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 an appropriate palette for a certain visualization task. 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 2004). 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 further information on HCL and other color spaces, contrasting in particular the properties of HSV and HCL colors. Section 3 introduces HCL-based qualitative, sequential, and diverging palettes and Section 4 shows some brief illustrations for each type of palette in statistical graphics. Section 5 provides 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. 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

1. *hue* (dominant wavelength)

2. *chroma* (colorfulness, intensity of color as compared to gray)

3. *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 $S, V \in [0, 100]$. HSV space has the shape of a single regular cone (often inflated to a regular cylinder). Vertical sections through this space are shown in the upper panel of Figure 1, depicting hue and saturation given different value levels. 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 which is most easily seen when converting the vertical sections to gray scale images in Figure 2. Clearly, the brightness of colors is not uniform over hues and saturations (given value)—therefore, HSV colors are often not considered to be perceptually based.
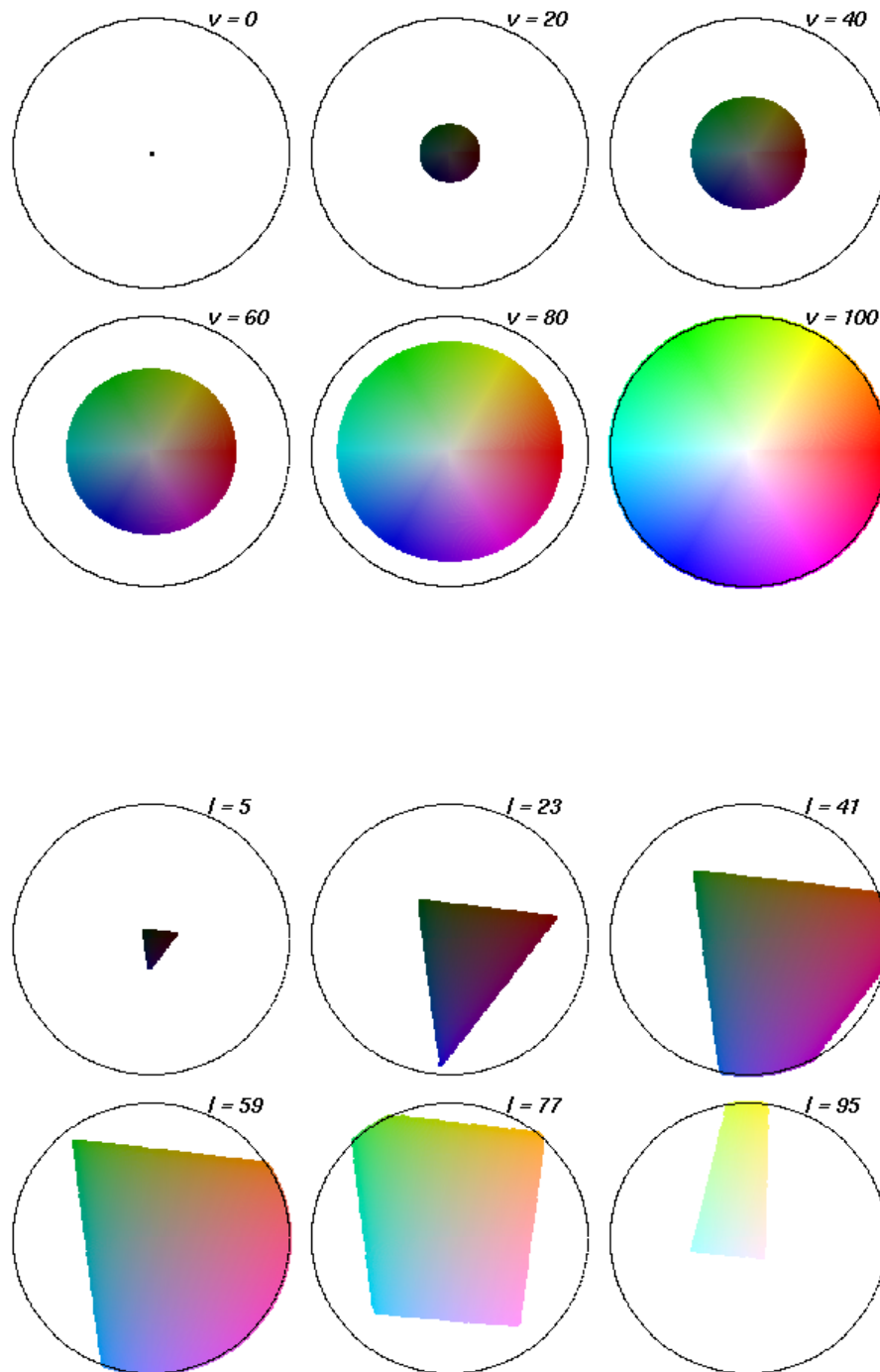
Figure 1: HSV and HCL space. The circles show the hue/saturation and hue/chroma plane, respectively, for varying levels of value and luminance.
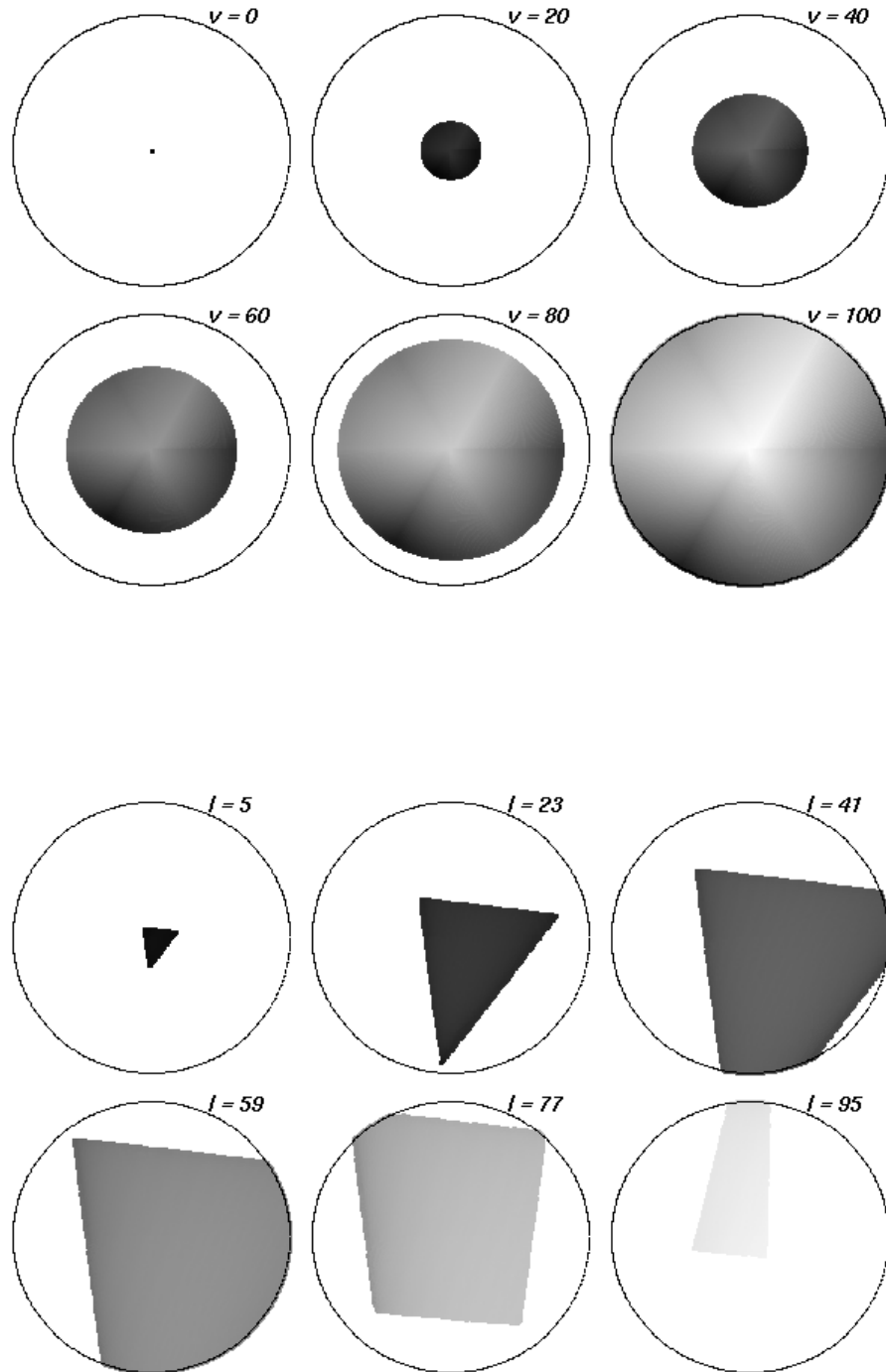
Figure 2:   HSV and HCL space (in gray levels).   The circles show the hue/saturation and hue/chroma plane, respectively, for varying levels of value and luminance.

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 $C, L \in [0, 100]$. HCL space has the shape of a distorted double cone: the admissible combinations of chroma and luminance coordinates (within the space's boundaries) depend on the hue chosen. The lower panel of Figure 1 shows vertical sections through this space: each of the resulting hue/chroma planes (given luminance) is now properly balanced towards the same gray (going from black to white with increasing luminance) which becomes obvious when converting the colors to a gray scale as in Figure 2. This 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.

