ColorCAT: Guided Design of Colormaps for Combined Analysis Tasks

Sebastian Mittelstädt, Dominik Jäckle, Florian Stoffel, Daniel A. Keim

University of Konstanz

Abstract

Colormap design is challenging because the encoding must match the requirements of data and analysis tasks as well as the perception of the target user. A number of well-known tools exist to support the design of colormaps. ColorBrewer [HB03], for example, is a great resource to select colors for qualitative, sequential, and diverging data. PRAVDAColor [BRT95] and Tominski et al. [TFS08], for example, provide valuable guidelines for single analysis tasks such as localization, identification, and comparison. However, for solving real world problems in most practical applications, single elementary analysis tasks are not sufficient but need to be combined. In this paper, we propose a methodology and tool to design colormaps for combined analysis tasks. We define color mapping requirements and develop a set of design guidelines. The visualization expert is integrated in the design process to incorporate his/her design requirements, which may depend on the application, culture, and aesthetics. Our ColorCAT tool guides novice and expert designers through the creation of colormaps and allows the exploration of the design space of color mapping for combined analysis tasks.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications

1. Introduction & Related Work

The elementary analysis tasks of data visualizations are *localization*, *identification*, and *comparison* of data values [AA06], which corresponds to the *search* tasks *locate* and *browse* to *identify* and *compare* data values in the multi-level typology of Brehmer and Munzner [BM13]. This was introduced for color mapping strategies by Tominski *et al.* [TFS08], who focuses rather on different data transformations than on designing colormaps. The challenge is that different tasks have different requirements for the visual encoding. For example, *comparing* data values requires that perceived distances match data distances. This is typically accomplished with unipolar colormaps that do not vary over different hues. These colormaps are the results of todays tools for continuous (sequential) data [BRT95, HB03, WVVW*08]. However, these colormaps are insufficient in the task of *identifying* data values (e.g., read metric quantities) because they do not provide many distinct colors [War88]. Color scales that are effective in *identification* must vary over multiple hues [War88], but this distorts perceptual linearity and biases the analysts in the *comparison* task. The complexity for designing colormaps increases if tasks are combined, e.g., to *identify* and *compare* data values, which is a typical task in real applications.

Most of the existing guidelines and tools are data-driven [War88, HB03, Bre15]. There exist also task-driven guidelines and tools [BRT95, Rhe00] but they focus on single tasks. We argue that this is not enough, since real analysis tasks typically require the combination of different elementary tasks. There exist algorithms [LH92, Kei00, KRC02, WGM*08, LSS12, MBS*14] for sophisticated colormaps that may cover single task combinations. However, these algorithms are based on complex color spaces and optimization problems. The colormap designer has no influence on the outcome of optimized results that, e.g., may lack in aesthetics [WGM*08] but also may not be in-line with the mental model of domain experts since the ordering of colors depends on culture and domain. This results in inappropriate colormap selection since there is no available tool that supports designers in the creation of colormaps for their analysis task. Further, colormaps for color-blind persons require additional strategies [Oli13, SMO*13] or recoloring methods [KOF08]. Mittelstädt *et al.* [MBS*14] defined data-dependent quality metrics for mapping data relations to color and Bernard *et al.* [BSM*15] perceptual-metrics for static 2D colormaps. Since both focus on encoding data relations and not on encoding single data dimensions, the approaches do not define

The definitive version is available at http://diglib.eg.org/ .
DOI: 10.2312/eurovisshort.20151135

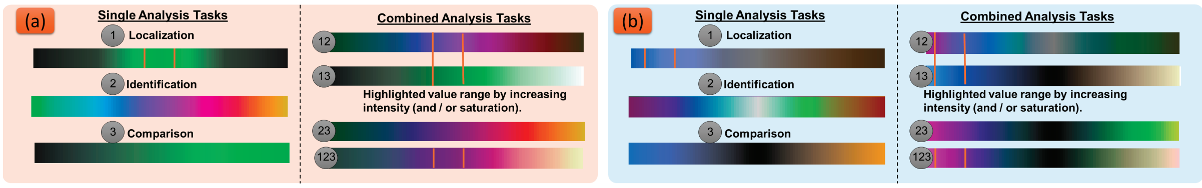

Figure 1: *Colormaps for analysis tasks (1 localization, 2 identification, 3 comparison) and their combinations for sequential **(a)** and diverging **(b)** data. The colormaps vary perceptual linear over hue, saturation, and intensity according to the task combination. All colormaps presented here are color-blind safe besides 2, which maximizes JNDs for normal color vision.*

requirements and quality metrics for the elementary analysis tasks to *identify*, *localize*, and *compare* data values.

