-density regions and not to the large low-density regions surrounding them.

A similar example is presented in Figure 2 (taken from Kneib 2006, Figure 5, left), depicting posterior mode estimates for childhood mortality in different regions of Nigeria. Deviations from a model for childhoold mortality are brought out by shading a map according to the corresponding residuals, revealing decreased mortality in the south-west and increased mortality in the north-east. Kneib (2006) uses an HSV-based palette coding the residuals by the hue, going from green via yellow to red. Our HCL-based palette also employs green and red hues for negative and postive residuals respectively, but codes neutral values (around 0) by a neutral light gray. Compared to the HSV-based palette, this offers again a number of advantages: only the interesting areas are highlighted by high-chroma colors; flashy colors are avoided making it easier to look at the display for a longer time; positive and negative residuals with the same absolute size receive the same perceptual weight by being balanced towards the same gray.

Although problematic for many tasks, pie charts can be useful for visualizing whether a set of pie segments constitutes a majority. A typical application is shown in Figure 3, visualizing the distribution of seats in the German parliament “Bundestag” following the 2005 election. In this election, five parties were able to obtain enough votes to enter the Bundestag—however, neither the governing coalition of SPD and Grüne nor the opposition of CDU/CSU and FDP could assemble a majority. Given that no party would enter a coalition with the leftists “Die Linke”, this lead to a big coalition of CDU/CSU and SPD. In graphical displays, the parties are usually matched by using colors as metaphors: red for the social democrats SPD, black for the conservative CDU/CSU,

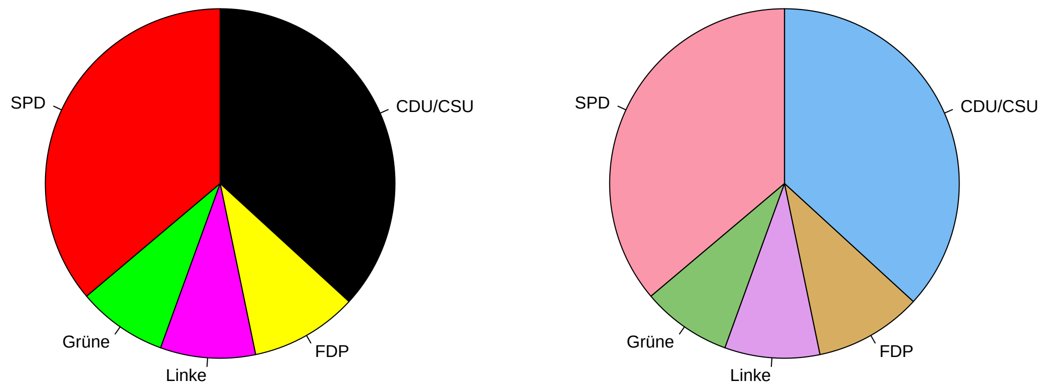

Figure 3: Seats in the German parliament.

yellow for the liberal FDP, green for the green party "Die Grünen" and purple for the leftist party "Die Linke". The left panel shows fully saturated HSV colors as usually found in the (German) media whereas the right panel uses less flashy HCL colors with the same hues (except for the CDU/CSU where a blue hue instead of the extreme "color" black is used). The advantage of the latter is again that they are easier to hold in focus for a longer time. Furthermore, they are all balanced towards the same gray and have the same amount of color, resulting in a perceptually balanced palette that does not introduce undesired graphical distractions. While it could be argued that this pie chart is such a simple display that a well-balanced palette is not so important and that flashy colors are to be preferred, the balancing properties of the HCL-based palette are very important in more complex displays such as the mosaic display in Figure 4. This re-uses the same colors and shows the distribution of votes in the 2005 election stratified by province (Bundesland). The order of provinces is from north to south, first for the 10 western provinces (the former Federal Republic of Germany, FRG), then for the 6 eastern provinces (the former German Democratic Republic, GDR). Clearly, the SPD performed better in the north and the CDU/CSU better in the south; furthermore, Die Linke performed particularly well in the eastern provinces and in Saarland.

Color choice is usually much more important in graphical displays with shaded areas compared to displays with only points or lines. However, color choice becomes much more relevant also in scatter plots when there are many points. As an example, Figure 5 depicts a scatter plot with three clusters similar to the one shown in Celeux, Hurn, and Robert (2000, Figure 3). To indicate cluster membership, three different colors (green, yellow, red) are employed. The HSV colors in the left panel are again very flashy and differ strongly with respect to luminance: the yellow is much lighter and hardly visible. In contrast, the HCL colors in the right panel use the same hues, but are balanced with respect to chroma and luminance, i.e., the amount of color and gray.