3. Color palettes

3.1. 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 9) or highlighted mosaic displays (see Figure 10).

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. As Figure 1 illustrates, the range of hues available depends on the combination of chroma and luminance chosen. Figure 3 depicts how three colors are chosen, given $C = 50$ and $L = 70$.

Various strategies for choosing the hues in a certain palette are conceivable. A simple and intuitive one is to use colors as metaphors for categories (e.g., for political parties), another approach would be to use segments from the color wheel corresponding to nearby or distant colors. The latter is shown in Figure 4 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.

3.2. 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 11).

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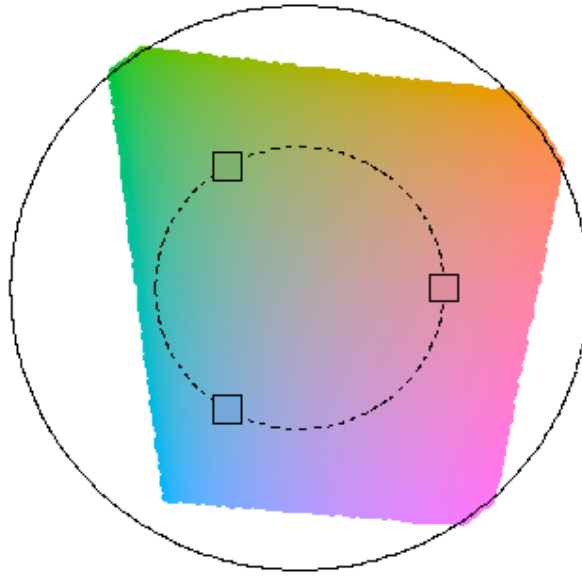
$$(H, 0, 90 - i \cdot 60),$$

Figure 3: Constructing qualitative palettes. In the hue/chroma plane for $L = 70$, the dashed circle correponds to a radius $C = 50$ with chosen angles $H = 0, 120, 240$.
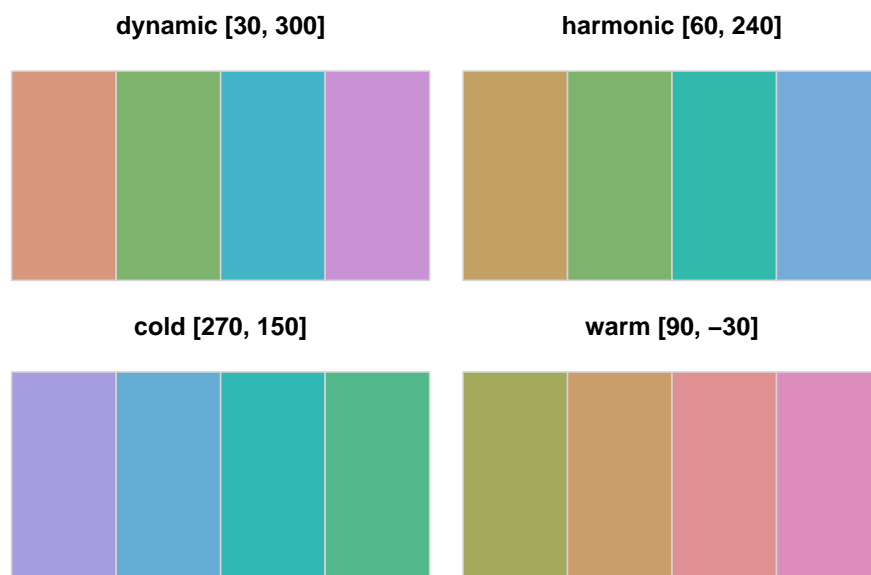

Figure 4: Examples for qualitative palettes. Hue is varied in different intervals for given $C = 50$ and $L = 70$.

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 $f(i)$ that controls whether luminance is increased quickly with intensity or not. We found $f(i) = i^p$ 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.,

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

This strategy is depicted in the left panel of Figure 5 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 6. The right panel of Figure 5 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:

$$(H_2 - i \cdot (H_1 - H_2), C_{\max} - i^{p_1} \cdot (C_{\max} - C_{\min}), L_{\max} - i^{p_2} \cdot (L_{\max} - L_{\min})).$$

One application would be 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 increase rather quickly for low values of i and then only slowly for higher values of i. This can be achieved by choosing a power $p_1 < 1$. Such a strategy is adopted for the palettes shown in the lower three rows in Figure 6 using different pairs of hues as well as different chroma and luminance contrasts. The exact parameter combinations are provided in the appendix.

