SPRWeb: Preserving Subjective Responses to Website Colour Schemes through Automatic Recolouring

David R. Flatla[1], **Katharina Reinecke**[2], **Carl Gutwin**[1] and **Krzysztof Z. Gajos**[2]

[1]University of Saskatchewan
Saskatoon, Saskatchewan, Canada

[2]Harvard University
Cambridge, Massachusetts, USA

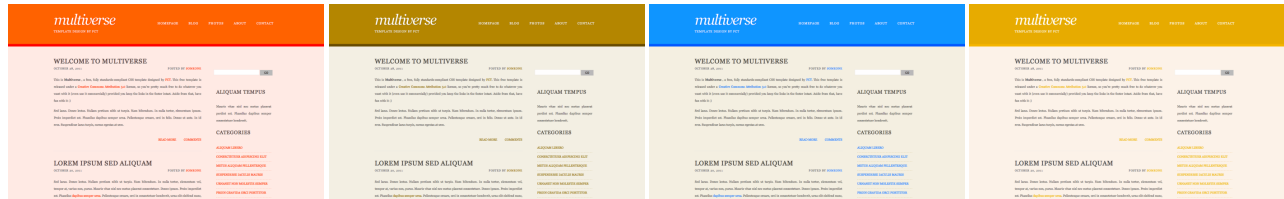david.flatla@usask.ca, reinecke@seas.harvard.edu, gutwin@cs.usask.ca, kgajos@seas.harvard.edu

Figure 1. Left to right: Website with 'warm' colour scheme, deuteranopic CVD simulation, standard recolouring, SPRWeb recolouring.

ABSTRACT

Colours are an important part of user experiences on the Web. Colour schemes influence the aesthetics, first impressions and long-term engagement with websites. However, five percent of people perceive a subset of all colours because they have colour vision deficiency (CVD), resulting in an unequal and less-rich user experience on the Web. Traditionally, people with CVD have been supported by recolouring tools that improve colour differentiability, but do not consider the subjective properties of colour schemes while recolouring. To address this, we developed SPRWeb, a tool that recolours websites to preserve subjective responses *and* improve colour differentiability – thus enabling users with CVD to have similar online experiences. To develop SPRWeb, we extended existing models of non-CVD subjective responses to CVD, then used this extended model to steer the recolouring process. In a lab study, we found that SPRWeb did significantly better than a standard recolouring tool at preserving the temperature and naturalness of websites, while achieving similar weight and differentiability preservation. We also found that recolouring did not preserve activity, and hypothesize that visual complexity influences activity more than colour. SPRWeb is the first tool to automatically preserve the subjective and perceptual properties of website colour schemes thereby equalizing the colour-based web experience for people with CVD.

Author Keywords

colour; colour vision deficiency; automatic recolouring

ACM Classification Keywords

K.4.2 Social Issues: Assistive technologies

INTRODUCTION

Website colours influence our emotional reactions to the Web – they help us identify familiar brands and products, as well as contribute to the aesthetic appeal of websites. In addition to familiarity and aesthetics, colour schemes also help convey numerous subjective qualities of a website (e.g., the 'warmth', 'busyness' and 'heaviness' of a website) [6, 2].

While designers have long realized the role website colours play, it has generally been assumed that all users have identical subjective responses to website colour choices. However, five percent of the world's population has some degree of colour vision deficiency (CVD) [4], that restricts their colour perception to a subset of the colours perceived by people with typical colour vision. As a result, users with CVD have an unequal and presumably less-rich user experience on the Web.

Recolouring tools have been developed that modify website colours to make them more accessible for people with CVD by substituting original colours with perceptually-similar (to maintain naturalness) yet suitably-differentiable replacement colours. However, no existing recolouring tools are designed to also preserve the subjective-response properties of the original website colour scheme for users with CVD.

To address this, we have developed a new recolouring tool called SPRWeb (pronounced SuPeRWeb: the Subjective response–Preserving Recolouring tool for Websites). SPRWeb utilizes a two-pass, hill-climbing optimization that finds replacement colours for websites that reflect the accessibility goals of previous recolouring tools, but also preserves the original colour's subjective responses. By preserving the naturalness, differentiability and subjective-response properties of the original website colours, SPRWeb is the first recolouring tool designed to provide users with CVD an experience equivalent to that of people with typical colour vision.

In a user study, we compared the perceptual (naturalness, differentiability) and subjective-response (activity, temperature, weight) preservation performance of SPRWeb to that of a recently-published, publicly-available recolouring tool by

Kuhn et al. [18]. We found that SPRWeb was significantly better at maintaining naturalness and temperature, and was no worse at maintaining differentiability and weight. We also found that activity was not influenced by either recolouring tool, possibly because participants rated the visual complexity of the website instead of its colour when reporting activity.

This work makes three contributions. First, it provides a new tool, SPRWeb, that improves the accessibility of website colours for users with CVD while also preserving the subjective experience of the original site. Second, it provides empirical evidence that SPRWeb performs significantly better than the current state of the art for preserving website temperature and naturalness, and no worse for preserving weight, activity and differentiability. Third, it provides designers with a new understanding of subjective responses to colour choices for the large population of users with CVD, and shows that existing models can be used to predict responses for both users with CVD and those with typical colour vision.

BACKGROUND
Colour Vision Deficiencies
Trichromatic human colour vision begins with the stimulation of three retinal photoreceptors (cones) that are each sensitive to different wavelengths [4, 26]. Colour vision deficiency arises when one of the cones is either missing (*dichromacy*) or has a shifted peak sensitivity (*anomalous trichromacy*) [24]. Human colour perception can be described by two orthogonal axes – a red-green axis and a blue-yellow axis. The most common forms of inherited CVD (>99.9% [1]) result in reduced discrimination along the red-green axis (full reduction in dichromacy, partial reduction in anomalous trichromacy [4]). We will focus SPRWeb on red-green CVD due to its prevalence, but nothing in SPRWeb's design prevent its extension to blue-yellow CVD or monochromatism (greyscale CVD resulting from two or three missing cones).

Accessibility-Enhancing Recolouring Tools
Many recolouring tools have been developed, but few address website recolouring directly. We recognize that image-based recolouring tools are able to recolour an image of a website, but this necessarily removes the interactive content (e.g., drop-down menus) of the website. As such, we will only discuss one image-based recolouring tool here.