As our final example, Figure 6 (taken from Friendly 2002, Figure 11, left/middle) visualizes the cross-tabulation of hair end eye color of 592 students in a mosaic display. Clearly, hair and eye color are not independent and the pattern of association is highlighted by means of residual-based shading. Cells associated with Pearson residuals whose absolute value exceeds (2 or) 4 are shaded (light) blue/red. This shows that there are significantly more students with black hair and brown eyes, blond hair and blue eyes, red hair and green eyes than expected under independence. Conversely, fewer students than expected have blond hair and brown/hazel eyes or black hair and blue eyes. Comparing the HSV and HCL colors in Figure 6, it is shown again that the HCL colors are better balanced (between the red and blue colors) and less flashy. Even the cells associated with small residuals (below 2) are a bit easier to read when shaded in a light grey rather than white. More details and extensions to residual-based shadings can be found in Zeileis, Meyer, and Hornik (2007).

Figure 4: Votes in the German election 2005 (by province).

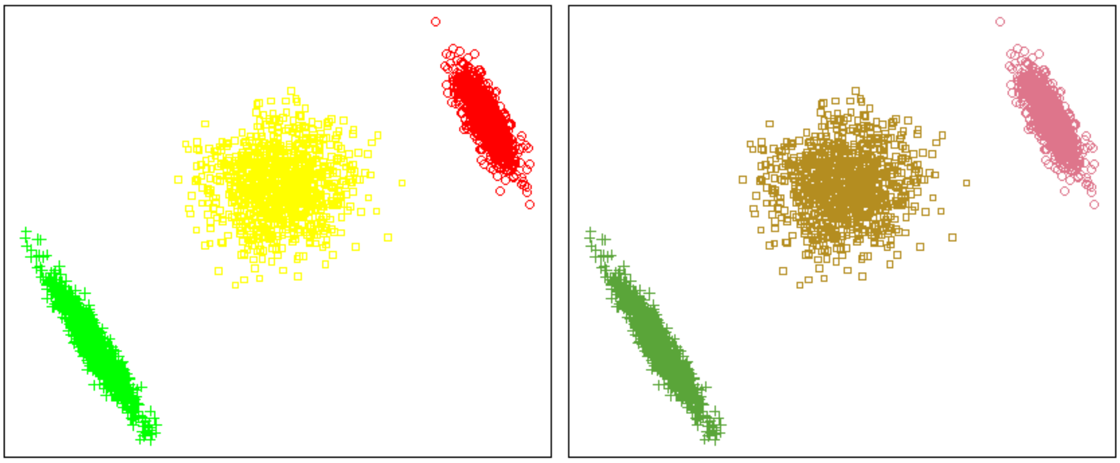

Figure 5: Scatter plot with three clusters.

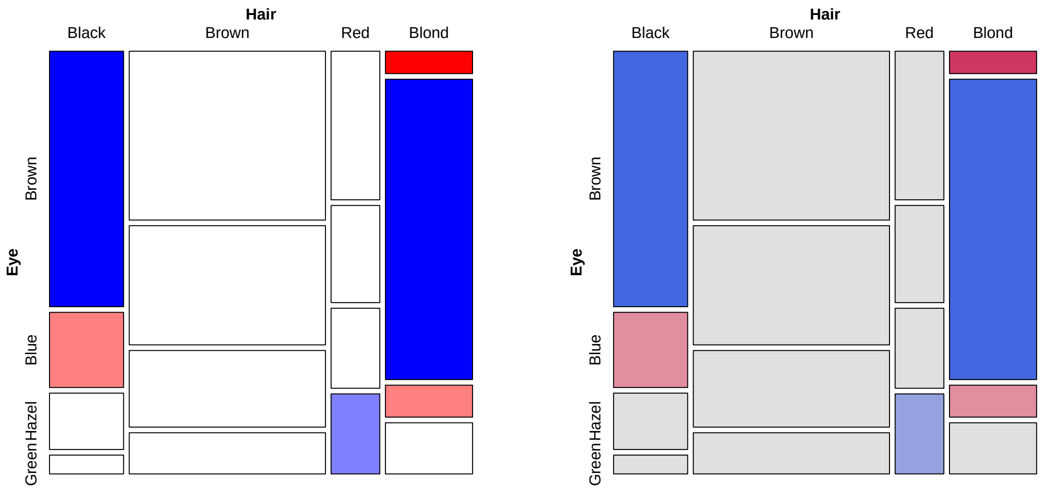

Figure 6: Extended mosaic display for haireye data.

3. Color spaces

For choosing color palettes it is helpful to have an idea how human color vision evolved. It has been hypothesized that it developed in three distinct stages: 1. perception of *light/dark* contrasts (monochrome only), 2. *yellow/blue* contrasts (usually associated with our notion of warm/cold colors), 3. *green/red* contrasts (helpful for assessing the ripeness of fruit). See Ihaka (2003) for more details and references.