3.3. 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 12) or classification maps (see Figure 13). 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 7 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 5 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 7 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.

Figure 8 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.
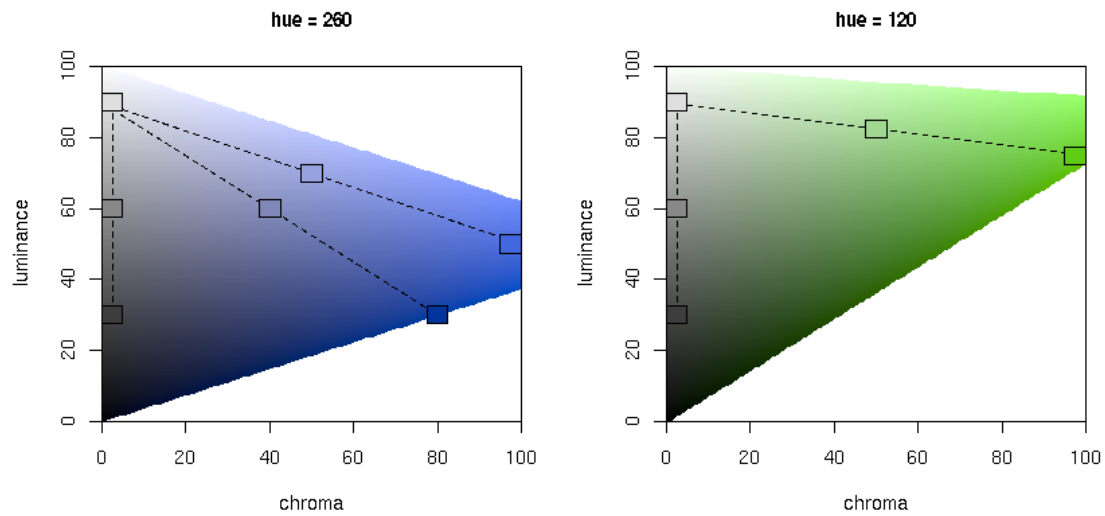
Figure 5: Constructing sequential palettes. The chroma/luminance plane is shown for two hues $H = 260$ (left) and $H = 120$ (right). Colors are chosen by varying either only luminance or both luminance and chroma.
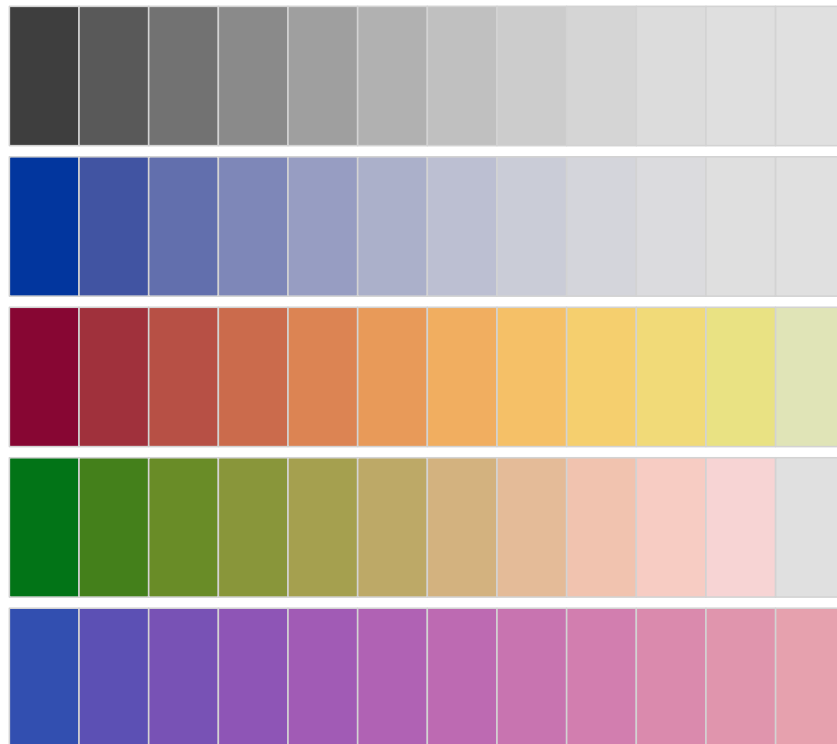