We see a gap of defining the color mapping requirements for elementary analysis tasks and their combinations; and further, we see the need for a tool that guides designers through the design space of colormaps. Therefore, we provide *ColorCAT*, which is a tool to support visualization experts in the design of colormaps. The designer has to specify the data-type (sequential or diverging) and the analysis task combination. *ColorCAT* determines the requirements for the colormap according to the specified data and task properties, and automatically generates a suitable (color-blind safe) colormap. The designer may modify the colormap according to the application but also in terms of culture and aesthetics.

In this paper, we claim the following contributions: 1) A definition of requirements for different analysis tasks and their combinations; 2) Quality metrics for one dimensional colormaps to support these requirements and; 3) provide color-blind safe color maps for each task combination; 4) We contribute *ColorCAT* for guided design of colormaps.

2. Color Mapping Requirements

The challenge is that the different tasks have conflicting requirements for colormaps. We extend the guidelines [TFS08, MBS*14, BSM*15] by defining colormap requirements and quality metrics for (combined) elementary analysis tasks and provide means for high qualitative colormaps. Figure 1 shows examples of colormaps created with *ColorCAT*. The encircled numbers link colormaps to the task combinations.

1 **R1-Localization.** This task is performed when the analysts wants to see “where” specific objects are *located* within the data [BM13], e.g., *visual query for the value 100 on the display*. Therefore, data values and ranges of high importance must be perceptually striking in the visualization (e.g., highlighting). To provide an appropriate color mapping, the visual importance VI of a color i must encode the data importance DI (Eq. 1). Studies showed that visual attention is predominantly steered by intensity and saturation [CYG04]. Thus, VI can be approximated by the arc of intensity I and saturation S (in the HSI color space [Kei00]), which is inline with the approach of Guo *et al.* [GGMZ05], Bernard *et al.* [BSM*15], and results of ColorBrewer [HB03].

$$QM_1 = \sum_i |DI(i) - VI(i)| \qquad VI(i) = \sqrt{I_i^2 + S_i^2} \tag{1}$$

QM_1 can be minimized by selecting one color hue and increase in intensity and saturation according to the specified data importance, e.g., from black to green (sequential) or blue and orange (diverging). *ColorCAT* lets the user specify the data importance by interactive spline charts and models intensity and saturation accordingly (see Section 3.4).

2 **R2-Identification.** This task is performed when the analyst *browses* or *explores* the data and reads values from color encoded objects on the screen [BM13], e.g., *estimate the value of the upper-left object*. A high number of perceptually distinct colors (JND, just-noticeable-difference [MEO94]) allows accurate identification of data values [War88, MSK14]. The task requires that the number of JNDs is maximized but all colors share equal visual importance (R1) to avoid the typical harmful effects of “rainbow” colormaps [RTB96] such as attention steering effects or intensity gaps. In order to measure the amount of JNDs, a colormap can be segmented such that the colors within each segment are perceptually equal to the centroid of the segment ($\Delta E(c_1, c_2) < JND$), but perceptually distinct ($\Delta E(c_1, c_2) > JND$) to the centroid of other segments (equivalent to *MacAdam ellipses* [Mac42] in color spaces or within a 2D colormap [BSM*15]). The number of segments corresponds to the number of JNDs.

$$\Delta E = \sqrt{(\Delta J/K_L)^2 + \Delta a^2 + \Delta b^2} \text{ see [LCL06]} \tag{2}$$

Thus, sequential colormaps must vary perceptually linear over the full range of hues with maximized and equalized (R1) intensity and saturation. An alternative to increase the number of distinct colors with equal visual importance is increasing intensity while decreasing saturation (Figure 2). This is recommended for color-blind persons for whom mixing red and green tones must be omitted. Accordingly, diverging colormaps increase in saturation but decrease in intensity.