Due to these three color axes, colors are typically described as locations in 3-dimensional spaces. However, human perception of color does not correspond to the physiological axes above, but rather to polar coordinates in the color plane (yellow/blue vs. green/red) plus a third light/dark axis. Thus, perceptually-based color spaces try to capture the perceptual axes

- *hue* (dominant wavelength)
- *chroma* (colorfulness, intensity of color as compared to gray)

- *luminance* (brightness, amount of gray)

A popular implementation of such a color space, available in many graphics and statistics software packages, are HSV colors. They are a simple transformation of RGB colors and are defined by a triplet (H, S, V) with $H \in [0, 360]$ and $S, V \in [0, 100]$. Although simple to specify and easily available in many computing environments, HSV colors have a fundamental drawback: its three dimensions map to the three dimensions of human color perception very poorly. The three dimensions are confounded: The brightness of colors is not uniform over hues and saturations (given value)—therefore, HSV colors are often not considered to be perceptually based.

To overcome these drawbacks, various color spaces have been suggested that properly map to the perception dimensions, the most prominent of which are the CIELUV and CIELAB spaces developed by the Commission Internationale de l'Éclairage (2004). Ihaka (2003) argues that CIELUV colors are typically preferred for use with emissive technologies such as computer screens which makes them an obvious candidate for implementation in statistical software packages. By taking polar coordinates in the UV plane of CIELUV, HCL colors are obtained, defined by a triplet (H, C, L) with $H \in [0, 360]$ and $C, L \in [0, 100]$. Given a certain luminance L, all colors resulting from different combinations of hue H and chroma C are balanced towards the same gray (and look the same when converted to a gray scale). However, the admissible combinations of chroma and luminance coordinates (within the space's boundaries) depend on the hue chosen. The reason for this is that some hues lead to light and others to dark colors, e.g., full chroma yellow is brighter (i.e., has higher luminance) than full chroma blue.

The balancing of HCL colors gives us the opportunity to conveniently choose color palettes which code categorical and/or numerical information by translating it to paths along the three perceptual axes. However, some care is required for dealing with the irregular shape of the HCL space which will be addressed in the following sections.

4. Color palettes

4.1. Qualitative palettes

Qualitative palettes are sets of colors for depicting different categories, i.e., for coding a categorical variable. Usually, these should give the same perceptual weight to each category so that no group is perceived to be larger or more important than any other one. Typical applications of qualitative palettes in statistics would be bar plots (see Ihaka 2003), pie charts (see Figure 3) or highlighted mosaic displays (see Figure 4).

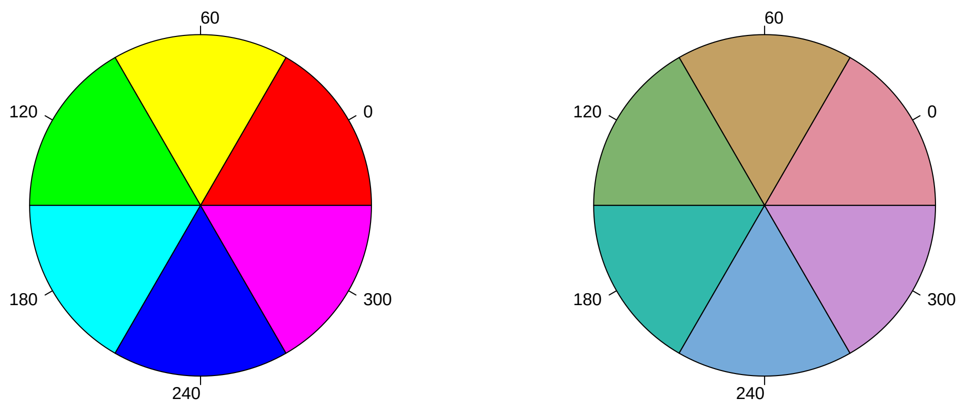

Figure 7: HSV-based and HCL-based color wheel.

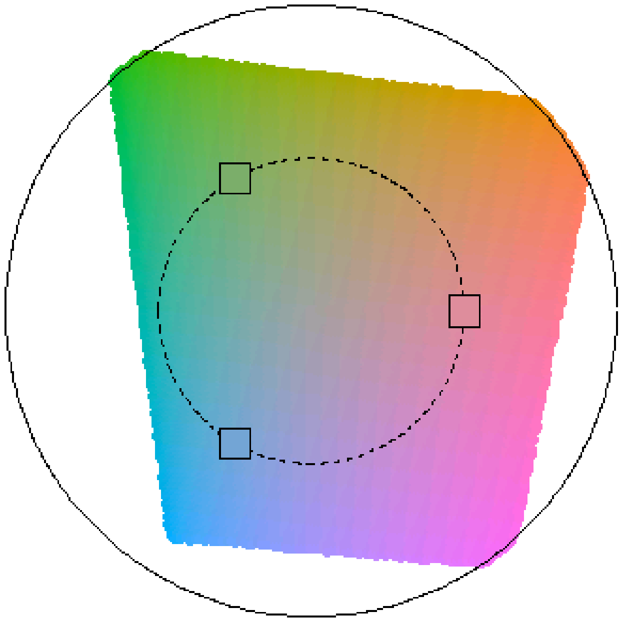

Figure 8: Constructing qualitative palettes. In the hue/chroma plane for $L = 70$, the dashed circle correponds to a radius $C = 50$ with chosen angles $H = 0, 120, 240$.