Figure 6: Examples for sequential palettes, varying only luminance (first panel), chroma and luminance (second panel), and hue, chroma and luminance (remaining panels).
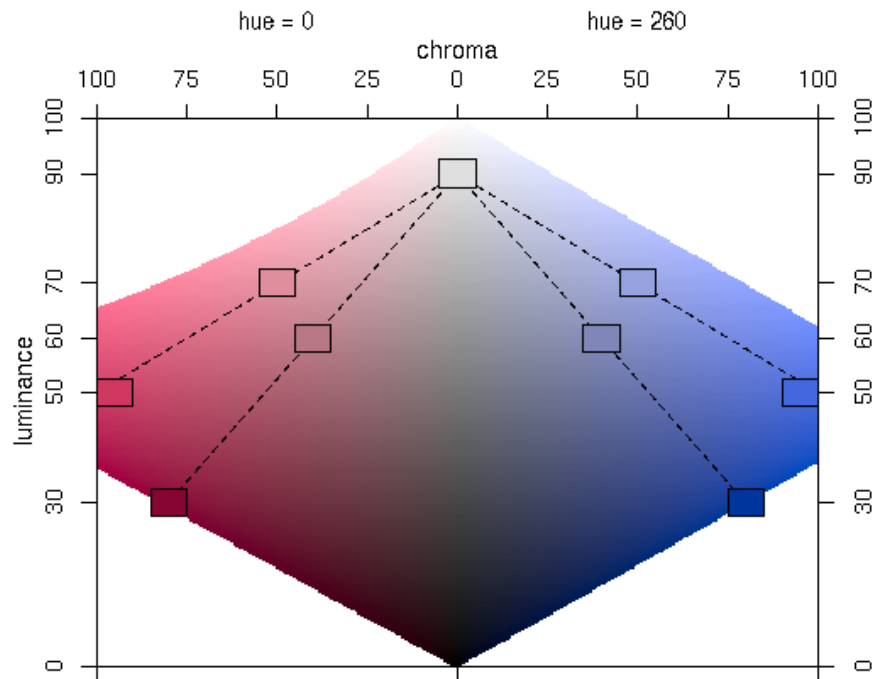
Figure 7: Constructing diverging palettes. The chroma/luminance plane is shown back to back hues $H = 0$ and $H = 260$. Colors are chosen by simultaneously varying luminance and chroma.

Figure 8: Examples for diverging palettes with different pairs of hues and decreasing luminance contrasts.

4. Illustrations

In this section, we show a collection of examples for the various types of palettes applied to statistical graphics. The first example illustrates qualitative palettes and visualizes data from the 2005 election for the German parliament "Bundestag". In this election, five parties were able to obtain enough votes to enter the Bundestag—the distribution of seats is depicted in a pie chart in Figure 9. The colors used are rough metaphors for the political parties, using a red hue $H = 0$ for the social democrats SPD, a blue hue $H = 240$ for the conservative CDU/CSU, a yellow hue $H = 60$ for the liberal FDP, a green hue $H = 120$ for the green party "Die Grünen" and a purple hue $H = 300$ for the leftist party "Die Linke". To obtain rather intense colors, we set chroma to $C = 60$ and luminance to $L = 75$ for all hues leading to a much better balanced palette compared to the fully saturated colors usually found in the (German) media (which instead of blue typically use the extreme "color" black for the CDU/CSU). The pie chart clearly shows that 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, this lead to a big coalition of CDU/CSU and SPD. Figure 10 shows the distribution of votes in this election stratified by province (Bundesland) in a highlighted mosaic display. 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.

To illustrate sequential palettes, a bivariate density estimation for the Old Faithful geyser eruptions data (Azzalini and Bowman 1990) is visualized. The Old Faithful geyser is one of the most popular sites in Yellowstone National Park and it is of some interest to understand the relation ship between the duration of a geyser eruption and the waiting time for this eruption. Figure 11 shows (two versions of) a heatmap of a bivariate kernel density estimate of duration and waiting time revealing a multi-modal bivariate 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). Both versions of the heatmap use a sequential palette as derived in Section 3.2 balanced towards the same gray levels with $L \in [30, 90]$ and $p_2 = 1.5$. The sequential palette in the left panel uses only gray colors (i.e., $C_{\max} = 0$) and the palette in the right panel additionally employs colors with $H \in [0, 90]$, $C \in [30, 80]$ and $p_1 = 1/5$. The parameters have been chosen such that the peaks in the density are brought out very clearly by using a rather large luminance contrast and $p_2 > 1$.

As pointed out in Section 3.3, diverging palettes are particularly useful when visualizing residuals or correlations (with natural neutral value 0) or probabilities in 2-class supervised learning (with neutral value 0.5). Examples for both situations are provided here. Figure 12 visualizes the outcome of a double-blind clinical trial investigating a new treatment for rheumatoid arthritis. The mosaic rectangles alone signal that the treatment lead to higher improvement compared to the placebo group; this effect is shown to be significant by the shading that codes the size of the Pearon residuals. Positive residuals, corresponding to more observations in the corresponding cell than expected under independence, are depicted in blue, negative residuals in red. Light colors signal significance at 10% level, full colors significance at 1% level. Hence, it can be concluded that there are significantly more marked improvements in the treated group and significantly fewer in the placebo group than would be expected under independence between treatment and improvement. More details can be found in Zeileis, Meyer, and Hornik (2005).

