

Methods for Compensating Contrast Effects in Information Visualization

S. Mittelstädt, A. Stoffel and D. A. Keim

University of Konstanz, Germany

Abstract

Color, as one of the most effective visual variables, is used in many techniques to encode and group data points according to different features. Relations between features and groups appear as visual patterns in the visualization. However, optical illusions may bias the perception at the first level of the analysis process. For instance, in pixel-based visualizations contrast effects make pixels appear brighter if surrounded by a darker area, which distorts the encoded metric quantity of the data points. Even if we are aware of these perceptual issues, our visual cognition system is not able to compensate these effects accurately. To overcome this limitation, we present a color optimization algorithm based on perceptual metrics and color perception models to reduce physiological contrast or color effects. We evaluate our technique with a user study and find that the technique doubles the accuracy of users comparing and estimating color encoded data values. Since the presented technique can be used in any application without adaption to the visualization itself, we are able to demonstrate its effectiveness on data visualizations in different domains.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms

1. Introduction

Color, after position, is among the most effective visual variable to encode information [Tre85]. Color is pre-attentively processed and thus, we perceive relations between features and groups of data points. Patterns are unconsciously detected and correlated by our perception. Often perception dominates our cognition and therefore, color mapping must obey perceptual processes in order to provide faithful data visualizations.

Optical illusions may bias the analysis process at perceptual and cognitive levels. In our cognition, illusions are caused by assumptions about the relation of visual objects which lead to unconscious inference. Perceptual illusions are caused in the processing of physical stimuli. Incoming light is first processed by nerve cells that measure the intensity of light of different wavelengths (short, mid and long wavelengths). This information is passed to cells that combine this information to perceive luminance and color. In order to detect and identify natural shapes, objects, and movement our visual perception has evolved to be good at detecting edges and determining different color hues. Amplifying contrasts is a means of achieving this.

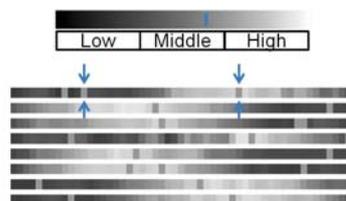


Figure 1: Pixel-cell based visualization for time series data. The highlighted values share the same data value of 0.6. Due to the surrounding regions in the visualization, contrast effects let us perceive the data points differently.

In data visualizations, these contrast effects can be harmful. Ware [War88] has quantified the bias of contrast effects, concluding that this can cause errors of up to 20 %, leading to analysts grouping wrong data points and detecting relations or extremes that do not exist in the data. Figures 1 and 2 demonstrate the extent of these contrast effect. As with most optical illusions, we cannot compensate for these effects even if we are aware that the representations are wrong.

There are sophisticated defined guidelines and discussions of colormap usage in the field. Colormaps have been devel-

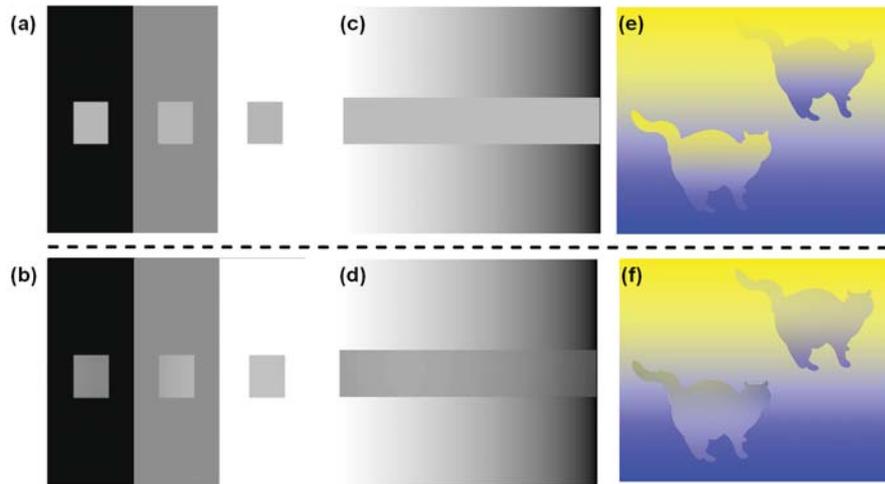


Figure 2: Standard examples of simultaneous contrast. (Top) Original images. The gray patches share the same gray values but are perceived differently (a) or as a gradient (c). The cats (e) share the same gradient from less saturated blue to yellow. (Bottom) Compensated images. The patches and cats are almost perceived equal. (b) Our method reduces the contrast effect on the patches to faithfully represent the gray value (3). (d) The method reduces the perceptually induced gradient (6). (f) Our method improves the global gradient and average color of the cats (7).

oped to avoid contrast effects and there is still a discussion whether such scales are intuitive and support enough distinct color levels to express continuous data. But of prime importance is that color perception is relative, so that the perceived color of an area of a display can only be predicted, if the surrounding colors are known. Thus, these guidelines cannot guarantee unbiased displays.

Our contribution. In this paper, we present a postprocessing technique for compensating contrast effects in visualizations as illustrated by the standard examples in Figure 2. We claim the following contributions: 1) A method for compensating physiological color effects based on color appearance models and optimization algorithms that can be used on any data visualization as a postprocessing step; 2) A definition of the optimization goal and the corresponding perceptual metrics; 3) A general heuristic to approximate the gradient of compensation; 4) An evaluation of the perception model and the compensation, based on realistic tasks and data.

While computational models for physiological effects exist, cognitive effects are far more complex and concrete theories or models have yet to be formulated. We therefore exclude cognitive effects from the scope of this paper.