Ichikawa et al. [16] published the first website recolouring tool that improves accessibility for people with CVD, but this tool does not attempt to maintain the original colours, nor their subjective responses. Iaccarino et al. [15] proposed the use of *edge services* to recolour website images in transit to the user, but this uses a recolouring technique that only enhances colour contrasts, and does not try to maintain naturalness, differentiability, or subjective responses. Nam et al. [20] proposed a system that attempts to maximize differentiability for CVD users while also steering the perceived temperature of the output, but does not attempt to maintain the original subjective responses. Extensions for Chrome (`www.daltonize.org`) and Firefox (`www.emblazonlife.com`) have also been released, but these only recolour website images and text to enhance differentiability, so do not address recolouring entire website colour schemes.

A recent and publicly-available image recolouring tool preserves the original image colours as much as possible (naturalness) while restoring the differentiability of colours for people with CVD [18]. This tool's public availability make it likely to be used by people with CVD, and its naturalness preservation is likely to also preserve some subjective response properties. As such, we adopted this recolouring tool as our benchmark for evaluating SPRWeb.

Modifying Subjective Responses to Images
Modifying images to change viewers' subjective responses has been explored by utilizing an exemplar image to recolour a given image. In general, these systems recolour an input image so that its constituent colours closely match the colours of an exemplar [12, 22] and this has been extended to allow the user to select an exemplar colour according to its 'mood' [28]. These techniques differ from what we are proposing here in two key ways: they do not seek to improve accessibility, and they do not attempt to specifically define the target mood for the recoloured image, opting to simply copy the mood from the exemplar image (or colour) to the recoloured image. One recent system does modify images to adopt a user-specified mood [27], but this system does not improve differentiability.

Subjective Responses to Colour
Colour is one of the first features users notice when looking at a webpage [13], and colour has been found to influence a website's perceived trustworthiness and visitor loyalty [6]. In addition, orange background colours have been found to have a positive effect on users' memory of a website's content compared to blue and grey [2], however, websites with a blue background were perceived to load more quickly than ones with a yellow or red background [10]. Warm colours have been found to consistently evoke more active feelings than cool colours [17, 13]. It is also widely accepted that the colour red speeds up the heart rate, thus eliciting excitement, serving as an attention grabber [10, 17]. In contrast, blue elicits relaxed feelings [5, 17]. Related to these findings, Gorn et al. found that ads containing higher levels of *chroma* (colourfulness) elicit greater feelings of excitement [11]. For *value* (lightness), higher-value background colours (e.g., pastels) in ads induce more relaxed feelings [11].

Based on these findings about particular hues and colour properties, models that predict how people subjectively respond to colours have been developed. Valdez and Mehrabian [25] developed a three-part model that relates the pleasure, arousal, and dominance of a colour to its hue, brightness and saturation. More recently, Ou et al. [21] developed a three-part model that uses the hue, chroma, and luminance of a colour to predict the colour's perceived temperature, activity, and weight. More details on this model are given now.

MODELS OF SUBJECTIVE RESPONSES TO COLOUR
The development of models of subjective responses to colour usually follows three steps. First, researchers gather empirical data from human participants on several subjective dimensions for a set of single colours. The subjective dimensions include scales such as tense-relaxed, hard-soft, or fresh-stale; often several scales with related semantic endpoints are given.

Activity (active-passive):
$$AP = -2.1 + 0.06\sqrt{(L^* - 50)^2 + (a^* - 3)^2 + (\frac{b^* - 17}{1.4})^2}$$
Temperature (warm-cool):
$$WC = -0.5 + 0.02(C^*)^{1.07}cos(H^* - 50°)$$
Weight (heavy-light):
$$HL = -1.8 + 0.04(100 - L^*) + 0.45cos(H^* - 100°)$$
Chroma: Hue:
$$C^* = \sqrt{(a^*)^2 + (b^*)^2} \qquad H^* = arctan(b^*, a^*)$$

Figure 2. Ou et al. [21] models for activity, weight, and temperature.

The second step involves factor analysis or principal components analysis to determine a smaller set of factors that best represent the larger set of dimensions. Third, models are built using the empirical data to predict people's responses to single colours along the reduced set of factors.

Following these steps, Ou et al. developed a model that predicts the subjective judgments of colour activity, temperature and weight [21]. Their model translates the colorimetric properties of any colour (expressed in CIE L*a*b* colour space coordinates [24]) into a prediction of the user's response along the three dimensions of activity, temperature, and weight. These functions are shown in Figure 2.

Activity (active-passive) predictions are calculated using the relative luminance (L*) and chroma of the colour compared to a reference muddy-yellow colour (L*a*b* = (50,3,17)). This colour has the most passive (negative) value and the farther colours are from it in L*a*b* space, the more active (positive) their predicted score is. Temperature (warm-cool) takes as input the chroma of the colour as well as its hue angle. Chroma determines the magnitude of the temperature prediction, and hue determines the sign – negative values indicate 'cool' colours, and positive values indicate 'warm' colours. Weight (heavy-light) takes as input the relative luminance and hue of the colour. Dark colours are predicted as 'heavy' (positive) and light colours are predicted as 'light' (negative).

Researchers have also used responses to single colours as the basis for models of responses to multiple colours [23]. The general finding of this work is that the responses to multiple colours is the average of each of the colours weighted by their occurrence. In an image, this amounts to averaging the subjective responses over all pixels.

As Ou et al's [21] and Solli and Lenz's [23] models are the most recently-published colour subjective response models – and they use the CIE L*a*b* colour space, which is commonly used in recolouring tools – we employed them in the development of SPRWeb. As a first step, we now present a formative study that extends Ou's model to people with CVD.

CVD USERS' SUBJECTIVE RESPONSES TO COLOUR
The models of subjective responses to colour described above are based on data collected from people with typical colour vision, but SPRWeb is intended to help people with CVD. We performed a formative study to investigate whether these existing models can be extended to individuals with red-green CVD. In this study, we first verified that Ou et al.'s model [21] is valid for a new sample of non-CVD participants, and then identified colours which elicited similar subjective responses from CVD and non-CVD participants. We hypothesized that the existing model holds for the new sample of people without CVD and that the model accurately predicts the subjective response of participants with CVD for colours that people with and without CVD perceive identically.

Apparatus and Procedure
We developed a testing system to gather participant responses to individual colours. Colours were presented as 10 cm circles on an L*=50.0 grey background. Participants rated each colour on a 7-point semantically-anchored scale in terms of three dimensions: active-passive (activity), warm-cool (temperature), and heavy-light (weight), following the main factors determined by Ou et al. [21].

Colours Presented
As discussed above, red-green CVD manifests as a compression of the red-green colour perception axis. Previous research has used findings from unilateral dichromats (individuals who are dichromatic in one eye, but are trichromatic in the other) to identify the set of colours that are perceived identically in dichromatic colour vision and typical colour vision. This set of colours is defined by a 575nm wavelength yellow, a 475nm wavelength blue, and the entire set of achromatic colours (greys) [3]. In a radial colour space with the achromatic axis running vertically up the middle, these wavelengths form two half-planes of colours when intersected with the achromatic axis, which can be approximated by a single plane of colours in CIE L*a*b* colour space [19] (the dichromacy-trichromacy equivalence plane, or DTEP).

