

# Screenshot-based Color Compatibility Assessment and Transfer for Web Pages

Ou Wu, Mengqiao Han

Received: date / Accepted: date

**Abstract** Colors play particularly important roles in both designing and accessing Web pages. A well-designed color scheme improves the visual aesthetic of Web pages and facilitates user interactions. As far as we know, existing studies on color compatibility assessment and enhancement focus on images, and the assessment and enhancement for Web colors are rare. In order to aid Web designers evaluate and choose colors, this paper investigates color compatibility assessment for Web pages and applies this assessment to Web color editing based on Web screenshots rather than source codes. This study consists of four parts. First, the roles of color design in Web pages are discussed and a screenshot-based approach is proposed for the analysis of Web page regions. Second, a new method for extracting the color theme of a Web page is proposed. Then, we construct an assessment model that attributes scores to the color compatibility of Web pages through transfer learning. Third, we examine Web color transfer and combine it with the learned compatibility assessment model to create a new application that recommends colors for Web design. Lastly, the evaluation results suggest that the constructed compatibility assessment model and the proposed color transfer technique are effective and are superior over conventional methods. User studies suggest that our color recommendation application can generate new Web page screenshots that have higher color compatibility scores in comparison with those of the original pages.

**Keywords** Color compatibility assessment · Color theme extraction · Color transfer · Screenshot · Web mining · Image processing · Transfer learning

## 1 Introduction

Color information such as hue, brightness, and saturation has the potential to affect users' perceptions, physiological reactions, emotional reactions or behavioral intentions [41]. Colors are also crucial to the success of Web pages because they directly affect how users

---

Ou Wu is with Center for Applied Mathematics, Tianjin University. wuou@tju.edu.cn

Mengqiao Han is with NLPR, Institute of Automation, Chinese Academy of Sciences. michellehan33@126.com

---

Address(es) of author(s) should be given

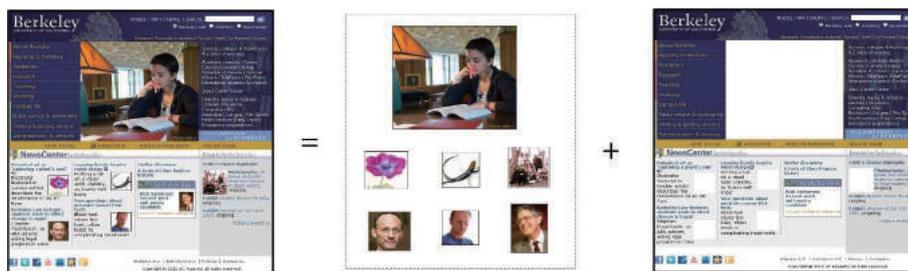
perceive the aesthetic and usability of Web pages [25,37]. Ineffective color combinations on graphic design impede user performance and satisfaction [18]. User studies also indicate that increased color appeal will result in greater trust and satisfaction on Web pages [5]. Given that color is particularly important for Web pages, numerous works have been conducted on Web page colors [17,40,4]. These existing studies either explore color design rules or consider colors as important aspects in the evaluation of Web design. To date, little research effort has been exerted toward the direct evaluation of Web colors, for example, the color compatibility of Web colors. Color compatibility refers to the compatibility degree of a set of colors. One of the most popular theories of color compatibility is the notion of hue templates, which generalizes Goethe's theory [11] by describing compatible colors as fixed rotations about the color wheel. An effective color compatibility assessment tool would be useful for many Web-related applications, especially those for choosing colors in Web design. Given that designers themselves do not always have impressions that are similar to users [31]), surveys of the perceptions of population samples are usually performed to evaluate Web design in terms of color choices. However, such an entire evaluation procedure is costly and time-consuming. Therefore, an objective third-party Web color compatibility evaluation tool that can produce reliable estimates of color compatibility of Web page designs can help designers obtain feedback at reduced costs and time.

O'Donovan et al. [28] recently initiated a pilot study on the construction of assessment models for the compatibility of color themes with the use of online datasets (from the Websites of Adobe Kuler and COLOURLovers). Such datasets consist of color themes and their associated compatibility ratings. These color themes were created by color experts, and the compatibility ratings were the results of votes cast by viewers (users). The constructed regression model proposed by O'Donovan et al. [28] can give scores to color themes, and their classification model predicts whether a color theme is compatible. They proposed several potential image editing applications, such as color theme optimization and color suggestion. Their experiments demonstrated the usefulness of the constructed color assessment modes in image editing.

The abovementioned work motivated us to investigate the assessment of Web pages in terms of color compatibility. Web page designers usually choose a combination of a small number of colors for a page. Therefore, an intuitive way to realize Web color compatibility assessment is to follow the approach proposed by O'Donovan et al. [28], that is, to obtain the color theme of a Web page<sup>1</sup> from its screenshot, and then assess its compatibility on the basis of the learned models by O'Donovan et al. [28]. However, we found this approach to be inappropriate for direct use with Web pages because Web page screenshots and images have obvious differences. As shown in Fig. 1, a Web page (screenshot) has several areas-denoted as the dynamic part-that display visual content, such as images, flash objects, videos, etc. The colors found in the dynamic part change with the content. The areas outside the dynamic part are called static. As the colors in the dynamic part change from time to time, the colors in the static part are the focus in Web design. The present study only assesses the colors of the static part. By contrast, image color assessment does not require a distinction between static and dynamic parts. Therefore, Web color compatibility assessment must be further studied. In this paper, regions of Web pages are analyzed, and a new construction approach for a Web color compatibility assessment model is proposed.

---

<sup>1</sup> In previous literature, color theme is a concept used in image/video analysis. Following the definition used in [24], the color theme of a Web page is defined as an ordered combination of 5 (or less) colors that can best represent the colors of that page.



**Fig. 1** A Web page (left), its dynamic part (middle), and its static part (right).

Aside from helping designers obtain feedback for colors, the color compatibility assessment for Web pages can be used together with automatic color editing techniques to help designers learn from well-designed Web pages. The Web provides an enormous repository of design knowledge: every Web page represents a concrete example of human creativity and aesthetics [19]. Existing well-designed Web pages, especially the top-ranked pages by Alexa, are valuable for inspiring designers in designing new pages. However, automatic Web color editing techniques are rare. Therefore, in this paper, we further investigate a new automatic Web color editing technique, namely, color transfer between Web page screenshots. Numerous computer graphic techniques are used for image design, such as color transfer between images. Similar to color transfer between images, given a source Web page and a reference Web page, Web color transfer<sup>2</sup> generates a new screenshot for the source Web page with new colors, such that the perception of the new colors in the source screenshot becomes similar to that of the reference page's screenshot. Intuitively, image color transfer technique can be used directly on screenshots. However, a Web screenshot cannot be simply considered an image. As discussed above, the colors of dynamic parts change from time to time, and colors in static parts are the focus of Web design. Therefore, only colors from static parts have to be considered in Web color transfer. Moreover, Web designers choose colors for Web pages to convey messages, invoke feelings, or accentuate areas of interest to potential users. The colors of different regions of a Web screenshot usually serve different functions. Web color transfer should consider the functions of colors. Consequently, image color transfer algorithms are inappropriate for direct use in Web screenshots. To our knowledge, automatic Web color transfer is a complex and unsolved problem. Thus, our work simplifies this problem by transferring colors of Web screenshots based on the region analysis, i.e., the static part is distinguished from the dynamic part for a Web screenshot.

In addition, Web color compatibility assessment and transfer are integrated as a new application for recommending colors to Web designers. This application can aid Web designers in obtaining color inspirations from existing professionally designed Web pages or other well-created visual arts, and may accelerate the whole design process. The new application may increase exchanges between Web mining and computer graphic communities so that more computer-aided Web design techniques could be developed.

The main contributions of our work are summarized as follows:

<sup>2</sup> Ideally, color transfer should ideally be investigated between Web pages. However, color transfer between pages requires the colors of each position of the Web pages and their corresponding positions in the source codes which is quite challenging. In addition, Web design has become increasingly graphical in nature [14]. Therefore, in terms of Web color design, a screenshot with new colors through color transfer can also inspire designers. In addition, the approach investigated for color transfer between Web page screenshots can also transfer the colors of images or videos into a Web page screenshot.