2. Related Work

The influence of contrast effects on data visualizations has been quantified and evaluated in [CM83, War88, Bre97], who indicate that their presence significantly influences users when they look at visualizations. There already exist strategies to handle these effects. Appropriate colormaps for specific tasks and specific data properties is a well discussed

topic in the literature and general guidelines on selecting color maps can be found in [RO86, War88, BRT95, RTB96]. In addition, colormaps for segmentation and categorical data have been discussed previously [Hea96, HB03]. Ware found that the metric task (reading metric quantities) is significantly influenced by contrast effects, when using a color scale along one chromatic/achromatic channel. Some solutions [War88, LH92, Kei00, KRC02] therefore propose *spiral* color maps, which maximize color differences by varying over hues with linear increasing intensity, in order to create color maps that alternate between our chromatic channels, thereby reducing the probability of contrast effects. A method for predicting simultaneous contrast on maps is presented by Brewer [Bre96] — the models and guidelines create a set of colors that to a high extent do not produce contrast effects in the final visualization. Recoloring tools [KOF08, FRGG13] have been developed that modify website colors to make them accessible for people with color vision deficiency while preserving the subjective response and color differentiability.

The perception of colors on the display is not taken into account in all the above techniques. Without knowing the surrounding color of a data point it is impossible to accurately estimate its perception a priori. We therefore propose a postprocessing method to cope with these effects.

How color is perceived and how to model illusions is still a focus of research. Hunt et al. [HP11] and Fairchild [Fai13] offer a full discussion of color appearance. The standardized color appearance model CIECAM02 [MFH*02] is based on the results of the Hunt model but is adapted for industrial practical usage. These models are still based on certain viewing conditions and fixed patch sizes. Further advances for

more complex stimuli of natural scenes and varying viewing conditions have been made in the iCAM framework by Fairchild et al. [FJ04].

Many computational models for brightness effects have been developed over the last decades. For instance, the anchoring theory of Gilchrist et al. [GKB*99] and Scission theory of Anderson et al. [AW05] try to explain these effects on high-level (cognitive) interpretation of a scene by segmenting and processing the complex scene into frameworks and layers. In contrast, low-level models such as the ODOG model of Blakeslee et al. [BPM05] successfully use simple Gaussian based convolutions to predict a variety of contrast effects such as simultaneous contrast, Mach bands, Hermann grid and White's illusion. A full introduction and discussion of achromatic vision can be found Gilchrist's book [Gil06].

3. Method for Compensating Contrast Effects

Our method to compensate physiological biases in data visualizations is an optimization process. The goal is to find an image I' that is perceived as the original image I and therefore faithfully represents the data. Our method finds I' in

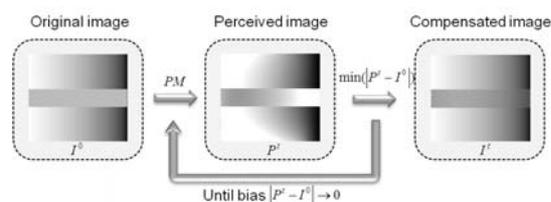


Figure 3: Schematic approach. The method iteratively reduces the difference between perceived and original image.

an iterative process illustrated in Figure 3: First, the method estimates in step t the bias for a given input image I^t at each pixel I_p^t . Then these effects are compensated by changing the color of pixel I_p^t and its surround S_p^t to reduce the bias. These steps are iteratively performed until the bias for all pixels is compensated. The bias at each pixel is estimated by a perception model PM that predicts the perceived image P^t of I^t . If the perceived image $PM(I^t, S^t) \mapsto P^t$ is equal to image I then the data is faithfully represented. The goal of compensation can thereby be defined by minimizing the difference of the perceived image and the original image (1).

$$\min(|PM(I, S) - I|) \quad (1)$$

A sound perception model PM is therefore the basis of the compensation method. Since vision research still advances to optimize or create new models to cover all physiological effects and some visualization techniques may suffer under different effects, we propose a method, where the perception model is an updateable module.

If the model PM is continuous, then the effects on one pixel I_p can be inverted and color I_p' , which is correctly perceived, can be estimated. Another solution is to change the

surrounding pixels S_p in order to faithfully represent I_p . Both work when considering a solitary pixel, but in a multi-pixel display the adjustment of one pixel will change the perception of other pixels. Therefore, the compensation method must find an optimum to compensate the effects for the whole visualization. Our process therefore requires perceptual metrics as cost functions to evaluate a solution I' , a sound perception model PM , and an optimization algorithm to find solutions that meet the compensation goal (1).

3.1. Cost functions

In the following sections, we define our cost functions based on perceptual metrics and different visualization tasks as defined by Tominski et al. [TFS08] that are based on a task model of Andrienko and Andrienko [AA06]. The visualization tasks are grouped in two levels: elementary and synoptic tasks. While elementary tasks address individual data points, synoptic tasks consider sets of values or data points.

Color distances are the base of all of our perceptual metrics. We use the DIN99 color space for distance measurements. This color space is an extension to CIELAB that accurately models small color distances and allows vector arithmetics. Its distance metric accords to CIEDE2000.

$$\Delta E(c_1, c_2) = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2} \quad (2)$$

3.1.1. Cost Functions for the Elementary Tasks

Any change in the perception of the original color of data points misleads users in the localization, identification and comparison tasks. Therefore, the costs have to increase if a data point is perceived differently as the original color. While the foreground F that holds the data must be accurately visualized, the background B can be used for additional space to reduce the effects on the foreground. We therefore differentiate between foreground and background.