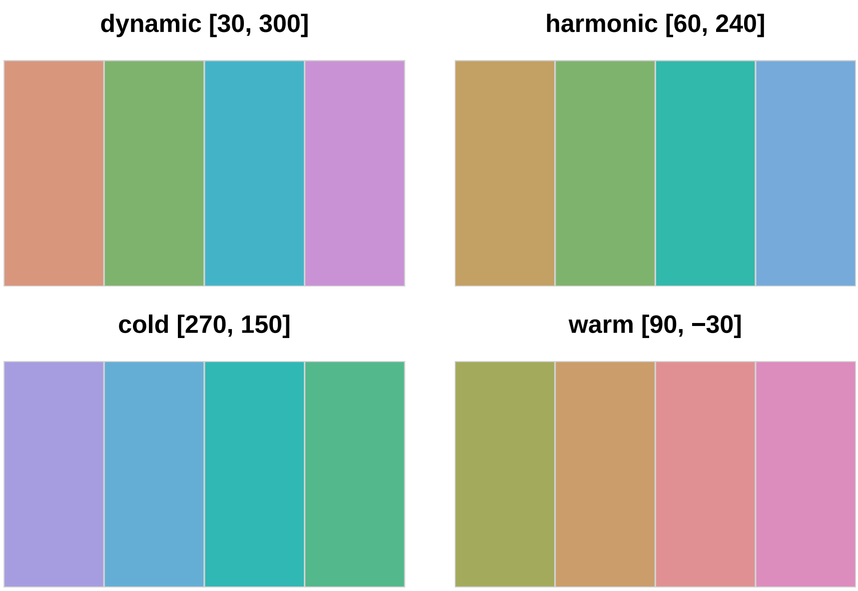

Figure 9: Examples for qualitative palettes. Hue is varied in different intervals for given $C = 50$ and $L = 70$.

Ihaka (2003) describes a simple strategy for choosing such palettes: chroma and luminance are kept fixed and only the hue is varied for obtaining different colors which are consequently all balanced towards the same gray. This effect is illustrated in Figure 7 which shows a color wheel obtained by varying the hue only in HSV coordinates $(H, 100, 100)$ and HCL coordinates $(H, 50, 70)$. Clearly, not only the hue but also the amount of chroma and luminance varies for the HSV wheel.

Figure 8 depicts how the HCL-based color wheel is constructed. It shows the hue/chroma plane of HCL space given a luminance of $L = 70$. Not all combinations of hue and chroma are admissible, however for a chroma of $C = 50$ a full color wheel can be obtained. For choosing the hues in a certain palette, various strategies are conceivable. A simple and intuitive one is to use colors as metaphors for categories (e.g., for political parties as in Figures 3 and 4), another approach would be to use segments from the color wheel corresponding to nearby or distant colors. The latter is shown in Figure 9 which depicts examples for generating qualitative sets of colors $(H, 50, 70)$. In the upper left panel colors from the full spectrum are used $(H = 30, 120, 210, 300)$ creating a 'dynamic' set of colors. The upper right panel shows a 'harmonic' set with $H = 60, 120, 180, 240$. Warm colors (from the blue/green part of the spectrum: $H = 270, 230, 190, 150$) and cold colors (from the yellow/red part of the spectrum: $H = 90, 50, 10, 330$) are shown in the lower left and right panel, respectively.

4.2. Sequential palettes

Sequential palettes are used for coding numerical information that ranges in a certain interval where low values are considered to be uninteresting and high values are interesting. Suppose we need to visualize an intensity or interestingness i which (without loss of generality) is scaled to the unit interval. A typical application in statistics are heatmaps (see Figure 1).

The simplest solution to this task is to employ light/dark contrasts, i.e., rely on the most basic perceptual axis. The interestingness can be coded by an increasing amount of gray corresponding to decreasing luminance in HCL space:

$$(H, 0, 90 - i \cdot 60),$$

where the hue H used does not matter, chroma is set to 0 (i.e., no color), and luminance ranges in $[30, 90]$ avoiding the extreme colors white $(L = 100)$ and black $(L = 0)$. Instead of going linearly from light to dark gray, luminance could also be increased nonlinearly, e.g., by some function $i' = f(i)$ that controls whether luminance is increased quickly with intensity or not. We found $i' = f(i) = i^p$ to be a convenient transformation where the power p can be varied to achieve different degrees of non-linearity.