1. The design roles of colors in Web pages are discussed. Web pages are designed to show information and facilitate user interaction. Thus, the screenshot of a Web page cannot be seen as a common image. This work summarizes the main roles of colors in Web pages, which suggests the distinct differences between the colors of Web screenshots and images. Image color editing techniques (e.g., image color transfer) cannot be used directly on the screenshots of Web pages if the roles of colors in design are considered. A new screenshot-based method is proposed for locating the static part of a Web page, given that current source code-based analysis techniques are insufficient for appearance analysis for some Web pages when not all the visual information is contained in source codes.
2. The color compatibility assessment for Web pages is investigated. A new approach is proposed to train a model to assess Web color compatibility. The approach is new in the following aspects. Only colors of the static part of a page are considered rather than the whole page. An effective clustering algorithm for extracting the color themes of Web pages is utilized. Subsequently, a transfer learning is used to learn assessment models.
3. Web color transfer, a new transfer technique between Web screenshots, is present and studied. Given that color transfer for images helps image designers learn from existing images in choosing and editing colors, color transfer between Web screenshots is investigated. The aforementioned color compatibility assessment is then integrated with color transfer to create a new application that recommends colors to Web designers.

To the best of our knowledge, this is the first time that automatic color compatibility assessment and transfer have been investigated for Web pages. The remainder of this paper is organized as follows. Section 2 introduces related works. Section 3 discusses the important roles of colors in Web pages. Then, the problems concerned in this paper are presented. Section 4 describes the key components of our construction framework for a color compatibility assessment model. Section 5 introduces our color transfer technique and the color recommendation application. Section 6 presents some discussions. Section 7 concludes this paper.

## 2 Related work

### 2.1 Color compatibility assessment

One of the focuses of color assessment is evaluating the compatibility of color combinations. Existing studies on color compatibility assessment can be roughly divided into three categories.

1. Single factor-based methods. These studies are based on color compatibility tools, such as color wheels [11]. [11] pointed out that contrasting colors, i.e., those found on opposite sides of the color wheel, are compatible. Some researchers have compiled color templates based on the color wheel to help color designers [3].
2. Multiple factor-based methods. These studies [22,36] assess colors depending on hues and other factors, such as saturation, lightness, etc. Unlike the previous category, this type of work attempts to develop quantitative analysis based on controlled laboratory experiments.
3. Learning-based methods. Machine learning is a powerful technique used to construct classification or scoring models. As discussed earlier, [28] conducted a pioneering study that trains a classification/scoring model to quantitatively rate the compatibility of color themes using color theme-rating data obtained online.

Other studies utilize the abovementioned color compatibility assessment theories to guide color enhancement or aesthetic evaluation. A system that assesses overall colors currently does not exist.

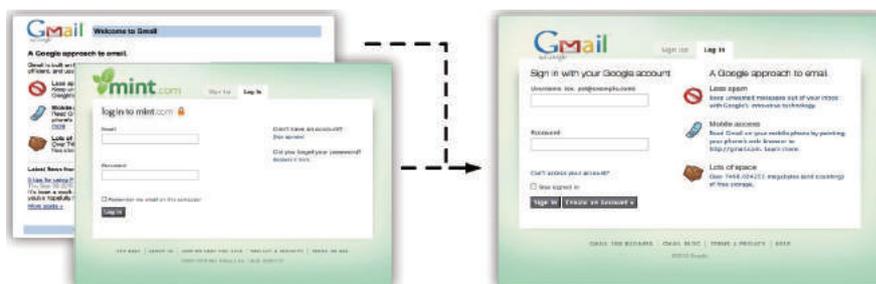
## 2.2 Color transfer

Color transfer is a classical image editing technique that retains the scene from a source image and applies the color style of a reference image to the source image [7,38,43,44]. Reinhard et al. [34] proposed the first color transfer algorithm based on the mean and standard deviation of color values from the source and reference images. Recently, researchers have developed numerous solutions that transfer the colors of pixels locally [32,39]. Color transfer has been successfully applied in image color editing and movie post-processing. Little headway has been made on color editing for Web pages, despite the fact that color is also one of the main concerns of Web design.

## 2.3 Web Quality Analysis

A large number of recent studies [15] have focused on analyzing the quality of Web pages. Ivory et al. [15] analyzed the metrics (e.g., link count, font count, and page size) of a large amount of Web pages and concluded that page-level metrics can accurately predict if a page will be highly rated. Reinecke et al. [33] developed computational models for measuring the perceived visual complexity and colorfulness of Web pages. Lafleur and Rummel [21] investigated the measurement of visual clutter of Web pages. Zhu and Gauch [45] incorporated six quality metrics (e.g., information-to-noise ratio) in Web information retrieval. In the current study, the regions and the colors of Web pages are analyzed. Therefore, the current study on region and color analysis also belongs to the area of Web quality analysis.

## 2.4 Automatic Web Color Edit



**Fig. 2** An example of data-driven Web design: automatically transfer the layout and style of the mint Web page to the Gmail page. This example is directly extracted from [20].

Kumar [19] proposed a classical data-driven Web design technique. Data-driven Web design can help Web designers obtain design patterns or inspirations from existing Web

pages. The Bricolage project [20] is the pioneering study on data-driven Web design. Bricolage can transfer the appearance style of one page to another as shown in Fig. 2. To our best, the Bricolage project does not introduce computer vision and computer graphics techniques, which have been widely used in the editing and enhancing the colors of images. [8] developed an interesting tool called SPRWeb to recolor Web pages that can preserve subjective responses and improve the color differentiability of Web pages. The subjective responses of the colors are measured by Ou et al. [30] emotion theory on color temperature and weight. SPRWeb can help users with color vision deficiencies view Web pages. This tool is based on a strict assumption that website images are not recolored and that the entire website color scheme is specified in its CSS file. However, this assumption usually does not hold because many Web pages' backgrounds consist of various images. Currently, colors in background images cannot be obtained by using CSS files. Thus, the strict assumption seriously weakens its applications for actual use.

Our proposed color transfer technique also offers a way to help Web color designers learn from existing Web pages. Nevertheless, our technique does not analyze the HTML source codes of Web pages and is based on applying computer graphic techniques to the screenshots of Web pages. Therefore, compared with the Bricolage project and SPRWeb which can be seen as source code-based technique, this study adopts a screenshot-based approach.

The source code-based approach is efficient and can obtain exact color mapping results when the structures of two pages are quite similar. Nevertheless, this approach suffers from the following limitations:

1. The analysis of HTML source codes has two main limitations [26]. First, HTML itself is still evolving (from Version 2.0 to Version 5.0) and when new versions are introduced, existing works will have to be amended repeatedly to adapt to new versions. Second, the complexity exhibited by HTML source codes of Web pages is ever-increasing. The underlying structures of current Web pages are more complicated than ever, making it more difficult for existing solutions to analyze the structures of their appearance. Moreover, designers usually use images to decorate Web pages and the visual presentation from these images cannot be obtained based on source code analysis.
2. The source code-based approach fails if some colors are not described in the HTML source codes.
3. The source code-based approach can not transfer colors for other forms of visual objects (e.g., images) into a Web page.

On the contrary, our proposed screenshot-based approach can alleviate these two limitations. Fig. 3 shows an example for the transfer of the colors of an image (i.e., Fig. 3(b)) into a Web page (i.e., Fig. 3(a)). The transferred result is shown in Fig. 3(c). In addition, our proposed approach provides an easy way of investigating intelligent Web design techniques. Various existing image color editing techniques (e.g., color harmonization [3] and data-driven color theme enhancement [42]) may be adapted to facilitate Web design based on a similar screenshot-based approach. Nevertheless, the screenshot-based approach is higher time-consuming than the HTML-based approach and may fail if the distributions of the reference colors and target colors are quite different. Therefore, in our point of view, an effective Web color transfer approach should integrate these two approaches into a whole.

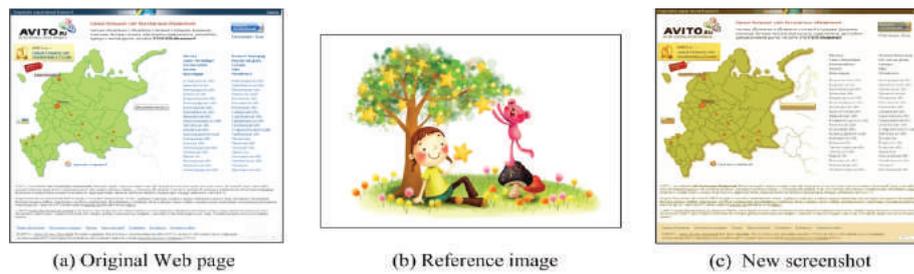


Fig. 3 Color transfer from an image to a Web page.