The first measure (3) computes the difference between the original color and the perceived color of each pixel in the foreground. This can also be measured in similar fashion for the background (4) in order to avoid disturbing halos or artifacts. In order to avoid further misinterpretation of the background, the cost function (5) measures the distance of a background color to the closest color $CM(P_p)$ of the color map. If the distance is very close ($\Delta E < 5$), users may interpret the background as a data point, which harms identification and localization of data values.

$$f_1(I, P) := \frac{1}{|F|} \sum_{p \in F} \Delta E(P_p, I_p) \quad (3)$$

$$f_2(I, P) := \frac{1}{|B|} \sum_{p \in B} \Delta E(P_p, I_p) \quad (4)$$

$$f_3(I, P) := \frac{1}{|B|} \sum_{p \in B} \max(0, 5 - \Delta E(P_p, CM(P_p))) \quad (5)$$

3.1.2. Cost Functions for the Synoptic Tasks

The faithful representation of single data values is ensured by the measures above. However, these measures do not preserve local and global relationships. The local measure (6) preserves the distance between each pixel and its neighbors and thus spatially connected structures. It effectively removes biased perceptual gradients as shown in Figure 2d). An issue with preserving global structures or color patterns is that they may not be spatially connected and not known a priori. This makes these patterns hard to preserve. The method can, however, ensure that the global relations and the average impression between regions of the image are faithfully represented such as the cats in Figure 2f). The constraint of faithfully representing each data point is therefore relaxed in order to preserve the global relationships of data items. This can be obtained by computing a Gaussian pyramid of the original and perceived images with $\delta = 2^\circ$ as first level according to the attentional center of the CIE standard observer (7). Each pixel is compared to its according position in the pyramid levels. We decided to use a Gaussian pyramid, because it measures the perception of the images at different resolutions and thereby accounts for structures of different granularity.

$$f_4(I, P) := \frac{1}{|F|} \sum_{p \in F} \sqrt{\frac{1}{|N_p|} \sum_{n \in N_p} (\Delta E(P_p, P_n) - \Delta E(I_p, I_n))^2} \quad (6)$$

$$f_5(I, P) := \frac{1}{|G|} \sum_{G_I \in G} \frac{1}{|F|} \sum_{p \in F} \Delta E(G_I(P)_p, G_I(I)_p) \quad (7)$$

3.1.3. Combinations of Cost Functions

Since all of the cost functions are based on the color distance measure, the cost functions do not have to be scaled and can be aggregated in a weighted sum.

$$f(I, I', PM) := \sum_{i=1}^{n=5} \lambda_i \cdot f_i(I, PM(I')) \quad (8)$$

The weights of the cost functions depend on the visualization task, data and colormap. A recommendation and evaluation of weights is still ongoing work. In our experience, equalizing the weights of all cost functions beside f_2 gives satisfactory results. We assign f_2 a very low weight since this allows the background to compensate effects in the foreground with regards to f_3 .

3.2. Perception Model

The iCAM framework has been developed as an enhancement of the standardized CIECAM02 color appearance model for image processing. It is robust in predicting color appearance and contrast effects, such as simultaneous contrast and chroma chrispening.

3.2.1. Revision of the iCAM Framework

The iCAM framework processes the image in the following way: First, the image is separated into three intensity images according to the response signals (short, medium, and long wavelengths) of the cones in the human eye. Second, each intensity image I is convolved with a Gaussian kernel to model the perceived surround S_p of each pixel. Then, the model adapts each color at pixel p of image I to its surround S_p and to the reference white $D65$ of sRGB. This local chromatic adaptation models a variety of contrast effects. The reference white can be used to model ambient light or other lightening conditions. The parameter ε controls the degree of adaption. The higher ε the more contrast is reduced. More details can be found elsewhere [FJ04, FJ05]. The constants are selected with $c_1 = 0.94$, $c_2 = 0.06$, and $c_3 =$ the maximum of the LMS channel [94.92, 103.54, 108, 74] in the CAT02 color space [MFH*02].

$$PM(I, S, \varepsilon) = \left(c_1 \cdot \frac{D65}{c_3 \cdot (S_p/I_p)^\varepsilon} + c_2 \right) \cdot I_p \quad (9)$$

3.2.2. Our Extensions to the iCAM Model

The size of the kernel that models the surround appearance is hereby a critical parameter. The kernel size is estimated in iCAM by 20% of the image width. According to [SCT*13] and our study, we found that this is insufficient since simultaneous contrasts are increased in images with high spatial frequency. We therefore adjust the kernel to the spatial frequency of the image. We filter the image by difference-of-Gaussians of varying sizes ($s_{i=0}^n = 2^i$ and with a fixed size ratio of 1 : 1.6), which accords human perception of edges in natural scenes [You87]. The root-mean-square response of each filter result indicates the presence of the according frequencies. The filter sizes are then pooled according to their responses, which approximates the average kernel size that will consider most of the spatial frequencies of the image.

The second critical parameter is the exponent that steers the adaption. There is a difference between bright centers with dark surrounds and dark centers with bright surrounds, which is reported by Blakeslee et al. [BM99]. We differentiate between these cases (10). In our experiments, we found that $\varepsilon_0 = 0.5$ (adapted from [FJ05]) and $\varepsilon_1 = 0.6$ provide satisfactory results.

$$PM'(I, S) := PM(I, S, \varepsilon), \text{ with } \begin{cases} \varepsilon = \varepsilon_0, & \text{if } S_p > I_p \\ \varepsilon = \varepsilon_1, & \text{else.} \end{cases} \quad (10)$$

3.3. Optimization Algorithms & Heuristics

The optimization goal (1) can be achieved by minimizing the sum of cost functions (8). The problem space is non-linear and may be non-convex with non-linear constraints depending on the perception model. The model itself may be continuous, differentiable and homogeneous. These properties significantly influence the selection of an efficient optimization algorithm. Since we do have a meaningful initialization

for the optimization and a rough approximation of the gradient of compensation, heuristics have great potential.