The DTEP defines the set of colours perceived identically by people with dichromatic CVD and typical colour vision. It also defines the entire set of colours perceived by people with dichromacy – all colours visible to people with typical colour vision 'map' to a representative colour on the DTEP for people with dichromacy. This explains why people with red-green CVD often cannot distinguish between colours that people with typical colour vision can. If two distinct colours map to an identical colour on the DTEP, they will be indistinguishable for someone with dichromacy.

Anomalous trichromacy results in a partial compression of the red-green colour perception axis; the more severe the anomalous trichromacy, the greater the compression [8]. Individuals with anomalous trichromacy experience a larger gamut of colours than those with dichromacy, but all colours are 'shifted' toward the DTEP. Colours that lie on the DTEP are not shifted, and are perceived identically in trichromacy and anomalous trichromacy. As a result, trichromatic, anomalous trichromatic, and dichromatic individuals all perceive colours on the DTEP identically.

Twenty-four colours were chosen uniformly from the DTEP and presented to the user three times in random order, once for each dimension of Ou et al.'s model. Each dimension presentation series began with five training trials to give 3x(5+24)=87 total trials.

Participants

We recruited 24 male participants (mean age 27.1 years) from the local community. Twelve participants had CVD (all red-green CVD), and twelve had typical colour vision. All participants were given the HRR colour-vision test [14] at the start of the session, and in all cases the results of the test matched the participants' statements about their own vision.

Study Design

The study measured three dependent variables – the participant scores for activity, temperature, and weight – for both CVD and non-CVD participants.

Our analysis then focused on two questions:

1. Do CVD and non-CVD participants have similar subjective responses to DTEP colours?

2. Does Ou's model predict CVD and non-CVD participant subjective responses to DTEP colours?

Results

Using the Pearson product-moment correlation coefficient, we determined the correlation between CVD and non-CVD participant responses and between predicted responses from Ou's model and our own CVD and non-CVD participants' responses. Our results showed a significant correlation between the subjective responses of CVD and non-CVD participants for the three dimensions of Ou's model (activity: r=.82, temperature: r=.96, and weight: r=.95, all p<.001).

We also found that the responses of CVD and non-CVD participants significantly correlated with Ou's model predictions. Specifically, non-CVD participants' responses were strongly correlated with Ou's model for activity (r=.80), temperature (r=.94), and weight (r=.91), all with p<.001. The correlations were similar for our participants with CVD for activity (r=.56, p<.01), temperature (r=.96, p<.001), and weight (r=.95, p<.001). We note that the correlation, although significant, is substantially lower for activity for participants with CVD. This may indicate a minor misalignment between Ou's model predictions and CVD responses. As an estimate, we compared the standard deviations of CVD and non-CVD responses for each colour and found that CVD participant responses had significantly higher deviations than non-CVD responses (p<.05). However, the correlations between CVD responses and Ou's model are still significant for activity, so we use this model to steer the recolouring algorithm of SPRWeb.

These results indicate that Ou's model will work well to predict the subjective responses of individuals with CVD for DTEP colours, affording SPRWeb a predictive model upon which to base its recolouring decisions. We now present the design and implementation details of SPRWeb.

SUBJECTIVE RESPONSE-PRESERVING RECOLOURING

Three steps are necessary to recolour a website: [1]

1. Process the cascading style sheet (CSS) file and extract every colour specified therein.

---

2. Find a mapping from the original colours to replacement colours that meet particular criteria.

3. Apply the replacement colours to the CSS file.

We will now explore each of these steps as implemented in SPRWeb in more detail.

Extract Colours from CSS File

Modern website design guidelines advocate the separation of website content from the website look and feel. To achieve this, Cascading Style Sheets (CSS) are used to specify the layout and colours of a website. To extract colours listed in a CSS file, SPRWeb utilizes regular expressions to find colours specified using hexadecimal, RGB, RGBA, HSL, HSLA, and explicitly named formats. Any alpha specifications (transparency) are saved for later, and each identified colour is converted to Java int RGB format (8-bits per red, green, and blue channel), and stored. Duplicate colours are only stored once, as all occurrences of a single colour (regardless of alpha) in a website should be recoloured to the same replacement colour.

Determine Replacement Colours

Following identification of each unique colour in the CSS file, SPRWeb selects a set of replacement colours from the dichromat-trichromat equivalency plane (DTEP). DTEP colours are chosen for three reasons. First, people with dichromatic CVD only perceive DTEP colours, so this is a natural choice for recolouring. Second, existing methods for computing colour differentiability for people with typical colour vision (e.g., Euclidean distance in CIE L*a*b*) are applicable to people with red-green CVD for DTEP colours. Third, our findings from the formative study show that people with red-green CVD have similar subjective responses as people with typical colour vision for DTEP colours.

To select DTEP replacement colours, we developed a constraint optimization technique that seeks to minimize a cost function that is computed by calculating a weighted sum of four individual costs derived from comparing the original colour set to a set of potential replacement colours. We now describe these four individual costs.

Perceptual naturalness: To help prevent dramatic shifts in colour during recolouring, the replacement colours should be as perceptually similar to the original colours as possible. SPRWeb utilizes Euclidean distance between two colours expressed in the CIE L*a*b* colour space [24] to represent perceptual difference. The perceptual naturalness cost (pn) value is the average CIE L*a*b* Euclidian distance between each original colour and its corresponding replacement colour:

$$pn = \frac{1}{N} * \sum_{i=1}^{N} \Delta_{Lab}(O_i, R_i)$$

where N is the number of original colours, Δ_{Lab} is the CIE L*a*b* Euclidean distance, and O and R are the original colour sets and replacement colour sets, respectively.