### 3 Web color analysis and problems

#### 3.1 Web color analysis

The colors investigated in our work are Web colors used in designing Web pages. Web colors are not randomly chosen and are usually not used solely for aesthetics, which is the main concern for image colors. A large number of design principles are used when choosing Web colors. According to these design principles, some important roles of Web colors are listed as follows:

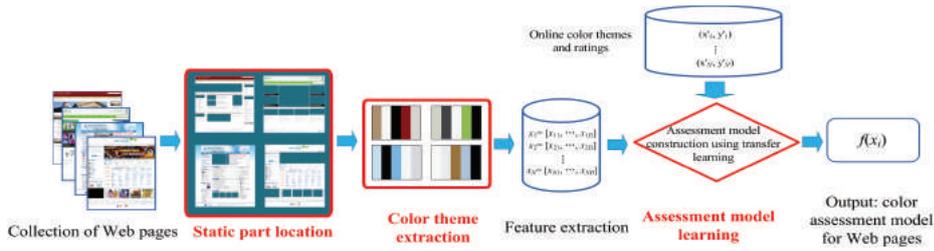
1. Facilitating user accessing and interaction. Web pages are generated to convey information to users. Thus, Web pages should have good readability. Background and font colors are usually chosen to be in contrast so that the text is readable. Colors are also used to help users know the functions of Web page elements (e.g., icons, buttons, and boxes) [23]. For example, elements with similar interaction purposes (e.g., navigation, input, news) are usually with the same colors in a Web page.
2. Describing and deliver important information. Various colors are used for texts and background images in Web pages. Differences in color often indicate the importance of pieces of information. For example, the colors of linked texts differ from the colors of adjacent texts. For Chinese Web pages, important texts are usually colored red.
3. Emotional appealing. Colors affect user emotions. Well-chosen colors in Web pages evoke positive emotions and make a user feel welcome, comfortable, relaxed, and secure<sup>3</sup>. According to color emotion theory [29], some colors are warm and portray feelings of energy, excitement and perkiness; some other colors are cool and make us feel calm, and possibly melancholic.

These three roles guide the color selection of Web designers. Naturally, Web colors and image colors have essential differences. The former usually has specific Web page design purposes. The roles of colors in Web pages should not be neglected when they are edited and enhanced.

#### 3.2 Problems

The aim of this study is to assess the color compatibility of Web pages based on their screenshots and transfer colors between Web page screenshots. According to the above analysis,

<sup>3</sup> <http://www.webdesign.org/web-design-basics/color-theory/the-importance-of-color-in-web-designing.12437.html>



**Fig. 4** The steps for constructing a color compatibility assessment model for Web pages.

this work focuses on designed colors which play important roles in Web page access and interaction. Compared with the corresponding image processing techniques, screenshot-based color compatibility assessment and transfer for Web pages have the following challenges:

1. Not all the regions of a Web page must be considered during assessment and transfer. A Web page usually contains several informative images that convey content to users. The colors of regions that contain informative images are not designed by designers and vary when the images change. Thus, the colors of these regions are not explored in this work. They should not be modified when assessing, editing, or enhancing the colors of a Web page and its screenshot.
2. The design purposes of colors should be considered. As stated above, colors in Web pages server different functions. Some colors are used to facilitate reading and interaction; some colors are used to attract potential users to buy products; while some others are used to convey the importance of standards or cultural meanings. Therefore, when editing or enhancing the colors in a Web page screenshot, the new colors should increase or at least not significantly weaken the functions of the original Web colors.
3. The amount of training data is insufficient. From the perspective of machine learning, constructing an effective automatic assessment model requires a large number of training Web pages. However, manually rating a large number of Web pages is a tedious job that is both time-consuming and error-prone. Therefore, this work should deal with learning using a limited number of rated Web pages.
4. The new colors generated should ideally be automatically transformed into the designed Web pages. The ultimate goal of Web color transfer is to generate new colors for a Web page instead of its screenshot. Therefore, once new colors are obtained, they should ideally be automatically transformed into the Web page being designed to produce a Web page with the new colors. Automatic transformation requires the colors of each position of the Web page and their corresponding positions in the source codes. Unfortunately, to our knowledge, no existing methods are capable of automatically transforming new colors into pages with many decorative background images.

Solving all the aforementioned challenges in one academic paper is unrealistic. This paper focuses on the first and third challenges. Therefore, the problems investigated in this paper are as follows. (1) Dividing a Web page into regions with static colors and regions with dynamic colors. (2) Extracting a color theme to represent the design colors of a Web page. (3) Constructing an assessment model for the color compatibility of Web pages. (4) Transferring colors between the screenshots of Web pages without considering the automatic transformation of new colors into Web pages. The following sections introduce our methodologies for solving these problems.

## 4 Web color compatibility assessment

### 4.1 An overall of the proposed approach

Our proposed approach falls under a conventional machine learning approach. It involves extracting features and training the assessment model (a regression function). The proposed approach is summarized in Fig. 4. Three main challenges (the red boxes in Fig. 4) are involved in constructing the model.

- Static part location. This step extracts the design areas of Web colors. Two screenshot-based methods are proposed by using computer vision techniques.
- Color theme extraction. Color theme extraction aims to represent the major colors in the static part of a Web page. Therefore, minor colors should be excluded.
- Assessment model learning. The training data should consist of the color themes of Web pages and their respective compatibility rating scores. Unfortunately, little such training data exist. The online color theme-rating data used by O’Donovan et al. [28] can be utilized as the training data. However, according to standard machine learning theory, the online data set is inappropriate for use because online color themes and Web color themes have different distributions.

To address the above challenges, a series of new methods are proposed. Processing source codes for locating the static part of a Web page can be avoided using a screenshot-based method. An effective clustering algorithm is introduced for color theme extraction to discover major colors as well as to exclude color outliers. Transfer learning is introduced into model learning to reweight the online training data and adapt the data distribution to that of Web color themes.

### 4.2 Static part location

There are two information sources available for the analysis of the visual presentation for a Web page. One is the source code and the other is the screenshot. As aforementioned analysis, source code-based approach has limitations. This study does not utilize source code-based analysis algorithms. Instead, we apply computer vision techniques to locate the static parts, which use the screenshots of Web pages directly.

Our proposed method is a block-by-block comparison method, which is based on blocks rather than pixels. If a Web page screenshot is divided into  $N * N$  blocks, we observed that the overall color differences between blocks in the same positions are far smaller than those between pixels in the same positions even when the sizes of dynamic parts vary. The output of this method is a set of image blocks from a sequence of temporal screenshots giving a URL<sup>4</sup>. Three steps are involved. (1) Each of the temporal screenshots is divided into  $N_1 * N_2$  blocks<sup>5</sup>. Assuming that  $I$  temporal screenshots are available, then, for each block position  $I$  image blocks are obtained after division. (2) We calculate the similarity between a block from the first screenshot and the corresponding blocks of the successive temporal screenshots. As a result,  $I-1$  similarities are obtained for each block. These similarities are then

<sup>4</sup> Temporal Web pages of a given URL can be obtained from <http://web.archive.org/> automatically.

<sup>5</sup> There have been several Web page block division algorithms. These algorithms are based on analyzing source codes. As this study does not intend to analyze HTML source codes, these existing algorithms are not used in this work.

averaged and normalized into  $[0, 1]$ . The similarity between two image blocks is calculated based on the earth mover distance (EMD) [35] between the color histograms. Assuming that the EMD is  $d$ , then the similarity is defined as  $\exp(-d)$ . (3) The blocks are sampled according to their average similarities and then used to construct the set of image blocks. The sampling strategy ensures that colors in the static parts are sampled with higher probabilities than those in the dynamic parts.

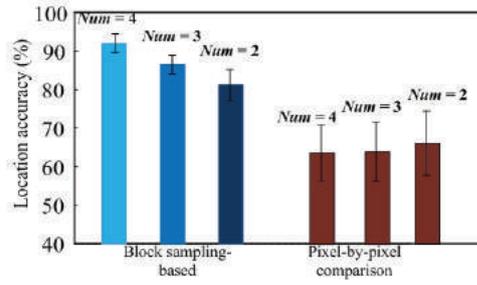
The results of this method on an exemplar temporal screenshot are shown in Fig. 5. The left image in Fig. 5 shows the static part (not including the black areas) and the right one shows the sampled blocks (not including the black areas). The blocks in the static part are sampled more than those in the dynamic parts. The sampled blocks do not affect the major colors of the static part, although some blocks in the static part are also not sampled as well. Most dynamic parts are not sampled (black areas).



**Fig. 5** The benchmark static part of a Web page (left) that was manually determined, and the sampled blocks of the Web page (right) obtained using our method. The black boxes were not included.