Heuristics. One of the most effective group of methods to optimize a function are gradient methods. By inverting the perception model the gradient of compensation can be estimated. Some models such as our model are not differentiable and thus, the gradient can only be approximated. This can be revealed by determining the direction (vector in the DIN99 color space) of contrast effects. For example, a bright patch of gray is perceived even brighter if the surrounding is dark. Thus, the perception model predicts an increase of intensity for the patch and a decrease of intensity for the surrounding. In order to compensate this effect, either the patch itself has to be made darker to reduce the effects on itself or the surrounding made brighter to compensate the effects on the patch.

Our iterative heuristic approximates I' in step t with I^t . It starts in step $t = 0$ with $I^0 = I$. The effects on one pixel ΔI_p^t can be estimated by calculating the difference of the perceived image $PM(I^t, S^t)$ to the original image I (11). If the model PM is roughly continuous and homogeneous, the effects will influence a pixel I_p^{t+1} in the same way. Therefore, the difference between perceived and original pixels can be reduced if, I_p^{t+1} is selected such that $I_p^{t+1} + \Delta I_p^t = I$. In the example, the bright patch will become darker (inverse direction of ΔI^t) in order to compensate the effects on itself.

$$\Delta I^t = PM(I^t, S^t) - I \quad (11)$$

A pixel I_p^t can also change its color to reduce the effects on other pixels. The pixel measures the effects $\Delta S(I_p^t)$ (12) on its surround S_p given by PM . The effects on surround pixels are summed according to the weights of the influence function $s(\Delta x, \Delta y)$ of the perception model (Gaussian kernel in iCAM). If pixel I_p^t adapts in the direction of $\Delta S(I_p^t)$, it will reduce the contrast effects for its surroundings. In the example, the dark surround will become brighter (direction of $\Delta S(I_p^t)$) in order to compensate the effects on the bright patch.

$$\Delta S(I_p^t) = \sum_{n \in S_p} \Delta I_n^t \cdot s(\Delta x_{p,n}, \Delta y_{p,n}) \quad (12)$$

The two different methods of compensation can be combined with Eq. (13). As describe above, inversion of the perception model is not enough, since changing one pixel will influence its surrounding. We calculate the compensated image I' with multiple iterations and control the step size of compensation with $\varphi_{1,2} \in [0, 1]$. The pixels of I^{t+1} will adapt to the change of their surround in I^t . In practice, we set $\varphi_1 > 0, \varphi_2 = 0$ for foreground pixels and $\varphi_1 = 0, \varphi_2 > 0$ for background pixels. In this way, the background, which does not contain valuable data, compensates the effects on the foreground.

$$I_p^{t+1} = I_p^t - \varphi_1 \cdot \Delta I_p^t + \varphi_2 \cdot \Delta S(I_p^t) \quad (13)$$

It should be noted that this computation of the gradient is independent of the perception model. The required influence function $s(\Delta x, \Delta y)$ can be approximated by Gaussians or difference-of-Gaussians for most computational models.

3.4. Instantiation of the Method

For our experiments and applications we use the following instantiation and parameters: We use simulated annealing with multiple threads. Each thread is initialized with the original image as starting point for the optimization. In each iteration the neighbor solution is determined by Eq. (13).

Parameterization. The perception model is parameterized as described in Section 3.2.2 with $\epsilon_0 = 0.5, \epsilon_1 = 0.6$. The step sizes of the iterations are set to $\varphi_1 = \alpha, \varphi_2 = 0$ for foreground pixels and $\varphi_1 = 0, \varphi_2 = \alpha$ for background pixels. We widen the search space by randomizing $\alpha \in [0, 1]$ in each thread. Our tasks are focused on the elementary metric readings and synaptic identification of local neighborhood trends. Therefore, we prioritize cost functions f_1-f_4 by equalizing their weights in the iterative steps and exclude cost function f_5 that preserves global relationships. For the final decision, we select the optimum solution that minimizes all cost functions f_1-f_5 .

Bounding constraints. Another issue is the limits of the display. sRGB does not support all the colors that are defined in perceptual color spaces such as DIN99. One solution is to integrate the borders of the defined sRGB colors as non-linear constraints in the optimization as presented in [LSS12]. We select a heuristic for our iterative method, which samples the defined sRGB color space in DIN99. Pixels that become undefined in one iteration are assigned to the perceptually closest color. In the next iterations, the pixel and its surrounding pixels will adapt to this constraint.

4. Evaluation

4.1. Experiment 1

The goal of the first experiment was to evaluate the perception model on the basis of pixel-cell based visualizations. We measured the accuracy of participants decoding colored information and compared these results to the predicted values of the perception model. Further, we measured the error between participants and the original data and estimated whether the accuracy of participants will improve, if compensation is applied. Our hypotheses were:

- H1** *Predicted values of the perception model are equal to participant results.* In the ideal case, the perception model decodes data values similar to a participant.
- H2** *Participants are as accurate or better with our method than with standard mapping.* Our method reduces the contrast effects and therefore improves the participant results.

Task. One-dimensional time series were visualized in a cell based visualization above and under an interactive drawing field (see Figure 4). The participants were asked to estimate the quantity of each pixel cell and the trend of the time series. They were able to redraw the trend by clicking on pixels, dragging the line (trend) and correcting any errors by simply adjusting the line accordingly. The colormap was shown above the field for reference and participants were told that the lowest and highest values in the colormap correspond to the lowest and highest positions in the field.

Such complex tasks include biases. We see a gap between perceiving the data values on the screen and expressing the cognitive processed information to the drawing field. Also the participant may be influenced by the line drawn before. However, participants have to cognitively process sets of data values and recognize trends, which is realistic compared to real data analysis scenarios.