Figure 2: *Colorblind-safe colormaps for identification.*

3 **R3-Comparison (absolute differences)** is about comparing two or more visual encoded objects [BM13] and to perceive the relative and absolute differences. This task requires that distances in data space are equal to perceived distances in the visual encoding [RTB96]. The

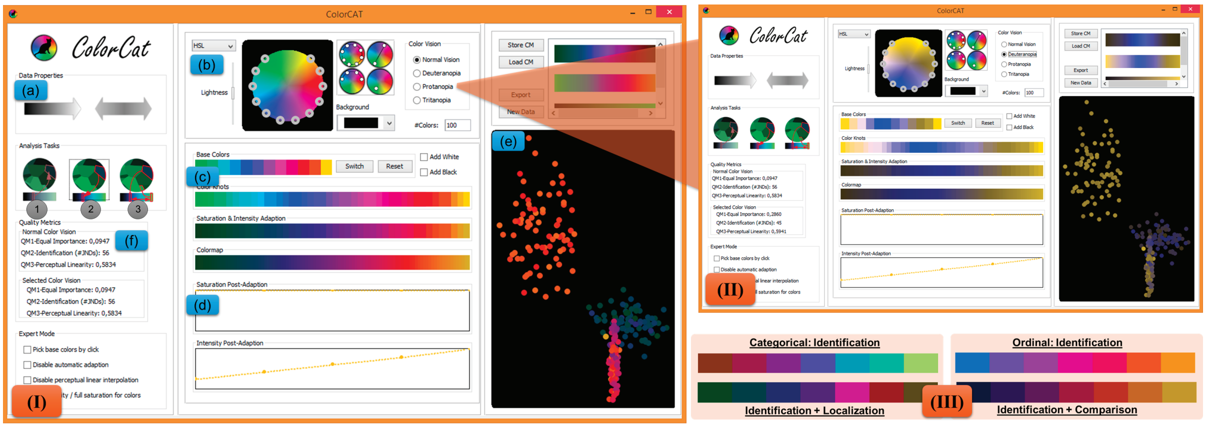

Figure 3: *(Ia) The user selects the data type and analysis task combination (1 Localization, 2 Identification, 3 Comparison) and ColorCAT suggests a colormap design based on the task requirements (see Figure 1). (Ib) The user can select different base colors in the color picker, which allows free adding, removing and rotating knots. (Ic) visualizes the steps of the colormap algorithm for the designer to understand how the base colors are modified to match the analysis task. (Id) Splines allow advanced users to modify intensity and saturation of colormaps. (Ie) The scatterplot allows visual inspection of colormaps and (If) reveals the quality of colormaps. (II) shows ColorCAT simulating red-green blindness. (III)Examples for categorical and ordinal colors.*

color encoding is faithful if the color distances $\Delta E(c_1, c_2)$ of two data values reflect the data distance $d(d_1, d_2)$. Therefore, perceptual linearity can be measured with the *Sammon's stress measure* (Eq. 3) [Sam69, MBS*14]. To provide perceptual linear colormaps for sequential and diverging data, the colormaps must vary from single color hues to black with perceptual linear decreasing saturation and intensity.

$$QM_3 = \frac{1}{\sum_{i<j} d(d_i, d_j)} \sum_{i<j} \frac{(d(d_i, d_j) - \Delta E(c_i, c_j))^2}{d(d_i, d_j)} \qquad (3)$$

R4-Comparison (relative differences). It is possible to preserve a perceptual ordering in a non-linear colormap. Colormaps that vary over multiple hues and linear over intensity are perceptually ordered and thus, enable relative judgments [War12] (see 23).

12 **R1 & R2.** It is possible to build combinations of the different tasks, e.g., the analysts wants to *locate* specific objects and at the same time to *identify* (*browse*) the values of other objects (this supports the *explore* task of Brehmer and Munzner [BM13]). To support identification, the colormaps must vary over hues with a maximum of saturation (R2). Intensity is increased to highlight the value ranges of interest (R1).