To conduct a quantitative evaluation, **150** homepages, mainly from companies, universities, governments, and personal sites, are collected. The URLs are introduced in Appendix B. To generate the ground truth, the static parts of the 150 pages are manually extracted. The manually extraction procedure is as follows. The three authors participated in the manual extraction. All are Chinese, male, and in the age of  $[30, 35]$ . Each participant manually extracted the static part for each page and each page is then associated with three static parts. The intersection of the three extracted parts is used as the ground-truth static part of a page. The screenshots in Figs. 5 (left) and 6 in the succeeding subsection are the manually extracted results. The location accuracy of a method on a page is defined as the proportion of correctly located pixels by the method. It is adopted as the evaluation criterion. Two methods are compared, namely, the block sampling-based, and a baseline method based on pixel-by-pixel comparison. For the pixel-by-pixel comparison method, pixels of the transformed screenshots of the same page collected over time with relatively slight variation in color are considered the static part. In the block sampling-based method, the number of block sampling times for a page is set to 2000, and both  $N_1$  and  $N_2$  are set to 40 based on a small validation set<sup>6</sup>.

<sup>6</sup> In the experiments, because it is difficult to conduct a theoretical analysis for the optimal values of the involved parameters, the parameters were chosen based on grid search on a small validation set. The block sampling times was chosen from the candidate set  $\{500, 1500, 2000, 2500, 3000\}$ . For  $N_1$  and  $N_2$ , a set of candidate values  $\{(20, 20), (25, 25), (30, 30), (35, 35), (40, 40), (45, 45), (50, 50)\}$  was tested. When



**Fig. 6** Location accuracies of the three competing methods for Web static part location.

Considering that the block sampling-based method adopts a sampling strategy, different runs of this method yield slightly different results. Therefore, this method is repeated ten times for each page and the average location accuracy is recorded. The average location accuracies over all the 50 pages achieved by the two competing methods are shown in Fig. 6. The results of both methods are presented when 2, 3, and 4 temporal screenshots are used. In Fig. 6, the three average location accuracies by the block-sampling method are higher than those of the competing method. Based on the  $t$ -test, the location accuracy achieved by the block sampling-based method is significantly larger than those of the baseline method ( $p < 0.01$ ).

The pixel-by-pixel comparison method fails because the dynamic part varies not only in its colors but also its size. Even if the size of the dynamic part of a Web page is changed only slightly, the performance of the pixel-by-pixel comparison will be decreased largely. Figure 7 gives an example of pixel-by-pixel comparison. The pixels whose colors change significantly are colored blue in Fig. 7 (c). Most pixels belonging to the static part are as also colored green in Fig. 7 (c). The reason is that the two images (for fox and people in (a) and (b)) have slightly different sizes, leading that the regions under the two images change slightly. Only a slight change will lead to the colors of the same positions of the two pages change significantly.



**Fig. 7** (a) and (b) are two pages with the same URL; (c) shows the pixels (colored by green) with different colors in (a) and (b).

the number of the sampling times was set as 2000, and the values of  $(N_1, N_2)$  were set as  $\{40, 40\}$ , the performance was the best.

### 4.3 Color theme extraction

In previous literature, color theme is a concept used in image/video analysis. Prior to presenting our proposed methodologies, the image color theme extraction is briefly reviewed.

#### 4.3.1 Color theme extraction for images

Two classical color theme extraction methods have been proposed in recent literatures. One is the aesthetics-aware method proposed by O’Donovan et al. [28], and the other is the user perception-aware method proposed by Lin and Hanrahan [24]. The goal of the aesthetics-aware method is to extract a representative color theme with a relatively high rate for an image. The color theme extracted through the method used by O’Donovan et al. [28] may not be the most representative color theme for an image. In contrast, our goal is restricted to the extraction of a color theme that can represent the colors of a Web page as much as possible, whether or not its rating is high. The goal of the user perception-aware method is to extract the color theme of an image that is as close to user perception as possible based on salient region detection for images. The visual attention of Web pages usually conforms to certain patterns, such as “F-shaped pattern” [27], across different pages. Therefore, salient region detection for Web pages is different from that of images. As a result, directly using the learned model based on salient region detection for images and image segmentation is inappropriate for Web page color theme extraction.

The two existing extraction methods described above share two similarities: (1) all colors in a color theme are assumed to be independent of contents/objects in an image; (2) the number of theme colors does not exceed five. Five colors are considered sufficiently to represent the colors of an image. Therefore, these two similarities are followed and the specific color roles in the static parts of Web pages are not further used in our work.

Following the two existing methods, this work also uses a clustering-based approach to extract the color themes of Web pages. A Web page usually contains different colors. Some colors (e.g., range red colors in the page in Fig. 4 (left)) have a very small proportion. Thus, the obtained five colors may drift much from the original colors of the page. In existing studies on color theme extraction, colors with a very small proportion are usually ignored. Ideally, colors with a very small proportion should be considered as color outliers in the color clustering. Nevertheless, conventional clustering techniques such as K-means are sensitive to outliers. To deal with this problem, we use the outlier-aware clustering algorithm proposed by Forero et al. [10].

#### 4.3.2 Outlier-aware clustering

[10] proposed the following clustering model which explicitly accounts for outliers:

$$\min_{M,O,U} \sum_{i=1}^n \sum_{k=1}^K u_{ik} \|x_i - m_k - o_i\|_2^2 + \lambda \sum_{i=1}^n \|o_i\|_2 \quad (1)$$

where  $x_i$  is a data point;  $u_{ik} = 1$  if  $x_i$  belongs to  $k$ -th cluster and  $u_{ik} = 0$  otherwise;  $m_k$  represents a centroid;  $\lambda \geq 0$  is an outlier-controlling parameter, such that the higher the value of  $\lambda$ , the less the number of the points detected as outliers; the vector  $o_i$  is deterministically nonzero if  $x_i$  corresponds to an outlier, and  $\mathbf{0}$  otherwise. When  $\lambda \rightarrow \infty$ , all the data are deemed outlier-free and the outlier-aware clustering equals to K-means. The value

of  $K$  represents the number of colors considered in a color theme. Following the settings in previous work [24],  $K$  equals five<sup>7</sup>.

#### 4.3.3 Color theme extraction with static part location

In this stage, five representative colors are obtained by applying outlier-aware clustering to the colors of pixels produced by the static part location of a Web screenshot. The proposed block sampling-based method may sample pixels in some blocks more than once. Therefore, a weighted outlier-aware clustering algorithm is required. Equation (1) is then transformed:

$$\min_{M,O,U} \sum_{i=1}^n \sum_{k=1}^5 w_i u_{ik} \|x_i - m_k - o_i\|_2^2 + \lambda \sum_{i=1}^n \|o_i\|_2 \quad (2)$$

where  $w_i$ <sup>8</sup> is the weight of the  $i$ -th pixel. The solution of Eq. (2) is similar to that of Eq. (1) with a trivial modification.

The five representative colors obtained should be combined from left to right to form a color theme. The possible left-to-right combinations of the five colors are  $5! = 120$ . Given that the differences between two adjacent colors in a color theme affect the rating of the color theme [28], randomly setting the relative spatial positions of the five representative colors is inappropriate. A simple method is proposed to capture the partial spatial information of the five representative colors. Once color clustering is completed, the pairwise physical distances among clusters are calculated. The pairwise physical distance between two clusters is the average of pairwise position distances of two pixels in the two clusters. The pairwise position distance of two pixels is the pairwise Euclidean distances of the positions of the pixels in the Web page. The pairwise physical distances among clusters indicates the spatial information of the five colors. For each of the 120 candidate color themes, a set of pairwise position distances can also be calculated. For example, given a color theme  $(c1, c2, c3, c4, c5)$ , then  $d(c1, c4) = 3$  and  $d(c1, c5) = 4$ . Finally, an optimal color combination from the 120 possible combinations is selected, such that the orders among its colors according its pairwise position distances are consistent with the orders from the pairwise physical distances for the clusters as much as possible.

Figure 8 shows three exemplar results of the extracted color themes based on different methods: K-means (in (a)), the outlier-aware method with image synthesized-based methods (in (b)), and the outlier-aware method with block sampling-based location (in (c)). The color themes from outlier-aware clustering + block sampling-based location are more insensitive to the colors of the dynamic part than the other three methods. For instance, in Fig. 8(2), only the color theme in (c) does not contain purple which does not appear in the static part. A comparison between the proposed method and the latest state-of-the-art image color theme extraction method can be found in the supplementary file of this submission.

#### 4.3.4 Evaluation

A number of **150** homepages were also tested and the URLs are present in the supplementary file of this submission. Six color theme extraction methods shown in Table 1 are tested on

<sup>7</sup> The choice of this number follows the choices in previous studies on image color theme extraction. In addition, a larger number requires more features which may lead to less accurate learning performances because of curse of dimensionality.