Furthermore, the intensity i could additionally be coded by colorfulness (chroma), e.g.,

$$(H, 0 + i' \cdot C_{\max}, L_{\max} - i' \cdot (L_{\max} - L_{\min})).$$

This strategy is depicted in the left panel of Figure 10 for a blue hue $H = 260$ and different combinations of maximal chroma ($C_{\max} = 0$, 80 and 100, respecitvely) and minimal luminance ($L_{\min} = 30$, 30 and 50, respectively). The first two combinations are also shown in the first two rows of Figure 11. The right panel of Figure 10 shows that the exact same strategy is not possible for the green hue $H = 120$. While the gray colors without chroma can be chosen in the same way, there is a stronger trade-off between using dark colors (with low luminance) and colorful colors (with high chroma). Hence, the second path from light gray to full green ends at a much lighter color with $L_{\min} = 75$.

To increase the contrast between the colors in the palette even further, the ideas from the previous sequential palettes can also be combined with qualitative palettes by simultaneously varying the hue as well:

$$(H_2 - i \cdot (H_1 - H_2), C_{\max} - i' \cdot (C_{\max} - C_{\min}), L_{\max} - i'' \cdot (L_{\max} - L_{\min})).$$

One application is an HCL-based version of "heat colors" that increase from a light yellow (e.g., $(90, 30, 90)$) to a full red (e.g., $(0, 100, 50)$). To make the change in hue visible, the chroma needs to

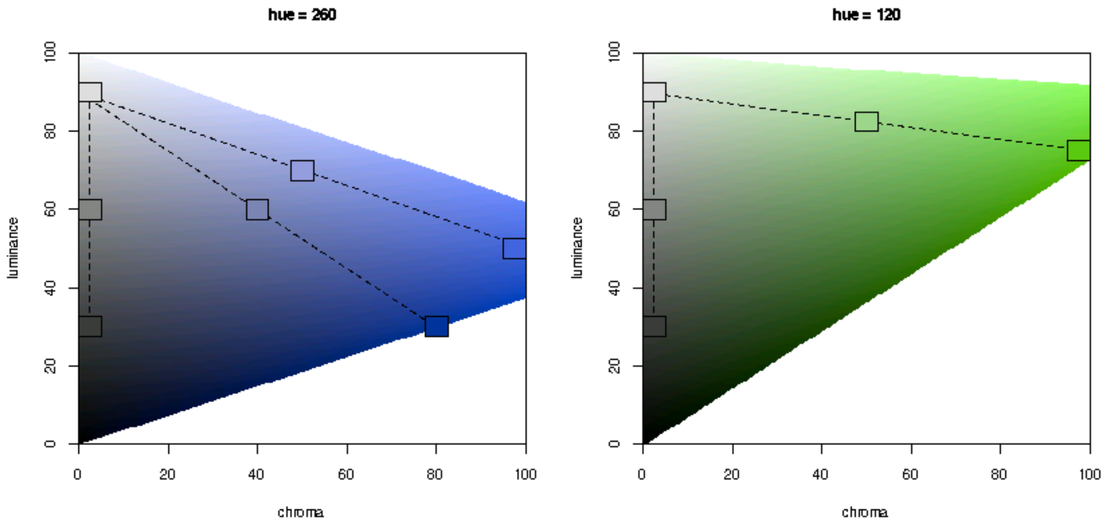

Figure 10: Constructing sequential palettes. The chroma/luminance plane is shown for two hues $H = 260$ (left) and $H = 120$ (right). Colors are chosen by varying either only luminance or both luminance and chroma.