23 **R2 & R3.** The most common analysis task combination is that the analyst wants to *identify* but also to *compare* data values. The challenge is to provide perceptual linearity and many distinct colors simultaneously. The results are the well-known *spiral colormaps* [LH92, Kei00, War12] for sequential data that vary over hues with a maximum of saturation (R2). The increasing intensity perceptually orders the colors and thereby reduces the bias of non-linearity on relative judgments (R4).

13 **R1 & R3.** This combination comprises the *localization* and *comparison* of data values. Complete perceptual linearity (R3) cannot be achieved because some value ranges must stand out due to highlighting. Therefore, colors must be perceptually ordered by linear increasing intensity (R4) and saturation increases to highlight the specified data objects (R1).

123 **R1, R2 & R3.** The combination of all tasks is not recommended because this results in colormaps with many trade-offs. By varying over hues with linear increasing intensity (R4) and a minimum of saturation (R2), values can be *identified* and *compared*. To support *localization*, saturation increases to the specified value range (R1).

3. *ColorCAT*: Guided Design of Colormaps

The idea of *ColorCAT* is that the designer specifies the properties of the data (sequential, diverging) and selects the analysis task combination (see Figure 3). *ColorCAT* then derives the requirements (R1,R2,R3,R4) for the selected tasks and models the intensity and saturation gradient of the colormap to minimize the quality metrics $QM1 - QM3$. The designer selects and orders base colors especially for *identification* tasks to provide multiple hues (R2). The colormap is generated by interpolating between the base colors in a perceptual uniform color space to maximize JNDs (R2) and to ensure perceptual linearity (R3) and/or orderliness (R4). Advanced designers are able to interactively change all properties of the colormap.

Categorical and Ordinal Data. All colormaps in this paper are designed to map continuous data. *ColorCAT* can also generate color encodings for categorical and ordinal data (see Figure 3(III)) since the requirements of Section 2 are valid for these data types as well. The user can specify the number of colors and thereby the number of categories. Categorical data can only be *identified* or *localized* since there exist no absolute or relative differences between categories.

3.1. Interactive Selection of Base Colors.

ColorCAT provides an interactive color picker that visualizes the intuitive HSL color space. *ColorCAT* supports the user by suggesting color orderings for, e.g., *spiral colormaps* and color harmonies for diverging colormaps. In this way, the designer can order the colors according to domain, culture, or user preferences (R4). Expert users can scan through the HSI and CIECAM02 color spaces and create customized colormaps beyond our guidelines.

3.2. Color Vision Deficiencies

In order to design colormaps that are color-blind safe, *ColorCAT* can be switched into three types of color blindness: protanopia, deuteranopia (most common), and tritanopia. All colors in *ColorCAT* will be simulated by the approach of Brettel *et al.* [BVM97] according to the selected deficiency and thus, colormap designers can perceive how the colormap will appear to a color-blind person (see Figure 3(II)). Additionally, *ColorCAT* indicates the quality by $QM1 - QM3$ of the colormaps for the selected color blindness. This mainly influences the base color selection since mixing red and green hues should be avoided. However, also these hues can be mixed if the difference in intensity between these hues is high.

3.3. Color Map Algorithm

ColorCAT uses the perceptual uniform color space CIECAM02-UCS [LCL06] to estimate perceived color differences and thus, create perceptual linear colormaps. However, perceptual uniform color spaces share the problem that interpolation especially along the lightness channel often leads to colors that are undefined in RGB [LSS12]. The HSI color space [Kei00] is not perceptually uniform but is defined (in RGB) for the creation of colormaps with perceptual linear increasing intensity. The trade-off is that all interpolations between base colors are calculated in CIECAM02 (in order to provide perceptual linearity) and HSI is used for modeling the intensity (and saturation) properties of a colormap.

The algorithm performs the following steps: 1) The base colors are extended by interpolation between the base colors in the HSI color space. This increases the number of "color knots" for interpolating intensity and saturation. 2) The algorithm assigns intensity and saturation to the "color knots" according to the analysis tasks. 3) For perceptual linearity, the algorithm computes the distances between all color knots with CIECAM02-UCS and places the knots according to their distances in the final (ordered) set of colors for the colormap. 4) The empty spaces between the knots are interpolated in CIECAM02 to provide a perceptual linear colormap.