<sup>8</sup> When the block sampling-based method is used, the value of  $w_i$  approximately proportions to the similarity of the block the  $i$ -th pixel locates.



**Fig. 8** Examples of color theme extraction for Web pages: in each example, (a) K-means, (b) outlier-aware clustering + image synthesized-based static part location, and (c) outlier-aware clustering + block sampling-based static part location.

the 150 Web pages. The image synthesized-based location generates a new screenshot by synthesizing a series of screenshots of a URL. This method is just for better evaluation the proposed method. For both the K-means and outlier-aware clustering algorithms, the number of clusters (i.e., parameter  $K$ ) is set as 5, and the initial centers are selected by the furthest-first initialization algorithm [12]. The parameter  $\lambda$  in outlier-aware clustering is set as 70 based on the search of a candidate set  $\{0.1, 1, 10, 20, 40, 70, 100\}$  on a small validation set.

**Table 1** The six competing methods for color theme extraction.

	K-means	outlier-aware clustering
without static part location	K-means without static part location	outlier-aware clustering without static part location
image synthesized-based location	K-means + image synthesized-based location	outlier-aware clustering + image synthesized-based location
block sampling-based location	K-means + block sampling-based location	outlier-aware clustering + block sampling-based location

As there is no ground truth, three graduate students, specifically two males and one female, were recruited using Email advertising from the experimental laboratory. All the participants are Chinese and age in [22, 27]. Each participant performed a simple online color blindness test<sup>9</sup> to ensure that the participant has normal color perception. The participants were invited to rank the performances of the six extraction methods on each page according to how well their extracted color themes represent the colors of the page. The more the consistency, the higher the rank. During a user-raking session, a Web page screenshot with its associated color themes obtained by the six extraction methods were randomly loaded. The participant was not aware of the correspondence between the color themes and the methods. The participant then viewed the screenshot and subsequently ranked the color themes. After ranking for a screenshot, the participant clicked “next” to load the succeeding random one. If the participant failed to rate a screenshot within a fixed amount of time, a screenshot was randomly selected among the un-rated screenshots and was loaded automatically. After all the screenshots’ themes were ranked, the participant’s rating task was concluded. Each

<sup>9</sup> <http://www.colormax.org/colorblind-test.htm>

participant was required to finish the ranking within three hours. Each participant was rewarded with 75RMB after finishing the ranking task. When all the participants finished their tasks, the average ranks of the performance orderings of the color theme extraction methods were calculated. The average ranks of the eight extraction methods are shown in Table 2. The ranking differences among the three participants are measured using the normalized Kendall tau distance [16]. The normalized Kendall tau distance ranges in  $[0, 1]$  and a lower value indicates a low difference (and a larger agreement). The normalized Kendall tau distances among the three participants are 0.1391, 0.1450, and 0.1604. All the disagreements among participants are less than 20%, indicating that the user disagreement among the three participants is insignificant.

**Table 2** The average performance Order of the six color theme extraction methods.

Color theme extraction method	Order
Outlier-aware clustering + block sampling-based location	1
K-means + block sampling-based location	2
Outlier-aware clustering + image synthesized-based location	3
K-means + image synthesized-based location	4
Outlier-aware clustering without static part location	5
K-means without static part location	6

To check the hypothesis that the proposed method of outlier-aware clustering + block sampling-based location is better than the other methods, the ANOVA-test is performed based on the participants' average ranks for all the six methods over the 150 pages. First, the methods with static part location significantly outperform the methods without ( $p < 0.05$ ). Second, the block sampling-based location significantly outperforms the other competing location methods ( $p < 0.05$ ). Therefore, the hypothesis is supported. The rest evaluations merely used the block sampling-based method for the static part location.

To further explore the performance of the outlier-aware clustering in color theme extraction, we compared the outlier-aware clustering with the latest state-of-the-art color theme extraction method for images proposed by Lin and Hanrahan [24]. The outlier-aware clustering method outperforms Lin and Hanrahan method on most experimental pages.

#### 4.4 Assessment model learning

The first two parts introduces the features and the learning algorithm. At last, the evaluation results are presented.

##### 4.4.1 Features

The features proposed in [28] are leveraged in this work with a slight modification. The mean values are weighted by the proportions of pixels in each of the five colors in the color theme obtained by the clustering algorithm. Finally, a feature vector with 334 dimensions is constructed to represent the color theme of a Web page. This vector contains comprehensive factors related to color compatibility, which can be divided three types. The first types of features are calculated in each space directly: the five colors themselves, colors sorted by lightness, differences between adjacent colors, sorted color differences, mean, standard

deviation, median, max, min, and max minus min across a single channel. Differences in hue are computed with wraparound. The second types of features are calculated based on the color histograms from the Kuler training set to produce scores for individual colors and pairs of colors. The third types of features are calculated for the distribution of hues from saturated colors and light colors in a color theme. Among the 334 dimensional features, numerous features are correlated. For example, the five colors and the colors sorted by lightness are nearly the same features. Consequently, the feature reduction technique, namely, principal component analysis (PCA) [2], is used in the experiment.

#### 4.4.2 Learning problem description

Unlike the work of [28] whose aim is to construct an assessment model for color themes created by color experts, our goal is to construct a model to assess Web page colors. On one hand, the input training data should consist of Web color themes and their associated ratings. Unfortunately, there are no sufficient rated data for Web page colors at present. On the other hand, learning a color assessment model for Web pages that directly uses rated color themes online is inappropriate.

Huang et al. [13] proposed an effective transfer learning algorithm that reweights the data in the source domain, in order to adapt the distribution of the re-weighted data to approach to the distribution of the target domain. Let  $\beta = \{\beta_i\} (i = 1, \dots, N')$  which denotes the weights of the samples in the source domain. It can be obtained by solving a quadratic problem.

After determining  $\beta$ , a color compatibility assessment function can then be obtained using a regression algorithm. This study uses the weighted LASSO regression method because of its good performance reported in [28].

#### 4.4.3 Evaluation

The evaluation of the utilized model learning methodologies is based on the following hypotheses:

- **H1**: When using the proposed transfer learning, the model based on more target source training data outperforms the model with less target source training data.
- **H2**: The feature de-correlation technique, PCA, improves the performance of a color compatibility assessment model.

**Experimental data** When the transfer learning is used, the training corpus should consist of two data sets, namely, source training set and target training set. Accordingly, the source training set is constructed as follows. All the three online color theme features and mean rating sets (Kuler, COLOURLovers, and Mturk) used in [28] are compiled to construct three candidate source training sets. Each of the three online sets is large, so when each one is used, 3000 samples are randomly selected from the online set to create its corresponding candidate source training set. Finally, three candidate source training sets are obtained.

To construct the target training set, top-3000 Web pages are collected based on the Alexa rankings [1] by deleting pages that are duplicated with nearly similar URLs (e.g., www.google.com and www.google.gr). To investigate whether more target data can improve the performance, an increasing number of training pages ( $500 * n, n = 1, \dots, 6$ ) by interval (the interval is  $3000 / (500 * n) - 1$ ) sampling from the top-3000 pages are used. Then the features of the sampled pages' color themes are extracted. Finally, six candidate target

training sets are obtained which are denoted as #1, #2, #3, #4, #5, and #6, respectively. In addition, we use ‘#0’ to denote a null target training data set in which the number of training page is zero.

For the test corpus, homepages are selected as our test data, and **500** homepages (including the pages used in Sections 4.2 and 4.3), mainly from companies, universities, governments, and personal sites, are collected. The URLs of all the experimental Web pages are available in an online supplementary file of this manuscript. A total of nine graduate students, specifically six males and three females, were invited to label the collected pages using Email advertising from our experimental laboratory. All the participants are Chinese and in the age of [22, 28]. Considering that the concept of color compatibility is not difficult to understand and we wanted the participants to access the pages freely, we did not give special instructions to the participants but only a simple training on how to use the labeling platform. Each participant was allowed to view one page within five seconds and assessed the color compatibility of the page from the five rating scores (1, 2, 3, 4, and 5). Here, “1” means very bad, and “5” means very good. The rating procedure for each participant was similar to those in user ranking in Section 4.3.4. The only difference is that for a given page in the user-rating procedure, an involved participant was required to rate the compatibility of the page’s colors instead of ranking it. Each participant was required to finish the rating within one and a half hours. Each participant was rewarded with 100RMB. After user rating, each page obtains nine scores. The average of these scores is adopted as the desired color compatibility score (or ground-truth score) of the corresponding test page. The color themes are extracted, and the features and the desired scores are the test corpus.