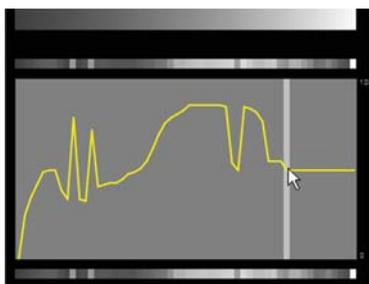


Figure 4: *Task of experiment 1. Participants had to draw a line (shown in yellow) that estimates the value at each point in the time series shown above and below the drawing field. The colormap was given for reference.*

Experiment Factors. The experimental factors were colormap and color mapping (with or without our method). We selected one achromatic scale and three chromatic scales shown in Figure 5 on the basis of the guidelines in [War88, BRT95, Kei00]. All colormaps were perceptually equally spaced by interpolations in the DIN99 color space. The achromatic scale continuously increased in lightness from dark gray to white. The first linear chromatic scale decreased in saturation from blue to white. The second linear scheme had equal lightness and only varied over hues from red to blue. The fourth *multihue* colormap varied over hues with continuously increasing intensity. Note, that some *multihue* sequences such as the common *rainbow* sequence share the problems of non-perceptual uniformity, false coloring and attention steering that are harmful in real analysis scenarios. These issues are reduced in the selected colormap of [Kei00] (see Section 2).

4.1.1. Experimental design

A user study with 40 participants was conducted. The experiment was split into two parts: In the first part, the data was collected to evaluate the perception model. In the second part, the mappings (with and without our compensation method) were alternated to measure how the results of participants change when compensation is applied.

First part. 20 participants performed the task on different colormaps. The participants were randomly assigned to one colormap and trained on the tool and colormap. Each participant fulfilled two phases with a break in-between. In each phase, the participants performed two training tasks to ensure that they understood the colormap correctly followed by three real tasks.

Second part. The goal was to measure the improvement of accuracy of participants with our compensation method. Therefore, the experiment followed a within-subject design and another 20 participants were randomly assigned to the colormaps. They were trained on the tool and then performed 2 phases with each 4 tasks alternating between the mappings in randomized order.

Participants. The participants (14 female, 26 male) were mixed graduated and non-graduated without background of information visualization but they were all familiar with infographics. The ages ranged from 19 to 57 with an average of 27. 37 participants had normal or corrected to normal color vision. Three male participants had deuteranopia.

Data. The data originated from power consumption measurements of a Smart Grid environment, where the analysis of past data requires a faithful visualization of large volumes of time series data. A set of 40 similar time series was selected from this data source for our experiment. In each task, a time series was randomly selected from this set and visualized. The time series were normalized.

Apparatus. The study was performed under controlled lab conditions. The monitor was color calibrated with a resolution of 1920×1200 pixels. Each pixel cell had a width of 0.125° of the viewing angle. The ambient light was controlled to normal daylight conditions, without reflectance on the screen. The viewing distance was approximately 60 cm. The introduction, training and tasks were web-based and standardized.

Metrics. We measured the error of participants in the task (estimating the quantity of each cell). We therefore computed the Euclidean distance between the original time series and the participant results. In order to evaluate the perception model, the result values of the perception model were predicted for each time series-colormap combination. The difference of participant results, as well as the difference of the perception model and participants were estimated by calculating the pairwise Euclidean distance. The time for a participant to finish each task was recorded.

4.1.2. Results

Figure 5 shows a summary of the results.

Evaluation of perception model. We found no significant difference (U-test: $p > 0.5$) between the pairwise participant-participant distances (median=1.51, iqr=0.84) and the pairwise model-participant distances (median=1.44, iqr=0.30).

Evaluation of compensation method. In our experiment, there was no significant difference (H-Test: $p > 0.6$) between the linear (intensity, saturation and hue) colormap-groups. The error of the linear-group (combined: median=0.98, iqr=0.34) was significantly greater (U-Test: $p < 0.01$) than the error of the multihue-group (median=0.80, iqr=0.24). The results of the within-subject comparisons (see Figure 5b)) revealed that there was no significant shift of error location and ratio of error scales between our method and a standard

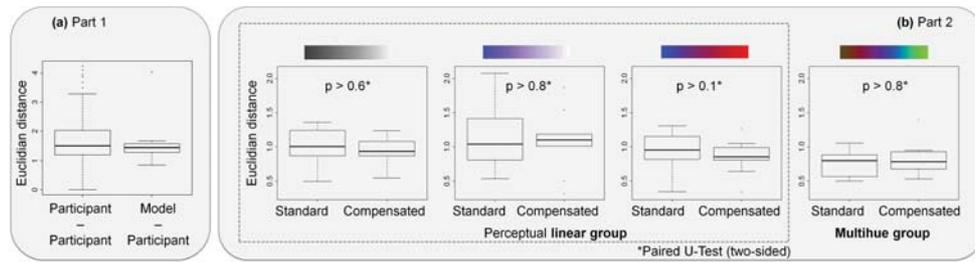


Figure 5: Results of experiment 1. (a) Pairwise distance measure for participant-participant and participant-model comparison. **(b)** Error of participants in the metric task using standard mapping and our compensation method.

mapping. However, the inter-quartile ranges and medians were smaller when compensation was applied for all groups.

Efficiency. The choice of colormap had significant influence on the efficiency of participants performing the task (H-Test: $p < 0.001$). The red-blue colormap group performed worse with a median of 209 seconds (iqr=84.75). The participants mentioned that they had problems to interpret differences in the middle of the colormap. The multihue group (median=94, iqr=33) was slower than the other linear groups (combined: median=81, iqr=56), however, there was no significant difference (U-Test: $p > 0.2$).

4.2. Experiment 2

The goal of our 2nd experiment was to measure the contrast effects on the estimation and comparison of metric quantities, and how participants improve when our method is applied. Our hypotheses were as follows:

H3 *Participants assign more data points to the correct data ranges with our method than with standard mapping.* Using the method, we expected participants not to overestimate and underestimate values in high frequent areas of the image.