3.4. Interactive Refinement & Interfaces

ColorCAT visualizes each step of the algorithm in order to enable the user to understand how the base colors are utilized and modified, which can be interactively changed in the color picker tool. Further, we provide an interactive spline chart that visualizes the intensity and saturation of the colormap. The user can modify the splines by interactively adding, removing, and moving control points of the intensity and saturation splines. Thus, the user can specify the data importance (see Section 2) for *localization* tasks. The quality metrics panel reveals the quality of the colormap in the according metrics. *ColorCAT* enables the user to store alternative colormaps in a list and provides a scatterplot of continuous data, which is encoded with the selected or currently modified colormap for visual inspection. The background color has high impact on the foreground color perception. Colors may blend with the background or strong contrasts change the appearance of single colors. Therefore, the designer can switch the background color of the scatterplot for visual inspection. Our colormaps work best on black backgrounds. We omitted the usage of white, which is often used to encode extreme values but would blend with white backgrounds. *ColorCAT* exports the colormaps in different formats (RGB and CIELAB color pallets, Java and Javascript arrays) in data files, but also can export this directly as JAVA classes. Exported classes can be directly used in JAVA based systems to visualize the colormap but also provide the interactive spline chart to modify the intensity and saturation properties of colormaps.

4. Discussion & Future Work

There exist sophisticated color mapping algorithms that outperform the colormaps of *ColorCAT* in single analysis tasks or data types. However, the advantage of *ColorCAT* is that it integrates the visualization expert in the design of colormaps. The expert intuitively combines different analysis tasks and modifies the colormap to match the target domain, user preferences, and culture, which is not possible for automatic methods. We argue that the integration of the visualization expert is more important for design processes of visualizations since the challenge of visualization design is to match the mental model of the target user. Contrast effects have high impact on color encodings. We therefore suggest applying the method of Mittelstädt *et al.* [MSK14] to avoid this issue or at least to add multiple base colors since varying over hues minimizes contrast effects as well [War88]. Aesthetic design goals are also very important in colormap design, because aesthetic designs reduce the stress of visual analysis tasks [WGM*08] and make the use of tools more enjoyable [Nor02]. It remains an open question how to satisfy perceptually-motivated metrics and to allow enough artistic freedom for colormap design simultaneously. Novel colormaps are presented by Samsel *et al.* [SPG*15] who derived the designs with trained artists. It would be interesting to extend *ColorCAT* in this direction.

5. Conclusion

In this paper, we introduce color mapping requirements and quality metrics for elementary analysis tasks and task combinations. Further, we present an approach to generate (also color-blind safe) colormaps for each task combination and provide *ColorCAT*, which comprises the automatic requirement analysis for task-based color design and interactively guides designers through the process of designing colormaps.

References

[AA06] Andrienko N., Andrienko G.: *Exploratory Analysis of Spatial and Temporal Data*. Springer, 2006. 1

[BM13] Brehmer M., Munzner T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2376–2385. 1, 2, 3

[Bre15] Brewer C. A.: Colorbrewer, 2015. (visited on 22/01/2015). URL: http://colorbrewer2.org/. 1

[BRT95] Bergman L. D., Rogowitz B., Treinish L. A.: A rule-based tool for assisting colormap selection. In *Proceedings of the IEEE Conference on Visualization* (1995), IEEE, pp. 118–125. 1

[BSM*15] Bernard J., Steiger M., Mittelstädt S., Thum S., Keim D. A., Kohlhammer J.: A survey and task-based quality assessment of static 2D color maps. In *Proceedings of SPIE 9397, Visualization and Data Analysis* (2015). 1, 2

[BVM97] Brettel H., Viénot F., Mollon J. D.: Computerized simulation of color appearance for dichromats. *JOSA A 14*, 10 (1997), 2647–2655. 4

[CYG04] Camgöz N., Yener C., Güvenc D.: Effects of hue, saturation, and brightness: Part 2: Attention. *Color Research & Application 29*, 1 (2004), 20–28. 2

[GGMZ05] Guo D., Gahegan M., MacEachren A. M., Zhou B.: Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach. *Cartography and Geographic Information Science 32*, 2 (2005), 113–132. 2

[HB03] Harrower M., Brewer C.: Colorbrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal 40*, 1 (2003), 27–37. 1, 2