**Method comparison** The parameters of the involved methods in training are searched via cross validation. The performance of an assessment model is measured as follows. At first, the model takes turns run on each sample in the test corpus and produces a predicted score. The residual sum of square errors (RSSE) is then calculated based on the predicted and ground-truth scores. The RSSE value indicates the performance of a model. A lower RSSE value achieved by an assessment model means that the model can more accurately score the color compatibility of a Web page’s screenshot.

**Table 3** The RSSE values of the competing models (We use the used data sets to represent the corresponding models).

	#0	#1	#2	#3	#4	#5	#6
Kuler	1.03	0.94	0.90	0.79	0.75	0.67	0.66
COLOURLovers	1.12	0.99	0.91	0.87	0.81	0.74	0.72
MTurk	1.15	1.05	0.97	0.95	0.88	0.80	0.77

**Results** Table 3 shows the regression errors achieved by the competing models on the 500 test Web pages in terms of RSSE using different candidate source (i.e., Kuler, COLOURLovers, and MTurk) and target (#0-6) training data. The results in Table 3 are used to test the hypothesis H1. In Table 3, when the number of target data increases, the RSSE values (regression errors) decrease. To check whether or not the decrements are significant, the *t*-test was performed. For all the candidate source training data, the decrements of the RSSE errors are significant ( $p < 0.05$ ). Therefore, H1 is supported and this result is reasonable because the increasing number of target data causes the distribution of target data to more affect the



**Fig. 9** The average user ratings and scores predicted by models constructed by two models constructed from the #0 and #6 candidate target training sets respectively.

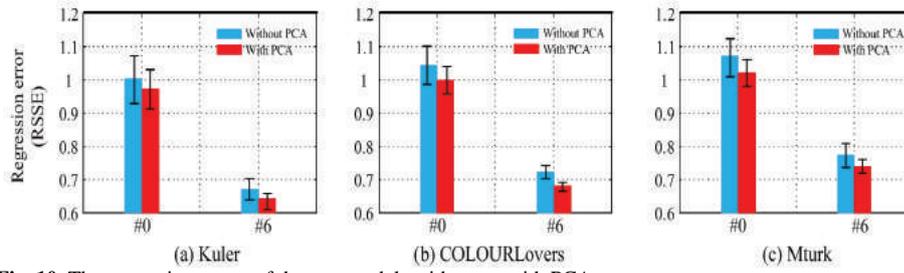
weights of the source data. In addition, the model obtained from the #0 candidate set is directly learnt from the online color theme-rating data. Therefore, the results in Table 3 also verify that ignoring the differences between the distributions of Web colors and online color themes is inappropriate.

Sample results are shown in Fig. 9. For each page in Fig. 9, its extracted color theme, its ground-truth score by users, and the scores obtained from models by using #0 and #6 candidate sets when MTurk data is used.

As previously presented in Section 4, some features are correlated. Thus, principal component analysis (PCA) [2] is further applied on the involved data sets to de-correlate the features. Figure 10 shows the results of the performances of the competing models when PCA is used using the three different candidate source training data sets. For the target training data, the #0 and #6 candidate sets are used. The average RSSE values of all the competing models decrease on all the three images. The  $t$ -test was performed for each pair of methods with or without PCA. In both Figs. 10 (b) and (c), the differences are significant ( $p < 0.05$ ). Therefore, H2 is partially supported and feature de-correlation can improve or at least not weaken the assessment performance.

## 5 Web color transfer and its application

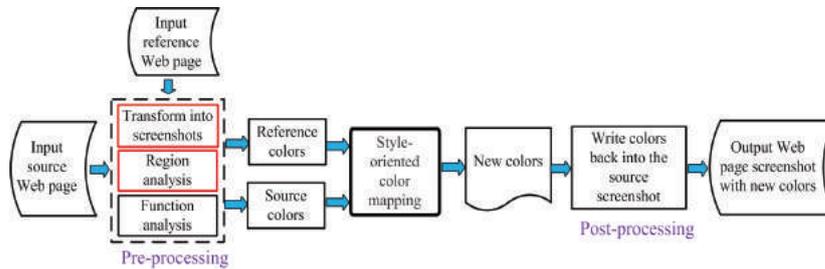
This section first introduces color transfer between Web screenshots. It is then used together with a color compatibility assessment model to automatically generate a number of new Web screenshots which can facilitate Web designers learn from existing well-designed pages.



**Fig. 10** The regression errors of the two models without or with PCA.

### 5.1 Web color transfer

The main technical line of Web color transfer based on differences between images and Web screenshots is shown in Fig. 11. This figure shows an additional pre-processing component in Web color transfer compared with the technical line of image color transfer. The color mapping step still applies the style-oriented color mapping in existing image color transfer methods. The region and function analysis steps aim to discard out-of-design colors and infer the functions of colors. Considering that the specific function (e.g., facilitating reading and interaction) analysis for colors is quite complex, this study only investigates the region analysis step. Ideally, the post-processing step should transform the new colors into the source Web page. Nevertheless, the transformation is also quite challenging<sup>10</sup>. Thus, this step only produces a screenshot with new colors. In terms of facilitating color design, the produced screenshot can also provide design clues and inspirations because designers usually choose colors on a screenshot sketch in the initial stage of Web design.

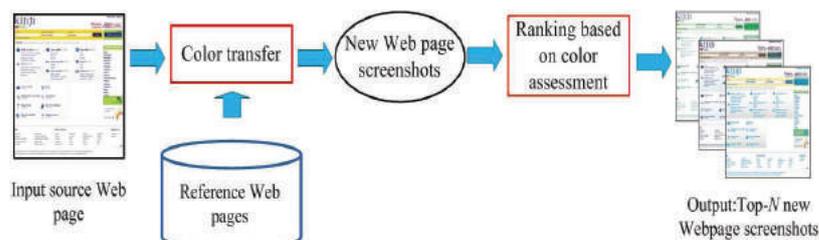


**Fig. 11** The main technical line of automatic Web color transfer.

### 5.2 Web color recommendation

Considering a typical application that a designer intends to change the colors of a Web page, if a technique can generate a number of new Web page screenshots based on the original page and a number of well-designed reference Web pages, and the new screenshots are ordered based on their estimated color compatibility scores, this technique may accelerate the

<sup>10</sup> The step of transforming new colors into a Web page relies on a powerful code analysis method that can infer the color of each position of a Web page. Unfortunately, to our knowledge, existing source code analysis techniques cannot implement these inferences.



**Fig. 12** The main steps of Web color recommendation.

design process and inspire the designer much. To this end, in this subsection, the above-mentioned Web color transfer approach is integrated with the aforementioned compatibility assessment into a new application, namely, color recommendation for Web design.

The main steps of color recommendation are presented in Fig. 12. The proposed approach automatically transfers the colors of the reference pages in the collections into an input Web page's screenshot. The top- $N$  transferred new Web page screenshots with higher color compatibility scores are then returned to users. The proposed application in this study includes a color compatibility assessment step, which can help users obtain high-quality transfer results. The dynamic and static parts of the source page's screenshot<sup>11</sup> are manually extracted by users to obtain more accurate results. This action is reasonable and acceptable considering that (1) only one source page (i.e., the input page) exists, and (2) when designing a page, the static part has usually been provided before the design for colors.

The three key components in the proposed application are detailed below.

- Constructing a reference Web page collection. This step collects Web pages with well-designed colors to be used in color transfer.
- Transferring colors based on the Web page database. This step takes turns to select one page in the Web page collection as the reference Web page. The static part of the reference page is extracted and color transfer is then performed.
- Ranking the transferred Web page screenshots based on color compatibility assessment. This step ranks the transferred Web page screenshots. A preferred screenshot list can be obtained according to the color compatibility scores of the transferred screenshots. Finally, top- $N$  screenshots are produced where  $N$  can be set by users.

The elements of the Web page collection can be any visual objects, such as images and paintings. The type of visual objects used in the collection depends on the application context or user choices.

### 5.3 Evaluation

This subsection compares our proposed Web color transfer method with the conventional image color transfer method for Web screenshots. Some cases studies are also provided to illuminate the potential value of the color recommendation application proposed in this study.

<sup>11</sup> The dynamic and static parts of the reference Web pages' screenshots are still extracted by the proposed algorithm. The number of reference pages can be large, whereas the number of source page is only one. Therefore, automatic extraction for the reference pages and manual extraction for the source page are reasonable. In fact, many computer graphic techniques are semi-automatic and require human interaction during running. If an accurate static part location algorithm is obtained in our future work, the new algorithm will be directly used in the approach.