Perceptual differentiability: In addition to perceptual naturalness, the differentiability between all original colours needs to be maintained in the replacement colour set. This maintains relationships between colours (e.g., text and background

---

[1] For simplicity, we assume that website images are not recoloured and that the entire website colour scheme is specified in its CSS file.

should be differentiable) and can be in opposition to perceptual naturalness (e.g., if two very distinct original colours map to the same DTEP colour to maintain perceptual naturalness). The perceptual differentiability cost (pd) value is the average absolute difference between the CIE L*a*b* Euclidean distance between every original colour and the CIE L*a*b* Euclidean distance between every potential replacement colour:

$$pd = \frac{2 * \sum_{i=1}^{N} \sum_{j=i+1}^{N} abs(\Delta_{Lab}(O_i, O_j) - \Delta_{Lab}(R_i, R_j))}{N * (N-1)}$$

Subjective-response naturalness: To preserve the non-CVD subjective responses (e.g., keeping warm colours warm) of the original colours for people with red-green CVD, SPRWeb uses a cost function similar to perceptual naturalness, but using the subjective response model developed by Ou et al. [21]. SPRWeb calculates the subjective-response difference between two colours by determining the temperature, activity, and weight for each colour, and computing the Euclidean distance between the colours as though temperature, activity, and weight occupied three orthogonal dimensions of a 'subjective-response space'. Although this technique is unproven, this 'subjective-response space' has been presented in previous research that utilized Ou et al's model [23], and we found it to work very effectively in pilot studies. To allow comparability between this cost value and the perceptual cost values described above, the 'subjective-response space' is scaled to have identical extreme values as the CIE L*a*b* gamut for the entire set of sRGB colours. The subjective-response naturalness cost (srn) is calculated by:

$$srn = \frac{1}{N} * \sum_{i=1}^{N} \Delta_{Ou}(O_i, R_i)$$

where Δ_{Ou} is the scaled Euclidean distance in the 'subjective-response space' described above.

Subjective-response differentiability: Similar to perceptual differentiability described above, the subjective-response differentiability needs to be maintained throughout recolouring as well. This helps preserve subjective-response relationships between colours (e.g., colour A is more active than colour B). It utilizes the same scaled 'subjective-response space' as the subjective-response naturalness value, and is calculated by:

$$srd = \frac{2 * \sum_{i=1}^{N} \sum_{j=i+1}^{N} abs(\Delta_{Ou}(O_i, O_j) - \Delta_{Ou}(R_i, R_j))}{N * (N-1)}$$

Utilizing these four cost values, the weighted cost function for a potential replacement colour set can be computed allowing evaluation and comparison of potential replacement colour sets. This total cost function is calculated by:

$$cost = W_{pn} * pn + W_{pd} * pd + W_{srn} * srn + W_{srd} * srd$$

where W_{xx} is the real-valued weighting factor for each cost value. Through pilot studies, we determined that $W_{pn} = W_{pd} = 1.0$ and $W_{srn} = W_{srd} = 2.0$ worked well.

Constraint optimization algorithm
SPR uses a two-pass hill-climbing algorithm to solve this optimization problem. For the first pass, a subset of 900 uniformly sampled DTEP colours are used to find an approximate solution. In the second pass, each replacement colour in the approximate solution is replaced with the set of DTEP colours that are less than five L*a*b* units from the original sample colour and a refined solution is computed. The second set of colours is designed to be large enough to facilitate the identification of an optimal or near-optimal set of replacement colours, but small enough to allow rapid stabilization.

In each step, a hill-climbing algorithm is used to find a set of replacement colours that minimizes the cost function described above. First, a set of possible replacement colours is generated (either the 900 DTEP subsample colours for the first pass, or the expanded sets of subsampled colours for the second pass). A set of replacement colours is then drawn randomly from the set of possible replacement colours. Each current replacement colour is then iteratively replaced with every one of the possible replacement colours. If the replacement reduces the cost function, then the replacement is kept, otherwise it is rejected. This process is repeated until two consecutive replacement sets are identical. Before each iteration, the order of the possible replacements is shuffled to help the algorithm stabilize and avoid local minima.

The constraint optimization algorithm converges quickly (typically < 100 iterations): on a 2.0 GHz machine with 4 GB RAM, the algorithm took less than one second for ten colours, and took about ten seconds for thirty colours.

Once the set of replacement colours have been identified, a Java Map data structure from original int RGB colours to replacement int RGB colours is generated. This map is used in the final recolouring step.

Apply Replacement Colours to CSS File
Using the original-to-replacement colour map generated above, the original CSS text is modified to replace any colour values, using six-digit hexadecimal format. This results in a modified CSS file that uses the replacement colours. When the website's HTML file is rendered by the browser, it will present the recoloured version of the website optimized to preserve perceptual and subjective-response naturalness and differentiability. An example SPRWeb recolouring is shown in the left (original) and right (SPRWeb) images in Figure 1.

EVALUATION: WIREFRAME WEBSITES
We evaluated SPRWeb by performing two studies that each examined the accessibility performance, theoretical performance, and user-study performance of SPRWeb compared to Kuhn et al's [18] recent and publicly-available recolouring tool. Kuhn's recolouring tool is an ideal comparator for SPRWeb because it is designed to maximally preserve naturalness while improving colour differentiability. Preserving naturalness and maintaining subjective responses can be complementary in many cases: if the original website colours can be preserved, then the original website colour subjective responses will also be preserved in many cases. Kuhn's recolouring tool also preserves the luminance of original
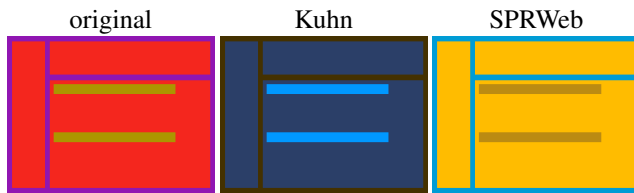
Figure 3. The original (left), Kuhn-recoloured (middle), and SPRWeb-recoloured (right) versions of the 'warm-active-heavy' wireframe website used in the wireframe website study.

Figure 4. Naturalness: Mean distance from original colours for Kuhn and SPRWeb ($\pm$ s.e.). SPRWeb generated replacement colours significantly closer to the original colours. Differentiability: Difference from original colour pair differences for Kuhn and SPRWeb ($\pm$ s.e.). SPRWeb maintains original colour differences at least as well as Kuhn.

colours during recolouring, potentially benefitting the maintenance of Ou's 'weight' dimension, as it is strongly correlated with luminance (according to Figure 2).

Websites are often visually complex presentations containing textual and graphical content, organized in a particular layout, using headers, tables, and lists each with their own fonts and styles. As such, the two studies used different levels of website complexity. The first study utilized abstract wireframes of websites – similar to 'back of the napkin' prototypes quickly generated by website designers to test multiple layouts and colour schemes. The second study utilized full websites gathered from CSS template collections on the Web. The design and results from the first study are presented now.