H4 *Participants increase the number of correct comparisons of data points with our method than with standard mapping.* We assumed that participants perceive the differences of data points correctly with our method.

Tasks. The task was inspired by Ware’s metric task [War88] for colored data visualizations. Ware showed the participants a visualization and asked participants to match the color at a given pixel (indicated with a cross-hair) with a given set of colors. We adapted this method to our experiment but extended it by direct comparison in the intentional presence of contrast effects. According to Ware [War88], participants have up to a 20 % error in the estimation of the correct quantity if a linear colormap is used. Since we were measuring cases of contrast effects, fine granular data ranges would not increase the expressiveness of our measurements. Therefore, the data range was split into three equally sized parts (low, middle, high). Two data points were marked that share the same data value but had different surrounds (see Figure 1). We ensured that the data values were at least 5 % away from

the closest different data range. Users were asked to assign the points to the data ranges. If a participant assigned the points correctly to same data range, we asked which of the data points was higher or whether they were equal.

4.2.1. Experimental design & metrics

The experiment was a within-subject design with 20 participants of the first experiment (part two). After finishing experiment 1, each participant was shown 32 different visualizations, as illustrated in Figure 1 for the time series data (encoded with their colormap of experiment 1). Our method was applied in half of the visualizations with randomized order of the mappings. For each of the 32 tasks we counted the number of correct data value assignments and comparisons.

4.2.2. Results

The summary of results is illustrated in Figure 6.

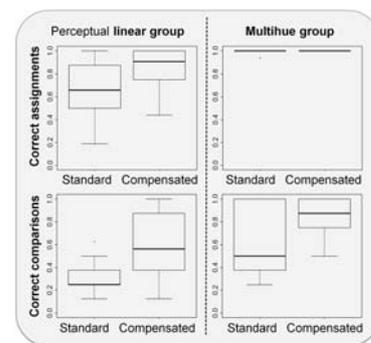


Figure 6: Results of experiment 2. (Top) The percentage of correct data value assignments increases if our method is applied on perceptual linear colormaps. **(Bottom)** The percentage of correct data value comparisons does increase with our method.

Estimation of quantity. As in experiment 1, there was no difference between the linear colormap groups (H-Test: $p > 0.3$). Again, the linear-groups (combined: median=0.81, iqr=0.44) assigned significantly less (U-test: $p < 0.001$) data points to their correct value range than the multihue group (median=1.0, iqr=0). The within-subject comparisons revealed a significant increase in the number of correct participants answers (paired U-Test: $p < 0.001$) with our method

(median=0.91, iqr=0.25) compared to standard mapping (median=0.66, iqr=0.38) for the linear groups.

Direct comparison. Users of all groups showed a significant increase (paired U-Test: $p < 0.05$) of correct answers with our method. In the linear groups the percentage of correct answers increased from (median=0.25, iqr=0.125) to (median=0.56, iqr=0.5) and in the multihue group from (median=0.5, iqr=0.5) to (median=0.875, iqr=0.22).

4.3. Discussion

We found that the predicted values of the perception model were close to the participant results. In general, the perception model was more accurate than participants. Uncertainties of the complex task and abilities to master the colormap and tool varied from participant to participant, which was expressed in the pairwise differences. Since there was no significant difference between the participant-participant and participant-model distances, we conclude that the model behaves like an average participant and thus, we can approve H1.

We can confirm the finding of Ware [War88] that colormaps that vary over hues with linear increasing intensity perform well for the metric task of reading metric quantities from a colored display. In all groups, the inter-quartile ranges and medians decreased with our method, which indicates that the accuracy of all participants was improved and thus, H2 can be approved. Even in combination with our method, grayscale mapping was not as accurate as the standard multihue colormap, which indicates that there might be a maximum of accuracy that can be achieved by each colormap. This confirms the guidelines of the community that colormaps must match the visualization task and data, even with our compensation method.

The second experiment revealed a significant influence of contrast effects on linear scales as already discussed and quantified by Ware [War88]. In contrast, participants were correct with the multihue colormap, because the granularity of data ranges was too coarse for this colormap. However, most of the participants were wrong in the comparison. Our method did significantly improve the correctness of participants in both tasks. Thus, we can approve H3 and H4.

Another issue is raised by the efficiency of participants. The multihue colormap did outperform the other mappings in terms of accuracy in the metric task and the low probability of contrast effects. However, participants were 14 % faster with the grayscale and saturation based colormaps than with the multihue colormap. This might be due to the issues of intuitiveness, since it is easier to identify trends in perceptual linear colormaps.

5. Applications

5.1. Purple America Map

The Purple America map of Vanderbei [Van12] visualizes the results of the 2012 presidential election in the United States. Figure 7 shows a detail of the western part of this map. The visualization uses color to show the proportion of

votes for Democrats (blue), Republicans (red), and others (green) for each congressional district. The color of a district results from a mixture of these colors corresponding to the percentage of votes. This color choice is solely based on the colors associated with the parties and does not consider perceptual effects.

This usage of blue and red leads to very strong color contrasts, which gives rise to small areas of one color appearing stronger when surrounded by the other color. For instance, the color of the districts Baine/ID (A) and Big Horn/MT (C) appear similar in Figure 7a). Looking into the numbers the distance between (A) and (C) is 5.7 %. However, this is the same distance of (A) to (B) with 4.3 % but they appear different. The contrast between the red surround of (A) causes the district to appear more bluish and thus, similar to (C).