**Fig. 13** The comparison of our color transfer method and the image color transfer method.

### 5.3.1 Evaluation for Web color transfer

Two color transfer methods are compared: (1) the image color transfer method proposed by Pitié et al. [32] and (2) the method employed in the present study that discards the dynamic part of the reference Web page. The style-oriented color mapping in the image color transfer method proposed by Pitié et al. [32] is used in our method. The image color transfer method differs from our method because it simply takes a Web page as an image and does not make any region or function analysis of the different areas of a page (screenshot).

Figure 13 shows five Web color transfer examples using the two competing methods. In all the five transfer examples, the colors of the new screenshots generated by the method employed in this study are more similar to the colors of the reference pages than those achieved by the image color transfer method. In the first, fourth, and fifth columns, the colors obtained by the Pitié et al. method are darker because the reference pages in these columns contain informative images with dark colors. The yellow color in the dynamic part of the second column is transferred into the downside of the static part. The colors in the dynamic part of the third column may also affect the transfer results.

The abovementioned comparison suggests that the colors of the dynamic part indeed affect the transfer results. Ignorance of the dynamic part colors results in that the transferred

colors are not perceived similar to those in the reference pages. The strategy that simply utilizes image color transfer techniques for Web color transfer is ineffective.

### 5.3.2 Evaluation for Web color recommendation

To verify the usefulness of our introduced Web color recommendation technique, a typical application context is considered: the designer of a Web page intends to improve the colors of the page. As described earlier, the reference Web pages in Web color transfer can be any forms of visual objects, including Web pages and images. Colors for ten Web pages from the collected Web pages and some cartoon images are transferred<sup>12</sup>. To assess the recommended Web screenshots, a user study was launched through an online user-study website in China [6]. In the present user study, in order to keep participants from the content, no Chinese Web pages were selected considering that all the participants were Chinese. In addition, participants were asked to focus on colors instead of content. Each participant was invited to rate the colors of 10 groups of Web page screenshots (including the screenshots in Figs. 14 to 19). In each group, four Web page screenshots that are made up of an original Web page and its three transferred result screenshots. Participants rated the color compatibility of each screenshot in the range of “1” to “5”. Here, “1” means very bad, and “5” means very good. A total of 110 participants (53 males and 57 females) rated the Web screenshots. Approximately 64.04% of the participants surf the Internet more than 16 hours a week. Approximately 12.36% of the participants surf the Internet between 8 and 16 hours a week and 23.6% surf the Internet lower than 8 hours. Approximately 5.46% are above the age of 40; 20.9% are in the age of 30 and 40; 52.74% are in the age of 20 and 20; 20.9% are below the age of 20.

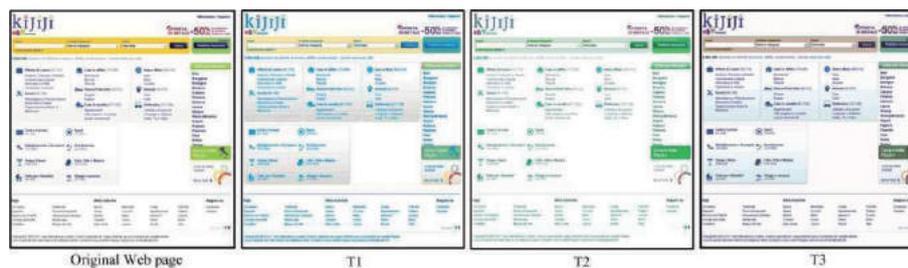
The results of the user study show that 56.67%<sup>13</sup> of the transferred Web screenshots have higher rating scores than the original pages. Figures 14 to 18 show five color transfer examples. The leftmost page in each example is the original page, whereas the remaining three screenshots (i.e., T1, T2, and T3) are the transferred results whose scores are ordered top 3 by employing the learned compatibility assessment model. Colors in some transferred screenshots have higher rating scores. For example, almost all the participants in the age of 15 and 20 provided the highest scores to the screenshot T1 in Fig. 16. Some transferred screenshots receive lower ratings. The primary reason is that the transfer colors in these screenshots have numerous artifacts. Fig. 19 shows a failed transfer example. All the transferred screenshots have lower rating scores than the original page. All the three transfer screenshots have color artifacts. The inter-rater reliability of the user ratings for all the groups is calculated using Fleiss’ kappa measure [9]. The average kappa value is 0.693, which means that the inter-rater reliability is relatively high. The screenshots with relatively low kappa values are those used cartoon images as reference. Cartoon images are usually with vivid colors that are usually appreciated by youngsters but may be unaccepted by old participants.

To further assess the usefulness of our color recommendation technique in Web color design (or modification), five Web designers including three male and two female are invited to evaluate the transferred Web screenshots. All the five designers are Adobe certified product experts and all of them have more than ten-year Web design experiences. All of

<sup>12</sup> Cartoon images are selected because they usually have vivid and brighter colors.

<sup>13</sup> The value 56.67% is based on the evaluation of the top-3 Web screenshots. If top-5 or top-10 Web screenshots are returned, the value will most likely increase. All the top- $N$  transferred screenshots are shown to designers. Whether or not they will be used depends on the choices of designers.

them are Chinese and in the age of 30 and 35. The ten original pages are used and each page is associated with a transferred screenshot which has the highest average rating score by users in the aforementioned online study. During the designer evaluation, the screenshot of the original page and its associated transferred screenshot were presented, each participant selected the screenshot (or screenshots) whose colors are more compatible. After the design evaluation procedure was finished, 50 judgments were obtained. Each participant was rewarded with 100RMB. In 74% of the judgments, the transferred screenshots judged to be with more compatible colors; in 10% of the judgments, the transferred screenshots are judged to be with identically compatible colors with the original ones. The results indicates that the color recommendation technique introduced in this study can offer designers new screenshots with better colors.



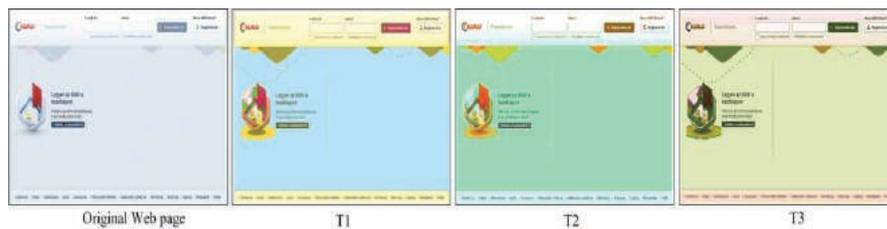
**Fig. 14** The original Web page (the leftmost) and three transferred screenshots. The color compatibility scores of the three transferred screenshots are ordered top 3. In our online user study, the average rating scores by 110 users of the two transferred screenshots (T2 and T3) are higher than that of the original page.



**Fig. 15** The original Web page (the leftmost). The original Web page (the leftmost). In our user study, the average rating score by 110 users of the second screenshot is higher than the original page.



**Fig. 16** The original Web page (the leftmost). In our online user study, the average rating scores by 110 users of the two transferred screenshots (T1 and T2) are higher than that of the original page.



**Fig. 17** The original Web page (the leftmost). In our online user study, the average rating scores by 110 users of the two transferred screenshots (T1 and T2) are higher than that of the original page.



**Fig. 18** The original Web page (the leftmost). In our user study, the average rating score by 110 users of the first screenshots (T1) is higher than the original one by 110 users.



**Fig. 19** The original Web page (the leftmost). In our user study, the average rating scores by 110 users of all the three screenshots are lower than the original one.

## 6 Discussion

The evaluations in Section 5 provide several initial qualitative analyses for different static part location methods, color theme extraction algorithms, assessment model construction approaches, and Web color transfer strategies. Some intuitive methods and existing image algorithms are compared in the evaluations. The results support the following conclusions. (1) The block-sampling static part location algorithm achieves the best division performances compared with other screenshot-based algorithms and an intuitive pixel-based algorithm. (2) The introduced method of outlier-aware clustering + block sampling-based location is superior to other methods discussed in Section 4.3 in color theme extraction for Web pages. (3) The color compatibility model learned by the proposed transfer learning with more target training data outperforms the model learned by the proposed transfer learning with less data.

(4) The PCA-based feature de-correlation improves or at least does not weaken the assessment accuracy. In addition, the colors generated by the proposed Web color transfer method are perceived more similar to the reference pages than those generated by the conventional image color transfer technique. The evaluation results suggest that the analysis, assessment, and editing for Web colors are different from those for image colors. New methods should be explored by considering the differences between Web page screenshots and images.

An online user study was launched to evaluate the colors of the new screenshots generated by the proposed color transfer and assessment framework. More than half of the top-3 generated new screenshots are judged to be better than the original Web pages in terms of colors. Considering this work is an initial attempt at bringing computer graphic techniques into Web color editing, and the results are promising.