Wireframe Websites Stimuli
To generate the wireframe websites for this study, we generated a three-colour website wireframe with a simple layout containing no text or images (Figure 3). To generate multiple versions of this website wireframe, eight colour schemes were applied to the website wireframe, giving eight original websites. To exercise the subjective response maintenance of each recolouring tool, we chose colour schemes that contained colours located at the extremes of Ou's subjective response model [21]. This model contains three dimensions (activity, temperature, weight), so all combinations of extremes at each end of each dimension gives eight total colour schemes. These colour schemes were chosen by digitizing the 532 three-element colour schemes in the Pantone Guide to Communicating with Color [7]. Each colour scheme was analyzed using Solli and Lenz's extension of Ou's model to colour images [23] to determine the extreme colour schemes.

In order to conduct comparisons, Kuhn and SPRWeb recoloured versions of the websites were needed. To generate Kuhn's, images of the eight original wireframe websites were submitted to Kuhn's image recolouring tool using the settings 'deutan' for CVD type, 'no' for exaggerate recolouring, and '32' for number of quantization levels. The original colour schemes were then submitted to SPRWeb to generate a replacement set of colours, which were then applied to the website wireframe to generate SPRWeb-recoloured versions.

Naturalness and Differentiability
SPRWeb is designed to preserve naturalness and improve colour differentiability in addition to maintaining subjective responses. As such, we first compared the naturalness-preservation and differentiability-restoration performance of SPRWeb to Kuhn's recolouring tool.
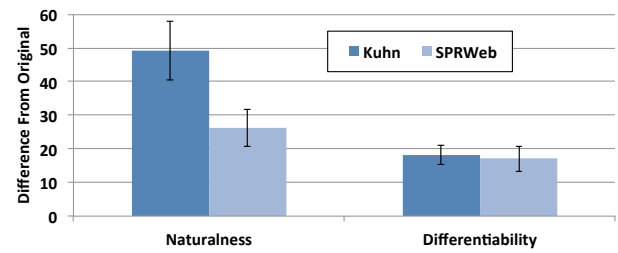
We converted the colours in each of the original, Kuhn, and SPRWeb wireframes to CIE L*a*b* coordinates. L*a*b* colour space is a perceptually-uniform colour space in which Euclidean distance is directly proportional to perceptual difference. To assess naturalness-preservation, we compared the L*a*b* Euclidean distance between each original colour and its respective Kuhn and SPRWeb replacement colours. Lower distances represent better naturalness-preservation, i.e., the replacement colour is more perceptually-similar to the original colour. This was performed for each colour in each original website and the average results are presented in Figure 4, left. Pairwise two-tailed t-tests (alpha = .05) showed that SPRWeb found replacement colours that are closer to the original colours than Kuhn's tool (p<.01).

To test differentiability, we calculated the Euclidean distance between each pair of colours in each original website and determined the absolute difference between this distance and the equivalent colour pair in the Kuhn and SPRWeb versions. Lower values represent better differentiability-restoration, i.e., the replacement colours are as differentiable (not more or less differentiable) as the original colours. Figure 4 (right) shows average differentiability by recolouring tool. No significant differences were found (p=.78).

Theoretical Performance
If Ou's model [21] correctly predicts the perceived activity, temperature, and weight of a website, then it can be used to make predictions about the user-study performance of SPRWeb. As discussed above, Ou's model predicts subjective responses for individual colours, and Solli and Lenz's extension of Ou's predictive model [23] predicts subjective responses for entire images. As websites can be represented as images, we used this latter model for this study. We determined the activity, temperature, and weight predictions of each original wireframe website as well as the Kuhn and SPRWeb-recoloured versions. These values were compared for each website by finding the absolute difference between the original predictions and the Kuhn and SPRWeb predictions. The average across all websites of these differences are presented in Figure 5. Although no significant differences were found using Bonferroni-corrected pair-wise two-tailed t-tests (alpha = .017; Activity: p=.02, Temperature: p=.03, Weight: p=.17), SPRWeb produced replacement colours with
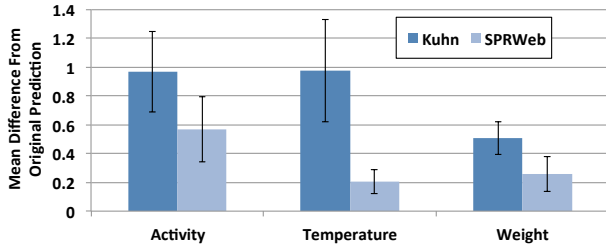
Figure 5. Mean differences between predictions of participant subjective responses ($\pm$ s.e.) for original and recoloured wireframe websites.
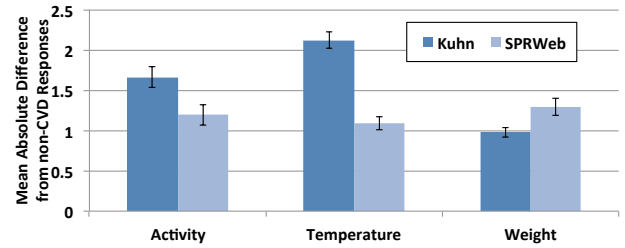

Figure 6. Mean absolute difference between CVD participant responses to recoloured websites and mean non-CVD participant responses to original websites ($\pm$ s.e.).

predicted subjective responses substantially closer to the original colour subjective responses.

Based on this theoretical analysis (and the assumption that the underlying models are correct) we hypothesized that SPRWeb would improve on Kuhn's preservation of activity, temperature, and weight for wireframe recolouring.

User-Study Performance

Utilizing the experimental design and participants (24; 12 with CVD) from the formative study, we conducted a user study to compare participants' reported subjective responses for the original and recoloured websites. We modified the single-colour system to display the website wireframes described above. Participants rated the temperature, activity, and weight of each website wireframe on a seven-point semantically-anchored scale. Participants with typical colour vision rated the original website wireframes ((8 websites + 5 training) x 3 dimensions = 39 trials). Participants with CVD reported their subjective responses for the Kuhn and SPRWeb-recoloured website wireframes ((16 websites + 5 training) x 3 dimensions = 63 trials).

We compared the non-CVD responses for original websites to CVD participant responses for the recoloured websites by finding the absolute difference between each CVD response for a recoloured website wireframe and the average non-CVD response for the corresponding original website wireframe. These differences were averaged over all website wireframes, and are presented in Figure 6.

RM-ANOVA showed a significant main effect of Recolourer on the score difference ($F_{1,11}$=23.12, p<.005), but there was also a significant interaction between Recolourer and Dimension ($F_{2,22}$=37.46, p<.001). Post-hoc pairwise t-tests (2-tailed, Bonferroni corrected) showed that Kuhn and SPRWeb were significantly different on each dimension (all p<.016). As shown in Figure 6, CVD participants' responses for Temperature and Activity were significantly closer to the non-CVD mean with SPRWeb. For Weight, the responses were significantly closer to Kuhn's recolourer but the magnitude of the difference was smaller than for Temperature and Activity.