Figure 13 shows a heatmap with the fit of a support vector machine (SVM) to an artificial 2-class supervised learning example: a mixture of two bivariate normal distributions with different means and covariance matrices. The circles and triangles show the original observations, solid symbols correspond to the support vectors found. The shading underlying the plot visualizes the fitted decision values: values around 0 are on the decision boundary and are shaded in light gray, while regions that are firmly classified to one or the other class are shaded in full blue and red respectively.
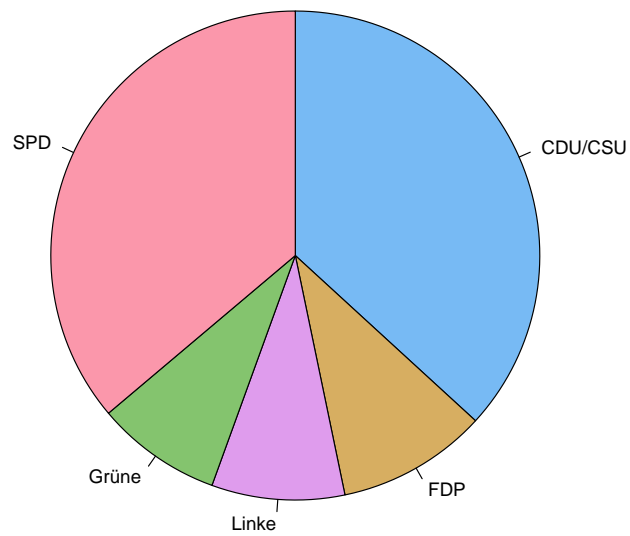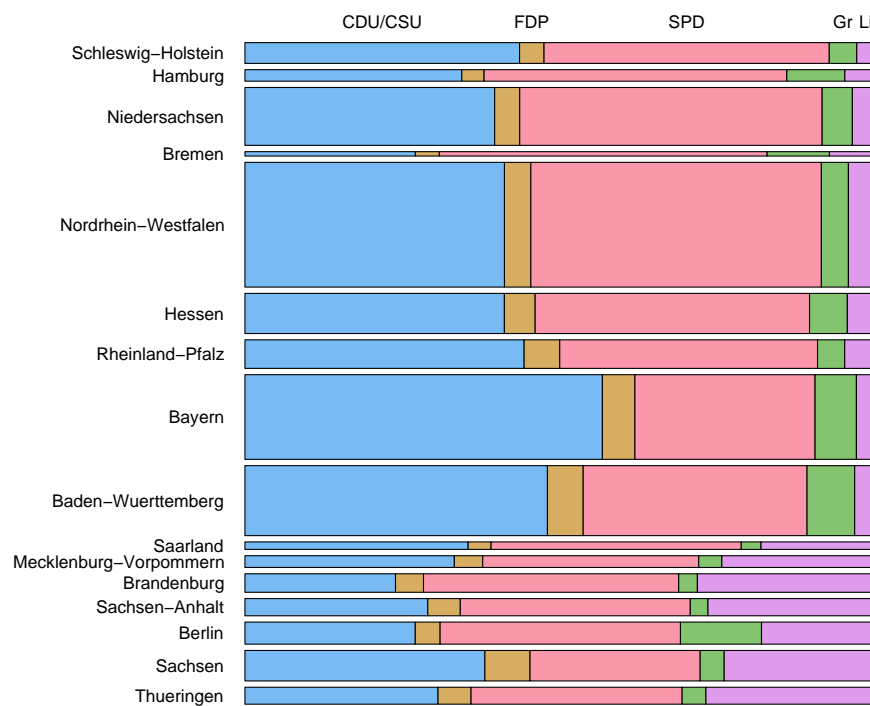
Figure 9: Seats in the German parliament.