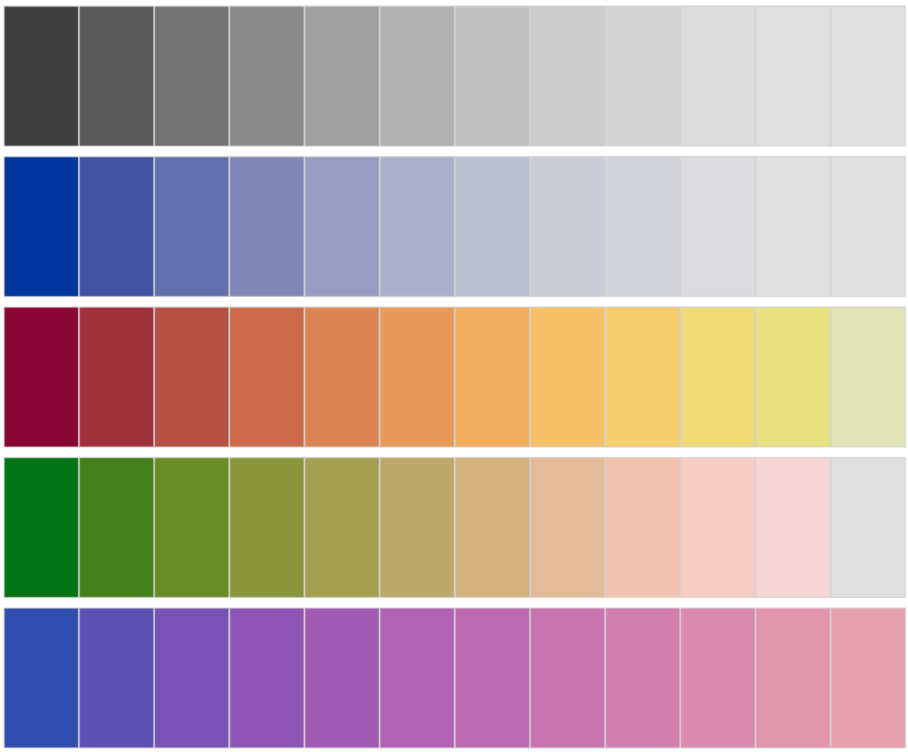

Figure 11: Examples for sequential palettes, varying only luminance (first panel), chroma and luminance (second panel), and hue, chroma and luminance (remaining panels).

increase rather quickly for low values of i and then only slowly for higher values of i. This can be achieved by choosing an appropriate transformation i' for chroma and a different transformation i'' for the luminance. Such a strategy is adopted for the palettes shown in the lower three rows in Figure 11 using different pairs of hues as well as different chroma and luminance contrasts.

4.3. Diverging palettes

Diverging palettes are also used for coding numerical information ranging in a certain interval—however, this interval includes a neutral value. Examples for this include residuals or correlations (both with the neutral value 0) or binary classification probabilities (with neutral value 0.5) that could be visualized in mosaic plots (see Figure 6), classification maps or model-based shading in maps such as Figure 2. Analogously to the previous section, we suppose that we want to visualize an intensity or interestingness i from the interval $[-1, 1]$ (without loss of generality).

Given useful sequential palettes, deriving diverging palettes is easy: two different hues are chosen for adding color to the same amount of 'gray' at a given intensity $|i|$. Figure 12 shows the chroma/luminance plane back to back for the hues $H = 0$ and 260 with two different paths—giving slightly different emphasis on luminance or chroma contrasts—from a full red over a neutral grey to a full blue. As Figure 10 illustrates, the pair of hues should be chosen carefully because the admissible values in the chrome/luminance plane differ across hues. Clearly, for deriving symmetric palettes, only colors from the intersection of the admissible chroma/luminance planes can be used. The particular hues $H = 0$ and 260 used in Figure 12 were chosen because they correspond to similar geometric shapes in the chroma/luminance plane, allowing for both large chroma and luminance contrasts. If potential viewers of the resulting graphic might be color-blind, the pair of hues should be taken from the yellow/blue axis of the color wheel rather than the green/red axis as contrasts on the latter axis are more difficult to distinguish for color-blind people (Lumley 2006).

Figure 13 shows various examples of conceivable combinations of hue, chroma and luminance. The first palette uses a broader range on the luminance axis whereas the others mostly rely on chroma contrasts.

5. Software

Implementing the different color palettes suggested in the previous section is extremely easy if the software environment chosen already provides an implementation of HCL colors: from the formulas provided above the HCL coordinates for a palette can be conveniently computed. A bit more work is required if the software package does not yet provide an HCL implementation. In that case, additional functionality is needed for translating HCL coordinates to the software package's color system which may vary between different packages, but standardized RGB (sRGB) is often used. The typical way of coordinate conversion is to go first from HCL to CIELUV by simply transforming the polar H and C coordinates back to the original U and V. Subsequently, CIELUV is converted to CIEXYZ which in turn is converted to sRGB (with the latter conversion depending on the device used for display). The details of these conversions are somewhat technical and tedious (and hence omitted here), however the conversion formulas are still straightfoward to implement and can, for example, be found in Wikipedia (2006b) or Poynton (2000).

The R system for statistical computing (R Development Core Team 2006) already comes with an open-source implementation of HCL (and other color spaces) in the package **colorspace** (Ihaka 2006). The coordinate transformations mentioned above are contained in C code within **colorspace** that are easy to port to other statistical software systems. Based on **colorspace**, the package **vcd** (Meyer, Zeileis, and Hornik 2006) provides implementations of all palettes discussed above. Qualitative palettes are provided by `rainbow_hcl()` (named after the HSV-based function `rainbow()` in base R). Sequential palettes based on a single hue are implemented in the function `sequential_hcl()` while `heat_hcl()` offers sequential palettes based on a range of hues. Diverging palettes can be obtained by `diverge_hcl()`. Technical documentation along with a