Examination of the responses supports these findings. Pearson product-moment correlation coefficients showed that the CVD participant responses to SPRWeb-recoloured website wireframes closely correlated with non-CVD participant responses to the originals for Temperature (SPR: r=.88, Kuhn: r=.69) and Activity (SPR: r=.93, Kuhn: r=-.37), but re-

sponses to Kuhn-recoloured website wireframes were closer for Weight (SPR: r=.73, Kuhn: r=.97).

Given that there was a difference between our predicted results for Weight and the results determined from the user study, we opted to modify SPRWeb before proceeding with the full-website study. The only substantial difference between the optimization algorithms for Kuhn and SPRWeb (besides SPRWeb's obvious inclusion of subjective response maintenance components) is that Kuhn's recolouring tool strictly enforces luminance preservation during recolouring. This restriction limits the optimization search to isoluminant cross-sections of the DTEP, allowing substantial performance gains. As SPRWeb did not achieve superior results using only Ou's predictions of Weight, we added a luminance preservation component (lm) to the cost function of SPRWeb:

$$lm = \frac{1}{N} * \sum_{i=1}^{N} abs(lum(O_i) - lum(R_i))$$

where lum finds the luminance of the supplied colour. To prevent this addition from overly-constraining the recolouring process, we opted to make this component not as strict as Kuhn's luminance preservation method, allowing SPRWeb to deviate from the original colour's luminance. To achieve this, we utilized a weighting factor of 1.1 for this component.

EVALUATION: REALISTIC WEBSITES

After adding the luminance preservation component to SPRWeb, we compared the subjective response performance of SPRWeb and Kuhn's recolouring tool for full websites.

Full Website Stimuli

To generate full websites for this study, we took three image-free 'representative' websites from `http://templated.org`. As using three websites to represent the rich diversity of the Web is potentially oversimplifying, we carefully chose one information page, one forum-style page, and one blog; each with either a light, medium, or dark colour scheme (see Figure 7). The light website contained 16 unique colours, the medium had 31, and the dark had 17 colours.

We applied one of six colour themes to each of these websites to generate 18 websites in total. To apply a colour theme, we parsed the CSS file for each website and extracted each colour. Each colour was converted to luminance, chroma, and hue angle values. The hue angle was set to one of 0, 20, 160, 180, 200, or 340 and the chroma was doubled. The hues

Figure 7. 'Light' website with 20° colour theme (left); 'medium' website with 180° colour them (middle); 'dark' website with 200° colour theme.
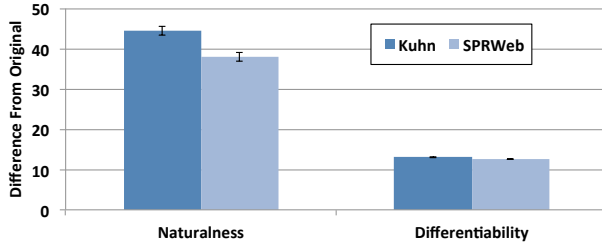

Figure 8. Naturalness: Mean distance from original colours for Kuhn and SPRWeb ($\pm$ s.e.). SPRWeb generated replacement colours significantly closer to the original colours. Differentiability: Difference from original colour pair differences for Kuhn and SPRWeb ($\pm$ s.e.). SPRWeb maintains original colour differences significantly better than Kuhn.

chosen ensure that recolouring will be performed because these hues represent colours that pose difficulties for individuals with red-green CVD. We amplified the chroma of each colour to increase the colourfulness of each website, helping elicit strong subjective responses from Ou's model and participants. Note that luminance was not modified, thereby preserving each original websites' lightness.

Keeping with the website wireframe study, each of the 18 original websites was then recoloured using Kuhn's recolouring tool and SPRWeb. We used the same settings for Kuhn as above (deutan, no exaggeration, 32 levels) to generate Kuhn-recoloured images of each website's CSS colours. The resulting replacement colour set was applied to the CSS file to create a Kuhn-recoloured website. Each original website was also passed to SPRWeb, which performed its standard recolouring as described above. Every original and recoloured website was then rendered in a browser and screenshots were taken to be used as stimuli in the user study.

Naturalness and Differentiability
We compared the naturalness-preserving and differentiability-restoring performance of Kuhn's recolouring tool and SPRWeb. Colours for each website were extracted from the CSS file for each original and recoloured website. These colours were analyzed as in the wireframe website study and the results are presented in Figure 8. Paired two-tailed t-tests found significant differences for naturalness (p<.01) and differentiability (p<.01).

Theoretical Performance
In the theoretical performance evaluation for the wireframe website study, we found that the subjective response predictions using Solli and Lenz's model [23] for weight did not correspond with the findings from the user study. In spite of this, the predictive model did accurately predict user responses for activity and temperature. As such, we once again developed hypotheses regarding the user study results using
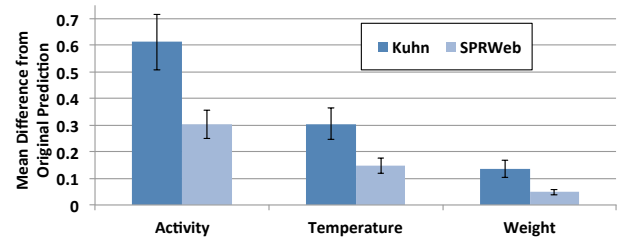

Figure 9. Mean differences between predictions of participant subjective responses ($\pm$ s.e.) for original and recoloured full websites.

this model for the current study. We analyzed the Sollie and Lenz model predictions for the original, Kuhn-recoloured, and SPRWeb-recoloured websites and recorded the absolute difference between original website predictions and predictions for each recoloured website for activity, temperature, and weight. The averages across all websites of these differences are presented in Figure 9. Using paired two-tailed t-tests (alpha = 0.017), these differences were shown to be significant (Activity: p<.001, Temperature: p=0.01, Weight: p<.001), in favour of SPRWeb.

Based on these theoretical performance results, and in consideration of the modifications stimulated by the wireframe website theoretical evaluation, we hypothesize that SPRWeb will produce recoloured full websites that reflect the activity and temperature of the original full website more accurately than Kuhn's recolouring tool. Based on the added luminance preservation parameter, we also predict that there will be no difference between the Kuhn-recoloured and SPRWeb-recoloured full websites for the weight subjective response.

User-Study Performance
We extended the experimental design from the wireframe website user study to incorporate the full websites described above. We recruited 24 new participants (12 with CVD, according to the HRR Plate Test [14]) from the community and had them report their subjective responses to the websites. Participants once again rated temperature, activity, and weight on a seven-point semantically-anchored scale. Participants with typical colour vision rated the original websites ((18 websites + 5 training) x 3 dimensions = 69 trials). Participants with CVD rated the recoloured websites ((36 websites + 5 training) x 3 dimensions = 123 trials).