[Kei00] Keim D.: Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Transactions on Visualization and Computer Graphics 6*, 1 (2000), 59–78. 1, 2, 3, 4

[KOF08] Kuhn G. R., Oliveira M. M., Fernandes L. A.: An efficient naturalness-preserving image-recoloring method for dichromats. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1747–1754. 1

[KRC02] Kindlmann G., Reinhard E., Creem S.: Face-based luminance matching for perceptual colormap generation. In *Proceedings of the conference on Visualization* (2002), IEEE Computer Society, pp. 299–306. 1

[LCL06] Luo M. R., Cui G., Li C.: Uniform colour spaces based on CIECAM02 colour appearance model. *Color Research & Application 31*, 4 (2006), 320–330. 2, 4

[LH92] Levkowitz H., Herman G. T.: The design and evaluation of color scales for image data. *IEEE Computer Graphics and Applications 12*, 1 (1992), 72–80. 1, 3

[LSS12] Lee S., Sips M., Seidel H.-P.: Perceptually-driven visibility optimization for categorical data visualization. *IEEE Transactions on Visualization and Computer Graphics* (2012), 1746–1757. 1, 4

[Mac42] MacAdam D. L.: Visual sensitivities to color differences in daylight. *JOSA 32*, 5 (1942), 247–273. 2

[MBS*14] Mittelstädt S., Bernard J., Schreck T., Steiger M., Kohlhammer J., Keim D. A.: Revisiting Perceptually Optimized Color Mapping for High-Dimensional Data Analysis. *In Proceedings of the Eurographics Conference on Visualization (EuroVis 2014, Short Paper)* (2014), 91–95. 1, 2, 3

[MEO94] Mahy M., Eycken L., Oosterlinck A.: Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Color Research & Application 19*, 2 (1994), 105–121. 2

[MSK14] Mittelstädt S., Stoffel A., Keim D. A.: Methods for Compensating Contrast Effects in Information Visualization. *Computer Graphics Forum 33*, 3 (2014), 231–240. 2, 4

[Nor02] Norman D.: Emotion & design: attractive things work better. *Interactions 9*, 4 (2002), 36–42. 4

[Oli13] Oliveira M. M.: Towards more accessible visualizations for color-vision-deficient individuals. *Computing in Science & Engineering 15*, 5 (2013), 80–87. 1

[Rhe00] Rheingans P.: Task-based color scale design. In *28th AIPR Workshop: 3D Visualization for Data Exploration and Decision Making* (2000), International Society for Optics and Photonics, pp. 35–43. 1

[RTB96] Rogowitz B., Treinish L., Bryson S.: How not to lie with visualization. *Computers in Physics 10*, 3 (1996), 268–273. 2

[Sam69] Sammon J. W.: A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers 18*, 5 (1969), 401–409. 3

[SMO*13] Sajadi B., Majumder A., Oliveira M. M., Schneider R. G., Raskar R.: Using patterns to encode color information for dichromats. *IEEE Transactions on Visualization and Computer Graphics 19*, 1 (2013), 118–129. 1

[SPG*15] Samsel F., Petersen M., Geld T., Abram G., Wendelberger J., Ahrens J.: Colormaps that improve perception of high-resolution ocean data. In *Proceedings of ACM CHI 2015 Conference on Human Factors in Computing Systems (in press)* (2015). 4

[TFS08] Tominski C., Fuchs G., Schumann H.: Task-driven color coding. In *Information Visualisation* (2008), IEEE, pp. 373–380. 1, 2

[War88] Ware C.: Color sequences for univariate maps: Theory, experiments and principles. *IEEE Computer Graphics and Applications 8*, 5 (1988), 41–49. 1, 2, 4

[War12] Ware C.: *Information visualization: perception for design*. Elsevier, 2012. 3

[WGM*08] Wang L., Giesen J., McDonnell K. T., Zolliker P., Mueller K.: Color design for illustrative visualization. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1739–1754. 1, 4

[WVVW*08] Wijffelaars M., Vliegen R., Van Wijk J. J., Van Der Linden E.-J., : Generating color palettes using intuitive parameters. *Computer Graphics Forum 27* (2008), 743–750. 1