## 7 Conclusions

This study has investigated the color compatibility assessment of Web pages and the transfer of colors between Web page screenshots, which are useful in evaluating and selecting of the colors of a Web page. Several new techniques are introduced to address the main challenges for Web color analysis. First, computer vision techniques are used to divide a Web page screenshot into static and dynamic parts. Second, an outlier-aware clustering method and a transfer learning method are introduced to extract color themes of the static parts and construct the assessment model based on existing online color theme-rating data respectively. Third, the color transfer between Web page screenshots is also investigated. The constructed compatibility assessment model and the proposed Web color transfer technique are then applied into a new application, namely, color recommendation for Web design. Experiments and online user study results suggest the effectiveness of our proposed methodologies for Web color compatibility assessment and transfer and the initial success of our proposed color recommendation application. The methods that are directly borrowed from image processing areas are inferior to our approaches in the evaluations.

Although color compatibility is important, other factors (e.g., fonts and images) also greatly affect the visual appearance of a web page. Our future work will consider more cues to assess the visual appearance of web pages.

## References

1. Alexa: <http://www.alexa.com/topsites> 11/30/2011. (2011)
2. Cao, L.J., Chua, K.S., Chong, W.K.: A comparison of pca, kpca and ica for dimensionality reduction in support vector machine. *Neurocomputing* **55**(1–2), 321–336 (2003)
3. Cohen-Or, D., Sorkine, O., Gal, R., Levvand, T., Xu, Y.Q.: Color harmonization. *ACM Trans. Graph.* **25**(3), 624–630 (2006)
4. Coursaris, C.K., Lansing, E., Swierenga, S.J., Watrall, E.: An empirical investigation of color temperature and gender effects on web aesthetics. *Journal of Usability Studies* **3**(3), 103–117 (2008)
5. Cyra, D., Head, M., Larios, H.: Colour appeal in website design within and across cultures: A multi-method evaluation. *International Journal of Human-Computer Studies* **68**, 1–21 (2010)
6. DIAOCHA: <http://www.diaochapai.net/> (2015)
7. Dong, W., Bao, G., Zhang, X., Paul, J.C.: Fast local color transfer via dominant colors mapping. In: *ACM SIGGRAPH ASIA 2010 Sketches*, pp. 46:1–46:2. ACM (2010)
8. Flatla, D.R., Reinecke, K., Gutwin, C., Gajos, K.Z.: Sprweb: Preserving subjective responses to website colour schemes through automatic recolouring. In: *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pp. 2069–2078 (2013)
9. Fleiss, J.L.: Measuring nominal scale agreement among many raters. *Psychological Bulletin* **76**(5), 378–382 (1971)

10. Forero, P.A., Kekatos, V., Giannakis, G.B.: Robust clustering using outlier-sparsity regularization. *IEEE Trans. Sig. Proc.* **60**(8), 4163–4177 (2012)
11. Goethe, J.: *Theory of Colors* (1810)
12. Hochbaum, D., Shmoys, D.: A best possible heuristic for the k-center problem. *Math. of Operations Research* **10**(2), 180–184 (1985)
13. Huang, J., Alexander, J.S., Gretton, A., Borgwardt, K.M., Schölkopf, B.: Correcting sample selection bias by unlabeled data. In: *The Neural Information Processing Systems (NIPS)*, pp. 601–608 (2006)
14. Ivory, M.Y., Megraw, R.: Evolution of web site design patterns. *ACM Trans. Inf. Syst.* **23**(4), 463–497 (2005)
15. Ivory, M.Y., Sinha, R.R., Hearst, M.A.: Empirically validated web page design metrics. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 53–60 (2001)
16. Kendall, M.: A new measure of rank correlation. *Biometrika* **30**, 81–89 (1938)
17. Kondratova, I., Goldfarb, I.: Color your website: use of colors on the web. In: *International Conference on Usability and Internationalization*, pp. 123–132 (2007)
18. Koufaris, M.: Applying the technology acceptance model and flow theory to online consumer behavior. *Information Systems Research* **13**(2), 205–223 (2002)
19. Kumar, R.: Data-driven interaction web design. In: *ACM symposium on User interface software and technology*, pp. 2197–2206 (2012)
20. Kumar, R., Talton, J.O., Ahmad, S., Klemmer, S.R.: Bricolage: example-based retargeting for web design. In: *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pp. 2197–2206 (2011)
21. Lafleur, C., Rummel, B.: Predicting perceived screen clutter by feature congestion. In: *Mensch & Computer 2011: 11. fachübergreifende Konferenz für interaktive und kooperative Medien. überMEDIEN-ÜBERmorgen*, p. 101 (2011)
22. Lalonde, J.F., Hoiem, D., Efros, A.A., Rother, C., Winn, J., Criminisi, A.: Photo clip art. *ACM Trans. Graph.* **26**(3) (2007)
23. Lee, S., Koubek, R.J.: The effects of usability and web design attributes on user preference for e-commerce web sites. *Computers in Industry* **61**(4), 329–341 (2010)
24. Lin, S., Hanrahan, P.: Modeling how people extract color themes from images. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13*, pp. 3101–3110 (2013)
25. Ling, J., Van Schaik, P.: The effect of text and background colour on visual search of web pages. *Displays* **23**(5), 223–230 (2002)
26. Liu, W., Meng, X., Meng, W.: Vide: A vision-based approach for deep web data extraction. *IEEE Trans. Knowl. Data Eng.* pp. 447–460 (2010)
27. Nielsen, J.: F-shaped pattern for reading web content (2006). [Http://www.useit.com/alertbox/reading\\_pattern.html](http://www.useit.com/alertbox/reading_pattern.html)
28. O'Donovan, P., Agarwala, A., Hertzmann, A.: Color compatibility from large datasets. *ACM Trans. Graph.* **30**(4), 63:1–63:12 (2011)
29. Ou, L.C., Luo, M.R., Woodcock, A., Wright, A.: A study of colour emotion and colour preference. part i: Colour emotions for single colours. *Color Research & Application* **29**(3), 232–240 (2004)
30. Ou, L.C., Luo, R., Wright, A.W.A.: A study of colour emotion and colour preference. part i: Colour emotions for single colours. *Color Research & Application* **29**(3), 232–240 (2007)
31. Park, S., Choi, D., Kim, J.: Critical factors for the aesthetic fidelity of web pages: empirical studies with professional web designers and users. *Interacting with Computers* **16**, 351–376 (2004)
32. Pitié, F., Kokaram, A.C., Dahyot, R.: Automated colour grading using colour distribution transfer. *Comput. Vis. Image Underst.* **107**(1-2), 123–137 (2007)
33. Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, T., Gajos, K.Z.: Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. In: *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pp. 2049–2058 (2013)
34. Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. *IEEE Comput. Graph. Appl.* **21**(5), 34–41 (2001)
35. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vision* **40**(2), 99–121 (2000)
36. Schloss, K.B., Palmer, S.E.: Aesthetic response to color combinations: preference, harmony, and similarity. *Attention perception psychophysics* **73**(2), 551–571 (2011)
37. Schmidta, K.E., Liua, Y., Sridharanb, S.: Webpage aesthetics, performance and usability: Design variables and their effects. *Ergonomics* **52**, 631–643 (2009)
38. Su, Z., Zeng, K., Liu, L., Li, B., Luo, X.: Corruptive artifacts suppression for example-based color transfer. *IEEE Transactions on Multimedia* **16**(4), 988–999 (2014)
39. Tai, Y.W., Jia, J., Tang, C.K.: Local color transfer via probabilistic segmentation by expectation-maximization. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 747–754 (2005)

40. Thorlacius, L.: The role of aesthetics in web design. *Nordicom Review* **28**(1), 63–76 (2007)
41. Valdez, P., Mehrabian, A.: Effects of color on emotions. *Journal of Experimental Psychology* **123**(4), 394–409 (1994)
42. Wang, B., Yu, Y., Wong, T.T., Chen, C., Xu, Y.Q.: Data-driven image color theme enhancement. *ACM Trans. Graph.* **29**(6), 146:1–146:10 (2010)
43. Xiao, Y., Wan, L., Leung, C.S.: Example-based color transfer for gradient meshes. *IEEE Transactions on Multimedia* **15**(3), 549–560 (2013)
44. Yang, Y., Zhao, H., You, L., Tu, R., Wu, X., Jin, X.: Semantic portrait color transfer with internet images. *Multimedia Tools and Applications* pp. 1–19 (2015)
45. Zhu, X., Gauch, S.: Incorporating quality metrics in centralized/distributed information retrieval on the world wide web. In: *International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 288–295 (2000)