We compared the non-CVD responses for original websites to CVD participant responses for the recoloured websites (absolute difference between each CVD response for a recoloured website and the average non-CVD response for the corresponding original). These differences were averaged over all websites, and are presented in Figure 10.

Comparing the results for each recolouring tool in each dimension using paired two-tailed t-tests, CVD participants' responses for Temperature were significantly closer to the non-CVD mean with SPRWeb (p<.001), and there was no significant difference between SPRWeb and Kuhn's recolouring tool for Weight. In contradiction to our predictions however, Kuhn's recolouring tool results in slightly better preservation of Activity, although this difference is not significant (p=.21).
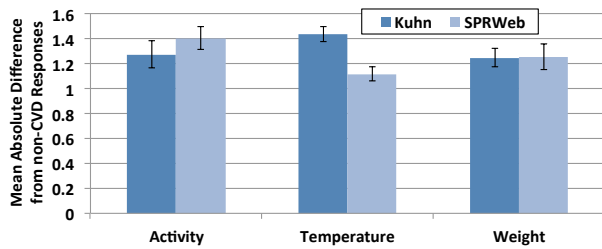
Figure 10. Mean absolute difference between CVD participant responses to recoloured websites and mean non-CVD participant responses to original websites (± s.e.).

Using Pearson product-moment correlation coefficients between average responses to original websites for non-CVD participants and recoloured websites for participants with CVD, the Temperature (SPR: r=.79, Kuhn: r=.16) and Weight (SPR: r=.82, Kuhn: r=.87) findings are supported, but correlations for Activity are both low (SPR: r=.44, Kuhn: r=.24). Kuhn and SPR recoloured websites did not elicit subjective responses that correlate with the original website responses.

Exploring correlations between the predicted responses from Solli and Lenz's model [23] and the subjective responses from the user study for Activity, we found that there was essentially no positive correlation in these relationships (original: -.25, Kuhn: .03, SPRWeb: -.47). As these were quite surprising results, we now turn to discussing the meaning, implications, and generalizations of these results for preserving subjective responses through automatic recolouring.

DISCUSSION

Our results demonstrate that:

1. Standard models of subjective responses to colour can be used in automatic recolouring tools, because CVD users respond similarly to people with typical colour vision;

2. SPRWeb performs much better than the state-of-the-art Kuhn recolouring tool at preserving original website temperature and naturalness, and is not significantly different for weight and differentiability. SPRWeb also performs better on maintaining activity for website wireframes;

3. Traditional models of colour subjective response do not adequately predict the perceived Activity of full websites.

In the following sections, we explain how SPRWeb achieved its performance gains and discuss how our results will generalize to broader contexts.

Explanation of Results

Temperature: SPRWeb was substantially better at preserving the perceived temperature of the original websites for both wireframes and realistic websites. Examination of some of the websites used in the studies shows that in many cases, Kuhn's recolouring tool replaced reds with blues, and purples with yellows. Ou's model for temperature (Figure 2) shows that hue determines the sign of a temperature prediction (positive for 'warm' and negative for 'cool'). When Kuhn performs this switch, warm colours become cool, and cool colours become warm. As SPRWeb is designed to keep the original colour temperature as much as possible, such dramatic switches in colour temperature do not occur, and hence

SPRWeb does consistently better than Kuhn at maintaining the original colour temperature.

Activity: SPRWeb performed better than Kuhn's recolouring tool at preserving Activity in the wireframe study. Kuhn's recolouring tool strictly enforces luminance preservation, and as a result, this recolouring tool has fewer options to choose replacement colours from. When two original colours have similar luminance, Kuhn's recolouring tool will push the colours apart, making one more colourful and one less. Ou's model for Activity relies heavily on chroma (colourfulness) of the colour, so changes in chroma will result in reduced Activity preservation. For full websites, SPRWeb and Kuhn had no significant difference in performance, and both performed poorly when considering correlations between original website user responses and the recoloured website responses. One possible reason for this is that Ou's model for Activity is overly simplistic, and does not account for the effects of visual layout and complexity on perception of Activity. This hypothesis about activity is supported by one user who remarked, "...the relative position of the color is important in terms of web site design as well. For example the dark blue in the bottom of the page has a different feeling with the same color on the header section of the site."

Weight: SPRWeb performed worse than Kuhn's recolouring tool for website wireframes, but the incorporation of a luminance preservation component to SPRWeb's cost function allowed comparable performance with Kuhn's recolouring tool in the subsequent full website study for Weight. This reveals that in addition to difficulties in predicting Activity for complex displays, Ou's model perhaps does not sufficiently incorporate the effect of luminance on perception of weight. The addition of luminance preservation to SPRWeb did not overly restrict its recolouring algorithm, however, allowing enough flexibility to still maintain Temperature for full websites.

Naturalness and Differentiability: SPRWeb achieves significantly better naturalness and equivalent colour differentiability to Kuhn. This is in part because of the hue-flipping performed by Kuhn during recolouring, but is also likely due to Kuhn's overly restrictive luminance preservation. Kuhn restricts replacement colours to be isoluminant with the original colours to improve performance (by reducing the solution space), but this also overly restricts the recolouring process, eliminating potential replacement colours that offer greater preservation of naturalness and greater restoration of differentiability. SPRWeb is less restricted in its recolouring algorithm, thereby allowing more flexibility and hence greater opportunities to find an optimal set of replacement colours.

Generalizability

SPRWeb currently recolours CSS files that specify website text and gradient colours, but most websites also use images. In addition, many paintings and visualizations use colours to elicit particular subjective responses. We plan to extend SPRWeb to full image recolouring in the future. Existing image recolouring tools use an optimization step that follows the same API as that of SPRWeb (representative colours in, replacement colours out), easing this extension. Once developed, combining the CSS recolouring present in the cur-

rent SPRWeb version with this image recolouring extension will allow the development of a web browser plug-in that recolours an entire website. Challenges lie in improving the performance of this plug-in to run in near real time, as well as gathering information about the type of CVD the user has.

The somewhat surprising results for Activity for full websites reveals that existing models of subjective responses lack consideration of the effect of layout and visual complexity on perceived activity. SPRWeb succeeded in preserving temperature and weight, but these dimensions may also be influenced by layout, spatial frequency, or other factors. We plan to explore extending the existing models in light of these new findings.