The result of our compensation method on this image is shown in Figure 7b). The compensation changed the color of almost all districts. The magnitudes of differences are shown in Figure 7c). In total, the compensation method reduced the contrasts in the image. Not only the contrasts between districts are reduced but also contrasts between the border of the map and white background. As an effect of the compensation the Purple America map appears even more purple than before. The real differences of the colors of the districts (A), (B), and (C) are visibly better. It is clear that the share of blue increases from (B), to (A) and (C), which is a more faithful representation of the data.

5.2. News Visualization

Another example of compensating contrast effects in a news visualization can be found in Figure 8. The original visualization in Figure 8a) plots a semi-transparent triangle for each news item. The color of the triangle (red, white, blue) is mapped to the sentiment of the item (negative, neutral, positive). The different rows represent different news categories at different days. The technique clusters news visually. When several items of news of the same category are published at the same time, the triangles overlap and a continuous block of news becomes visible.

In Figure 8a) we can see that the differing contrast between the triangles and their background alters the appearance of triangles. The perceived saturation of single triangles depends on their color. For instance, the triangles in the fourth row appear to have different saturations. Some appear as saturated as the groups of triangles in the first or third row, which is actually a perception error. Our compensation method takes the color contrasts into account and generates Figure 8b). In this image, the triangles in the fourth row appear to have the same saturation as in the areas with many triangles (in the first and third row). In the difference image in Figure 8c) one can see that the compensation method has changed the triangles differently to compensate for the contrast effects. Through the black background, we over estimate the brightness and differences of triangles in the original image, which is reduced by our method.

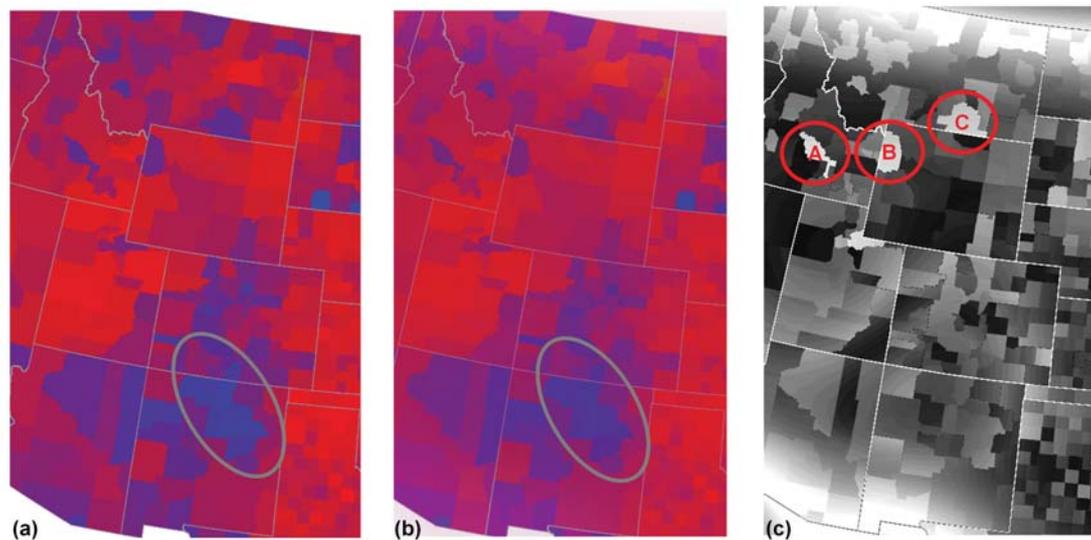


Figure 7: Detail of Vanderbei's Purple America map of the 2012 presidential election in the US [Van12]. The images show the original **a)** and compensated **b)** visualizations. The normalized difference between **a)** and **b)** is shown in **c)**. The labels in **c)** mark the districts or counties Baine/ID (A), Teton/WY (B), and Big Horn/MT (C). The colors in **a)** appear crisp and saturated due to color contrast effects. Our compensation removes color contrasts and the resulting image **b)** appears less saturated with dull colors but represents the data faithfully. In **b)** the share of blue increases from (B) to (A) and (C), which accords to the data.

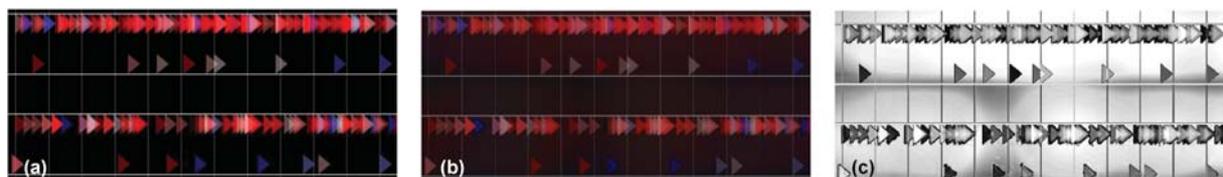


Figure 8: Detail of a news visualization in [WRM*09] using colored triangles with low alpha value for representing single news items in time. The images show the original **a)** and compensated **b)** visualizations. The normalized difference between **a)** and **b)** is shown in **c)**. Depending on the color, single news items in **a)** appear as bright as dense episodes of news. Our compensation in **b)** compensates the contrast effects and the corrects the impression of single news.

6. Limitations & Future Work

We have experienced some limitations of the perception model. Since the method behaves like a contrast low pass filter and uses a distinct kernel size for the whole image, we often have the impression that the final results of the method are not as clear as the original data visualizations. Fine achromatic structures sometimes show artifacts in these cases, because the method induces color. A locally adapted perception model would decrease these issues. Also, the influence of the background is overestimated by the model. Future work will look at methods to adjust the parameters to user, hardware, and visualization and also methods to preserve structural information that is known a priori. One concern with the presented method, however, is its runtime cost. It can be directly applied on any visualization, however, calculation of a perception model is very costly and thus, there is a clear need for more efficient optimizers. In our current

implementation the algorithm needs minutes to converge and therefore, cannot be applied to dynamic or interactive analysis tools. One potential solution is in efficient heuristics that only focus on relevant regions of the visualization and can exclude recalculation of the whole perceived image.