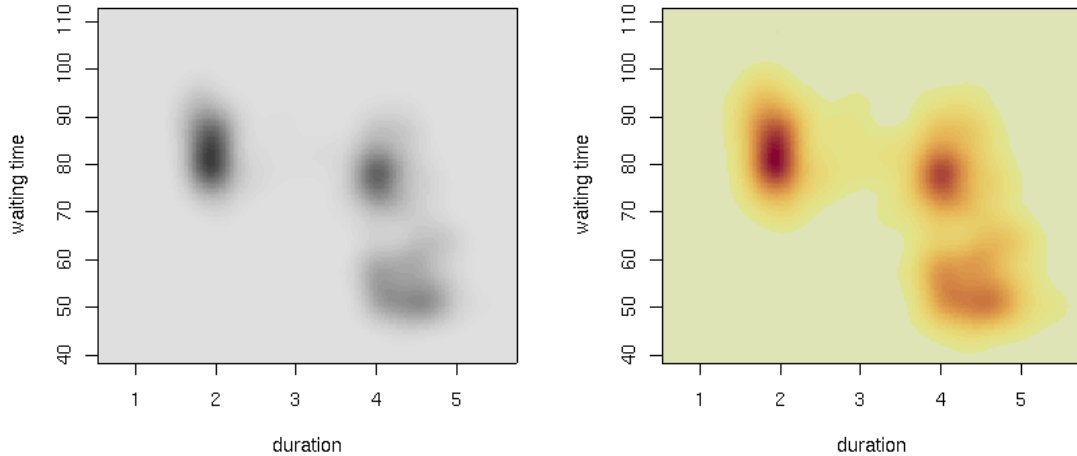Figure 10: Votes in the German election 2005.

Figure 11: Bivariate density estimation for Old Faithful geyser eruptions.

Both diverging palettes employ a smaller contrast on the luminance axis ($L \in [50, 90]$) and a larger chroma contrast ($C \in [0, 100]$) compared to the sequential palettes from the previous example. For Figure 13, we set p_1 and p_2 to the somewhat smaller value of 1.3 to bring out all deviations from the decision boundary and not only the extreme decision values. For Figure 12, the choice of p_1 and p_2 is less important because only 5 colors are selected—we use values $p_1 = p_2 = 1$.

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 2004) which we build upon for implementing all palettes described above. The functions outlined below are provided by the package **vcd** (Meyer, Zeileis, and Hornik 2006)—technical documentation for the functions along with a large collection of examples is available via `help("rainbow_hcl")`. These examples also include comparisons between existing R palettes (based on HSV colors) and the new HCL color palettes.

Qualitative palettes are implemented in **vcd** in the function

```
rainbow_hcl(n, c = 50, l = 70, start = 0, end = 360*(n-1)/n, ...)
```
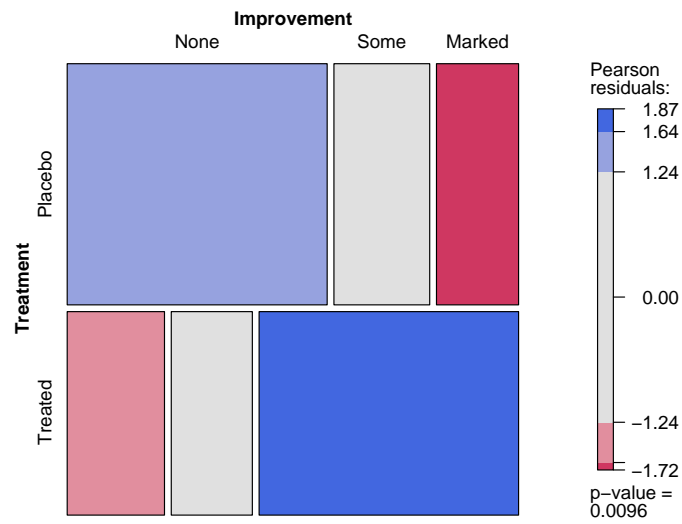
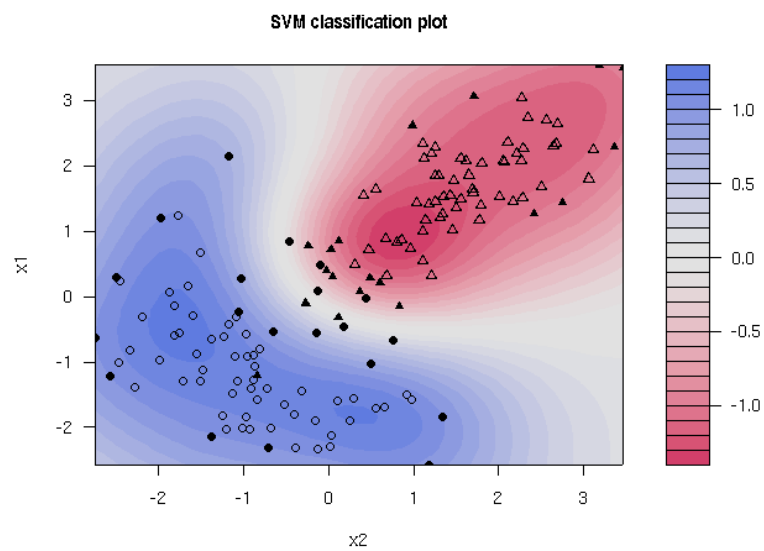Figure 12: Extended mosaic display for arthritis data.