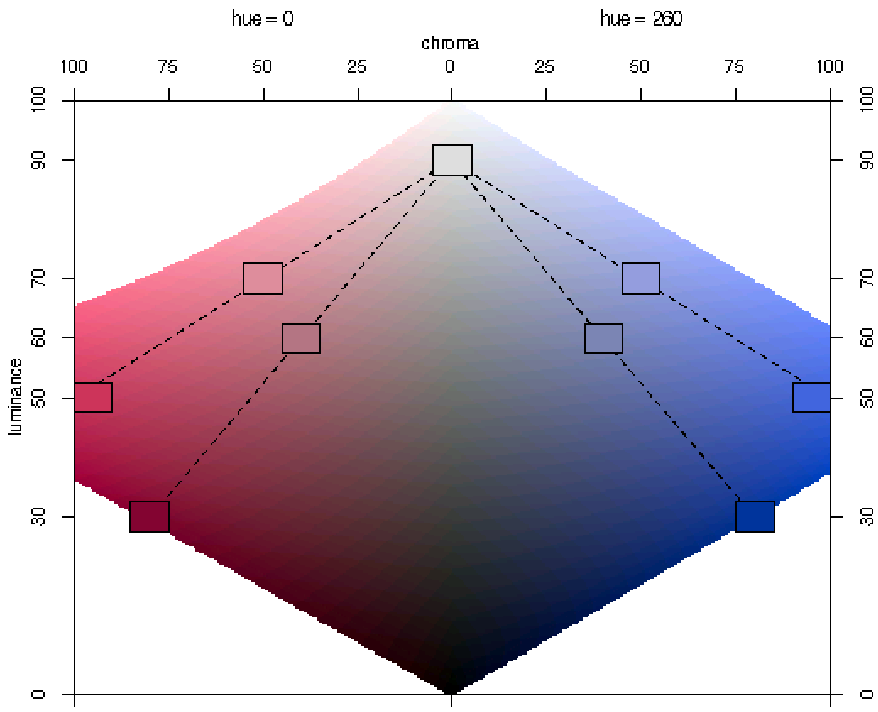

Figure 12: Constructing diverging palettes. The chroma/luminance plane is shown back to back hues $H = 0$ and $H = 260$. Colors are chosen by simultaneously varying luminance and chroma.

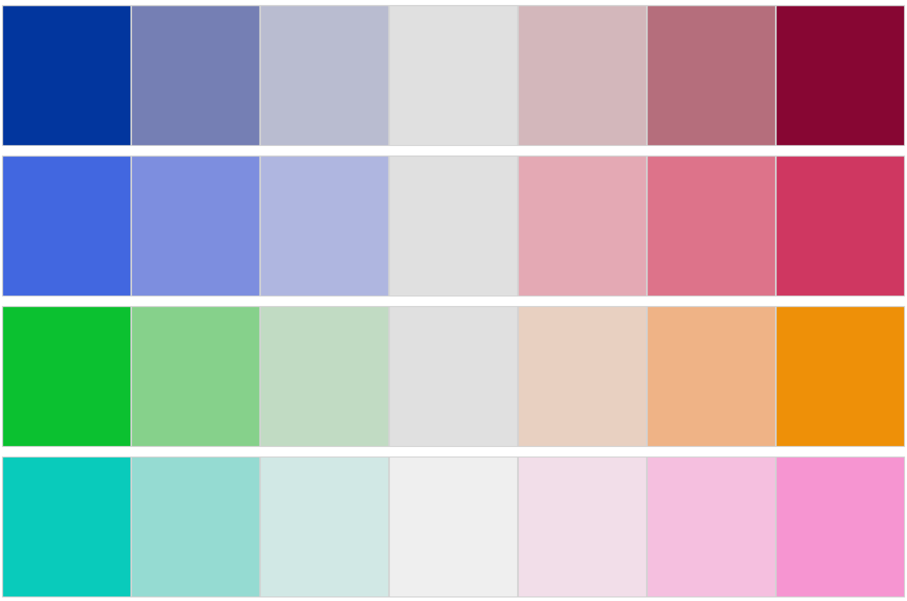

Figure 13: Examples for diverging palettes with different pairs of hues and decreasing luminance contrasts.

large collection of examples is available via `help("rainbow_hcl")`. R code for reproducing Figures 9, 11 and 13 (and some illustrations) can be accessed via `vignette("hcl-colors", package = "vcd")`.

6. Discussion

Many statistical graphics—especially when displayed on a computer screen, e.g., as in interactive usage, electronic papers or presentation slides—employ colors to code information about a certain variable. Despite this omnipresence of color, there is often only little guidance in statistical software packages on how to choose a palette appropriate for a particular visualization task—auspicious tools such as **ColorBrewer.org** notwithstanding. We try to address this problem by suggesting color schemes for coding categorical information (qualitative palettes) and numerical information (sequential and diverging palettes) based on the perceptually-based HCL color space.

We provide paths through HCL space along perceptual axes so that colors selected along these paths match perceptual dimensions. This gives the users the possibility to conveniently experiment with the HCL-based palettes by varying several graphical parameters. For qualitative palettes, these are the coordinates on the chroma and luminance axis, respectively, controlling whether the colors are light or dark and how colorful they are. For sequential and diverging palettes, the user can decide whether contrasts in the chroma or luminance direction (or both) should be employed. In our experience (as illustrated in Section 2), chroma contrasts work sufficiently well if a small set of colors is used. However, when a larger set of colors is used (e.g., for heatmaps where extreme values should be identifiable) it is much more important to have a big difference in luminance.