7. Conclusion

In this paper we present a method for compensating physiological color effects based on color appearance models and optimization algorithms. We present the necessary cost functions and heuristics to reach the optimization goal. Our experiments show that with our method users double their accuracy in comparison tasks and significantly improve their accuracy in identifying data values. We demonstrate that our method successfully compensates contrast effects in data visualization in order to provide a faithful display.

References

- [AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory Analysis of Spatial and Temporal Data*. Springer, 2006. 3
- [AW05] ANDERSON B. L., WINAWER J.: Image segmentation and lightness perception. *Nature* 434, 7029 (2005), 79–83. 3
- [BM99] BLAKESLEE B., MCCOURT M. E.: A multiscale spatial filtering account of the white effect, simultaneous brightness contrast and grating induction. *Vision Research* 39, 26 (1999), 4361–4377. 4
- [BPM05] BLAKESLEE B., PASIEKA W., MCCOURT M. E.: Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including white, howe and simultaneous brightness contrast. *Vision Research* 45, 5 (2005), 607–615. 3
- [Bre96] BREWER C. A.: Prediction of simultaneous contrast between map colors with hunt's model of color appearance. *Color Research and Application* 21, 3 (1996), 221–235. 2
- [Bre97] BREWER C. A.: Evaluation of a model for predicting simultaneous contrast on color maps. *The Professional Geographer* 49, 3 (1997), 280–294. 2
- [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. 2, 6
- [CM83] CLEVELAND W. S., MCGILL R.: A color-caused optical illusion on a statistical graph. *The American Statistician* 37, 2 (1983), 101–105. 2
- [Fai13] FAIRCHILD M. D.: *Color Appearance Models*. John Wiley & Sons, 2013. 2
- [FJ04] FAIRCHILD M. D., JOHNSON G. M.: iCAM framework for image appearance, differences, and quality. *Journal of Electronic Imaging* 13, 1 (2004), 126–138. 3, 4
- [FJ05] FAIRCHILD M. D., JOHNSON G. M.: Meet iCAM, 2005. (visited on 12/04/2013). URL: <http://www.cis.rit.edu/research/mcsl2/icam/>. 4
- [FRGG13] FLATLA D. R., REINECKE K., GUTWIN C., GAJOS K. Z.: SPRWeb: Preserving subjective responses to website colour schemes through automatic recolouring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), ACM, pp. 2069–2078. 2
- [Gil06] GILCHRIST A.: *Seeing Black and White*. Oxford University Press, 2006. 3
- [GKB*99] GILCHRIST A., KOSSYFIDIS C., BONATO F., AGOSTINI T., CATALIOTTI J., LI X., SPEHAR B., ANNAN V., ECONOMOU E.: An anchoring theory of lightness perception. *Psychological Review* 106, 4 (1999), 795–809. 3
- [HB03] HARROWER M., BREWER C.: Colorbrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal* 40, 1 (2003), 27–37. 2
- [Hea96] HEALEY C.: Choosing effective colours for data visualization. In *Proceedings of Visualization*. (1996), IEEE, pp. 263–270. 2
- [HP11] HUNT R. W. G., POINTER M. R.: *Measuring Colour*. John Wiley & Sons, 2011. 2
- [Kei00] KEIM D.: Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Transactions on Visualization and Computer Graphics* 6, 1 (2000), 59–78. 2, 6
- [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. 2
- [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. 2
- [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. 2
- [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. 5
- [MFH*02] MORONEY N., FAIRCHILD M. D., HUNT R. W., LI C., LUO M. R., NEWMAN T.: The CIECAM02 color appearance model. In *Proceedings of the Color and Imaging Conference* (2002), vol. 2002, Society for Imaging Science and Technology, pp. 23–27. 2, 4
- [RO86] ROBERTSON P. K., O'CALLAGHAN J. F.: The generation of color sequences for univariate and bivariate mapping. *IEEE Computer Graphics and Applications* 6, 2 (1986), 24–32. 2
- [RTB96] ROGOWITZ B., TREINISH L., BRYSON S.: How not to lie with visualization. *Computers in Physics* 10, 3 (1996), 268–273. 2
- [SCT*13] SHI V., CUI J., TRONCOSO X. G., MACKNIK S. L., MARTINEZ-CONDE S.: Effect of stimulus width on simultaneous contrast. *PeerJ* 1 (2013), 146. 4
- [TFS08] TOMINSKI C., FUCHS G., SCHUMANN H.: Task-driven color coding. In *Proceedings of the 12th International Conference on Information Visualisation* (2008), IEEE, pp. 373–380. 3
- [Tre85] TREISMAN A.: Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing* 31, 2 (1985), 156–177. 1
- [Van12] VANDERBEI R. J.: 2012 presidential election results, 2012. (visited on 11/28/2013). URL: <http://www.princeton.edu/~rvdb/JAVA/election2012/>. 8, 9
- [War88] WARE C.: Color sequences for univariate maps: Theory, experiments and principles. *IEEE Computer Graphics and Applications* 8, 5 (1988), 41–49. 1, 2, 6, 7, 8
- [WRM*09] WANNER F., ROHRDANTZ C., MANSMANN F., STOFFEL A., OELKE D., KRSTAJIC M., KEIM D. A., LUO D., YANG J., ATKINSON M.: Large-scale comparative sentiment analysis of news articles. Presented at the poster session at IEEE Information Visualization Conference 2009, Atlantic City, USA, 2009. 9
- [You87] YOUNG R. A.: The Gaussian derivative model for spatial vision: I. retinal mechanisms. *Spatial Vision* 2, 4 (1987), 273–293. 4