Figure 13: SVM classification plot.

where `n` controls the number of colors in the palette. The arguments `c` and `l` give the fixed chroma and luminance levels, respectively, and `start` and `end` specify the range of hue angles. The function is named after the base R function `rainbow()` which has a similar interface but chooses colors in HSV coordinates. It computes the HCL coordinates and transforms them to RGB by using the `hcl2hex()` function which is a convenience wrapper to the function `polarLUV()` from the **colorspace** package.

Sequential palettes based on a single hue are implemented in the function

```
sequential_hcl(n, h = 260, c = c(80, 0), l = c(30, 90), power = 1.5, ...)
```

where the first element of `c` and `l` give the starting chroma and luminance coordinate (by default colorful and dark) and the second element the ending coordinate (by default light gray). The `power` argument implements the parameter p from the i^p function. Sequential palettes using a range of hues are provided by

```
heat_hcl(n, h = c(0, 90), c = c(100, 30), l = c(50, 90), power = c(1/5, 1), ...)
```

named after the HSV-based R function `heat.colors()` and by default starts from a red and going to a yellow hue. The defaults in `heat_hcl()` are set differently compared to `sequential_hcl()` as to make the default HCL heat colors more similar to the HSV version. The defaults of `sequential_hcl()`, on the other hand, are set as to achieve a large contrast on the luminance axis.

Diverging palettes are implemented in the function

```
diverge_hcl(n, h = c(260, 0), c = 80, l = c(30, 90), power = 1.5, ...)
```

which has the same arguments as `sequential_hcl()` but takes a pair of hues `h`.

The R code for reproducing Figures 4, 6 and 8 is given in the appendix. A vignette that additionally explains how the figures from the illustrations section can be reproduced is contained in **vcd** and 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 4), 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. Another parameter in these palettes is the rate at which the intensity $|i|$ is increased from 0 to 1: Employing a function $|i|^p$ for this, small powers p can be used if palettes with a lot of color should be constructed (e.g., to clearly separate regions in a classification map). On the other hand, large powers p will result in palettes with less color (e.g., appropriate to highlight only extreme regions firmly classified in such a map).

Based on these conceptual 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

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.

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.

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 (2004). **colorspace**: *Colorspace Manipulation*. R package version 0.9.

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 1.0-0.

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.

Zeileis A, Meyer D, Hornik K (2005). "Residual-based Shadings for Visualizing (Conditional) Independence." *Report 20*, Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Research Report Series. URL http://epub.wu-wien.ac.at/dyn/openURL?id=oai:epub.wu-wien.ac.at:epub-wu-01_871.

A. Supplementary R code

All palettes discussed above are implemented in the R system for statistical computing and graphics (R Development Core Team 2006) in the package **vcd** (Meyer *et al.* 2006). The functions are based on the `polarLUV(l, c, h)` function proveded by the **colorspace** package (Ihaka 2004). A simple convenience function for displaying a certain palette is

```
pal <- function(col, border = "light gray", ...)
{
  n <- length(col)
  plot(0, 0, type = "n", xlim = c(0, 1), ylim = c(0, 1),
       axes = FALSE, xlab = "", ylab = "", ...)
  rect(0:(n-1)/n, 0, 1:n/n, 1, col = col, border = border)
}
```

The qualitative palettes in Figure 4 are generated via

```
pal(rainbow_hcl(4, start =  30, end = 300))
pal(rainbow_hcl(4, start =  60, end = 240))
pal(rainbow_hcl(4, start = 270, end = 150))
pal(rainbow_hcl(4, start =  90, end = -30))
```

The sequential palettes in Figure 6 are generated via

```
pal(sequential_hcl(12, c = 0, power = 2.2))
pal(sequential_hcl(12, power = 2.2))
pal(heat_hcl(12, c = c(80, 30), l = c(30, 90), power = c(1/5, 2)))
pal(terrain_hcl(12, c = c(65, 0), l = c(45, 90), power = c(1/2, 1.5)))
pal(heat_hcl(12, h = c(0, -100), l = c(75, 40), c = c(40, 80), power = 1))
```

The diverging palettes in Figure 8 are generated via

```
pal(diverge_hcl(7))
pal(diverge_hcl(7, c = 100, l = c(50, 90), power = 1))
pal(diverge_hcl(7, h = c(130, 43), c = 100, l = c(70, 90)))
pal(diverge_hcl(7, h = c(180, 330), c = 59, l = c(75, 95)))
```

Further examples and technical information on all palettes is available from the corresponding online documentation.