Based on these conceputal guidelines and the computational tools readily provided in the R system for statistical computing (and easily implemented in other statistical software packages), users can generate palettes varying these graphical parameters and thus adapting the colors to their particular graphical display.

Acknowledgements

We are thankful to David Meyer, Paul Murrell, Brian D. Ripley, Thomas Kneib, and Michael Höhle for feedback, suggestions and discussions.

References

Azzalini A, Bowman AW (1990). "A Look at Some Data on the Old Faithful Geyser." *Applied Statistics*, **39**, 357–365.

Brewer CA (1999). "Color Use Guidelines for Data Representation." In "Proceedings of the Section on Statistical Graphics, American Statistical Association," pp. 55–60. Alexandria, VA.

Celeux G, Hurn M, Robert CP (2000). "Computational and Inferential Difficulties with Mixture Posterior Distributions." *Journal of the American Statistical Association*, **95**(451), 957–970. Figure 3.

Cleveland WS, McGill R (1983). "A Color-caused Optical Illusion on a Statistical Graph." *The American Statistician*, **37**, 101–105.

Commission Internationale de l'Éclairage (2004). *Colorimetry*. Publication CIE 15:2004, Vienna, Austria, 3rd edition. ISBN 3-901-90633-9.

Friendly M (2002). "A Brief History of the Mosaic Display." *Journal of Computational and Graphical Statistics*, **11**(1), 89–107. Figure 11 (left, middle).

Gneiting T, Ševčíková H, Percival DB, Schlather M, Jiang Y (2006). "Fast and Exact Simulation of Large Gaussian Lattice Systems in $\mathbb{R}^2$: Exploring the Limits." *Journal of Computational and Graphical Statistics*, **15**(3), 483–501. Figures 1–4.

Harrower MA, Brewer CA (2003). "**ColorBrewer.org**: An Online Tool for Selecting Color Schemes for Maps." *The Cartographic Journal*, **40**, 27–37.

Ihaka R (2003). "Colour for Presentation Graphics." In K Hornik, F Leisch, A Zeileis (eds.), "Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria," ISSN 1609-395X, URL http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/.

Ihaka R (2006). **colorspace**: *Colorspace Manipulation*. R package version 0.95.

Kneib T (2006). "Mixed Model-based Inference in Geoadditive Hazard Regression for Interval-censored Survival Times." *Computational Statistics & Data Analysis*, **51**, 777–792. Figure 5 (left).

Lumley T (2006). "Color Coding and Color Blindness in Statistical Graphics." *ASA Statistical Computing & Graphics Newsletter*, **17**(2), 4–7.

Meier BJ, Spalter AM, Karelitz DB (2004). "Interactive Color Palette Tools." *IEEE Computer Graphics and Applications*, **24**(3), 64–72.

Meyer D, Zeileis A, Hornik K (2006). **vcd**: *Visualizing Categorical Data*. R package version 1.0-2.

Moretti G, Lyons P (2002). "Tools for the Selection of Colour Palettes." In "Proceedings of the New Zealand Symposium On Computer-Human Interaction (SIGCHI 2002)," University of Waikato, New Zealand.

Munsell AH (1905). *A Color Notation*. Munsell Color Company, Boston, Massachusetts.

Murrell P (2006). *R Graphics*. Chapmann & Hall/CRC, Boca Raton, Florida.

Poynton C (2000). "Frequently-Asked Questions About Color." URL http://www.poynton.com/ColorFAQ.html. Accessed 2006-09-14.

R Development Core Team (2006). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, URL http://www.R-project.org/.

Smith AR (1978). "Color Gamut Transform Pairs." *Computer Graphics*, **12**(3), 12–19. ACM SIGGRAPH 78 Conference Proceedings.

Wikipedia (2006a). "HSV Color Space — Wikipedia, The Free Encyclopedia." URL http://en.wikipedia.org/w/index.php?title=HSV_color_space&oldid=74735552. Accessed 2006-09-14.

Wikipedia (2006b). "Lab Color Space — Wikipedia, The Free Encyclopedia." URL http://en.wikipedia.org/w/index.php?title=Lab_color_space&oldid=72611029. Accessed 2006-09-14.

Yang YH, Buckley MJ, Dudoit S, Speed TP (2002). "Comparison of Methods for Image Analysis on cDNA Microarray Data." *Journal of Computational and Graphical Statistics*, **11**(1), 108–136. Figure 4a.

Zeileis A, Meyer D, Hornik K (2007). "Residual-based Shadings for Visualizing (Conditional) Independence." *Journal of Computational and Graphical Statistics*. Forthcoming, URL http://statmath.wu-wien.ac.at/~zeileis/papers/Zeileis+Meyer+Hornik-2007.pdf.