Among the major challenges in attempting to preserve subjective responses are cultural and interpersonal variations [6]. It is not known what determines our subjective responses, and whether they are learned or innate. We do know that people's responses contain outliers: e.g., some participants reported blue as warm, not cool. In the future, we plan to incorporate personalized models of subjective colour responses into SPRWeb to address these variations.

Last, our formative study explored subjective responses to a restricted set of colours, the DTEP. We are very interested to see how people with CVD respond to colours not on the DTEP. We hypothesize that the subjective responses for people with CVD will match the subjective responses of someone without CVD for the simulated appearance of the colour under consideration. To achieve this, we are currently working on tools that allow personalized colour vision simulation [8].

CONCLUSIONS AND FUTURE WORK

Colour vision deficiencies change the user experience of the Web. To help equalize the user experience of a wider range of users, we developed the SPRWeb tool that automatically recolours websites to preserve subjective responses for individuals with red-green CVD while improving colour differentiability. Our evaluation showed that SPRWeb performs much better than the state-of-the-art Kuhn recolourer.

In the future, our main goals are to extend SPRWeb so that it can recolour website images and to deploy SPRWeb as a browser plug-in. To recolour images, SPRWeb will need to incorporate image processing to find representative colours of images. We plan to utilize a recently-developed clustering technique [9] to find the representative colours for each image. These representative colours will then be passed to SPRWeb's existing optimization step to produce replacement colours. Utilizing the clusters of colours produced by this clustering technique, the replacement colours will then be applied to the image to produce a recoloured version. By merging the current CSS SPRWeb version and this proposed image version, a browser plug-in applicable to a wide variety of web content will be possible. Deployment of SPRWeb as a plug-in will allow us to study it in a range of real world situations.

REFERENCES

1. Birch, J. *Diagnosis of Defective Colour Vision*, 2 ed. Oxford University Press, 2001.
2. Bonnardela, N., Piolata, A., and Bigot, L. L. The Impact of Colour on Website Appeal and Users' Cognitive Processes. *Displays 32*, 2 (2011), 69–80.
3. Brettel, H., Viénot, F., and Mollon, J. D. Computerized Simulation of Color Appearance for Dichromats. *JOSA A 14*, 10 (1997), 2647–2655.
4. Cole, B. L. The Handicap of Abnormal Colour Vision. *Clinical and Experimental Optometry 87*, 4-5 (2004), 258–275.
5. Coursaris, C. K., Sweirenga, S. J., and Watrall, E. An Empirical Investigation of Color Temperature and Gender Effects on Web Aesthetics. *Journal of Usability Studies 3*, 3 (2008), 103–117.
6. Cyr, D., Head, M., and Larios, H. Colour Appeal in Website Design Within and Across Cultures: A Multi-Method Evaluation. *IJHCI 68*, 1 (2010), 1–21.
7. Eisemann, L. *Pantone's Guide to Communicating with Color*. Grafix Press, Ltd., 2000.
8. Flatla, D. R., and Gutwin, C. "So That's What You See!" Building Understanding with Personalized Simulations of Colour Vision Deficiency. In *Proc. ASSETS '12* (2012), 167–174.
9. Frey, B. J., and Dueck, D. Clustering by Passing Messages Between Data Points. *Science 315*, 5814 (2007), 972–976.
10. Gorn, G. J., Chattopadhyay, A., Sengupta, J., and Tripathi, S. Waiting for the Web: How Screen Color Affects Time Perception. *Journal of Marketing Research 41*, 2 (2004), 215–225.
11. Gorn, G. J., Chattopadhyay, A., Yi, T., and Dahl, D. W. Effects of Color as an Executional Cue in Advertising: They're in the Shade. *Management Science 43*, 10 (1997), 1387–1400.
12. Greenfield, G. R., and House, D. H. Image Recoloring Induced by Palette Color Associations. In *Proc. WSCG* (2003), 189–196.
13. Hall, R., and Hanna, P. The Impact of Web Page Text-Background Colour Combinations on Readability, Retention, Aesthetics and Behavioural Intention. *Behaviour & Information Technology 23*, 3 (2004), 183–195.
14. Hardy, L. H., Rand, G., and Rittler, M. C. HRR Polychromatic Plates. *JOSA 44*, 7 (1954), 509–521.
15. Iaccarino, G., Malandrino, D., and Scarano, V. Personalizable Edge Services for Web Accessibility. In *Proc. W4A* (2006), 23–32.
16. Ichikawa, M., Tanaka, K., Kondo, S., Hiroshima, K., Ichikawa, K., Tanabe, S., and Fukami, K. Web-Page Color Modification for Barrier-Free Color Vision with Genetic Algorithm. In *Proc. GECCO* (2003), 2134–2146.
17. Jacobs, K. W., and Frank E. Hustmyer, J. Effects of Four Psychological Primary Colors on GSR, Heart Rate and Respiration Rate. *Perceptual and Motor Skills 38*, 3 (1974), 763–766.
18. Kuhn, G., Oliveira, M., and Fernandes, L. An Efficient Naturalness-Preserving Image-Recoloring Method for Dichromats. *IEEE Trans. Vis. & Comp. Graph. 14*, 6 (2008), 1747–1754.
19. Meyer, G. W., and Greenberg, D. P. Color-Defective Vision and Computer Graphics Displays. *IEEE CG&A 8*, 5 (1988), 28–40.
20. Nam, J., Ro, Y. M., Huh, Y., and Kim, M. Visual Content Adaptation According to User Perception Characteristics. *IEEE Trans. Mult. 7*, 3 (2005), 435–445.
21. Ou, L., Luo, M., Woodcock, A., and Wright, A. A Study of Colour Emotion and Colour Preference. Part I: Colour Emotions for Single Colors. *Color Research & Application 29*, 3 (2004), 232–240.
22. Reinhard, E., Adhikhmin, M., Gooch, B., and Shirley, P. Color Transfer Between Images. *IEEE Comp. Graph. & App. 21*, 5 (2001), 34–41.
23. Solli, M., and Lenz, R. Color Emotions for Image Classification and Retrieval. In *Proc. CGIV* (2008).
24. Stone, M. *A Field Guide to Digital Color*. AK Peters, 2003.
25. Valdez, P., and Mehrabian, A. Effects of Color on Emotions. *Journal of Experimental Psychology 123*, 4 (1994), 394–409.
26. Wandell, B. A. *Foundations of Vision*. Sinauer Associates, 1995.
27. Wang, B., Yu, Y., Wong, T., Chen, C., and Xu, Y. Data-Driven Image Color Theme Enhancement. *ACM Trans. Graph. 29*, 6 (2010), 146.
28. Yang, C., and Peng, L. Automatic Mood-Transferring Between Color Images. *IEEE Comp. Graph. & App. 28*, 2 (2008), 52–61.