

Affective image adjustment with a single word

Xiaohui Wang · Jia Jia · Lianhong Cai

Published online: 27 October 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract We present a complete system that automatically adjusts image color to meet a desired emotion. It will be more convenient for users, especially for non-professional users, to adjust an image with a semantic user input, for example, to make it lovelier. The whole algorithm is fully automatic, without any user interactions, and the inputs are simply the original image and an affective word (e.g. lovely). To achieve this goal, we solve several non-trivial problems. First, in order to find the proper color themes (template of colors) to reflect the expression of the affective word, we exploit the theoretical and empirical concepts in famous art theories and build a color theme—affective word relation model allowing efficient selection of candidate themes. Furthermore, we propose a novel strategy to select the most suitable color theme among the candidates. Second, to adjust image colors, we propose the Radial Basis Functions (RBF) based interpolation method, which is more effective in many scenarios as evidenced in experiments. We also evaluate the system with comprehensive user studies and its capability is confirmed by the results.

Keywords Affective image adjustment · Color theme · Art theory

Electronic supplementary material The online version of this article (doi:10.1007/s00371-012-0755-3) contains supplementary material, which is available to authorized users.

X. Wang (✉) · J. Jia · L. Cai
Department of Computer Science and Technology, Tsinghua University, Beijing, China
e-mail: wangxh09@mails.tsinghua.edu.cn

1 Introduction

Every artistic picture conveys its own emotion. In our daily life, increasingly more digital pictures with various artistic styles are available with the rapid development of digital capture devices, interactive creation softwares, and the Internet. Meanwhile, the booming needs for editing these pictures for various reasons emerge, among which simply changing the emotions is in great demand. Note that most users do not have any systematic training in art. For these users, editing with professional softwares such as Photoshop is either impossible or time consuming. In this paper, we present a comprehensive system that automatically adjusts an image to meet a desired emotion (affective image adjustment). That is, the user simply inputs an image and an affective word (e.g. lovely), and the system automatically adjusts the image to the target emotion (e.g. change the image from the sad emotion to loveliness).

Psychological studies had concerned more about the connections between image color themes defined as template of colors and emotions, and convinced the strong correlations between them [1, 11–13]. For example, *romantic* often corresponds to the template of pink, light blue, and white; while *serious* is often expressed by dark colors, e.g. dark blue or deep grey. As a result, to express a specific emotion, color themes tend to be adjusted. However, the existing commercial softwares such as Photoshop do not directly support this adjustment and need tedious operations. It is a tough job for an untrained user to edit with these softwares.

Recent methods such as color transfer [25, 32], image colorization [5, 9, 16] and edit propagation [18, 33] can avoid previous painful operations. However, color transfer needs a similar image with the desired emotion, which is not easily obtained in most situations. Both colorization and edit propagation need user scribbles and specified colors to

control the whole process. Accordingly, rich experience is often required to carefully select reference colors and their positions.

More recently, Wang et al. presented a segmentation based color enhancement method with a given target theme [30]. This method is able to adjust an image to a given color theme with optional user scribbles. However, choosing a color theme for a desired emotion is hard for non-professional users. Tens of color themes may match an affective word, and sometimes just one of them fits best for a given image. It is painful to select from dozens of candidate color themes even for skilled users. Besides, their method depends on texture modeling which is itself a tough problem [15, 28], and is not suitable for images without obvious textures such as cartoon images or textures without a common color pattern such as clothes. It also means additional training for a new image category containing new types of textures. Furthermore, they rely on a highly non-linear hard optimization to balance the original, the target, and transcendental colors, and this sometimes may result in a blurring effect for the tradeoff among these different colors. Beyond that, their method requires optimal soft segmentations, which is hard to be automatic for textureless images. Although a manual tuning is allowed, this sometimes requires trivial user interactions.

Thus, an ideal approach should automatically adjust the image with a verbal instruction, without any knowledge of art or trivial user interactions. Besides, it should also be common and suitable for Internet images, without the need of a database, which restricts it to a specific range of images.

This is an original work which targets affective image adjustment for users without any experience, who only need to input an image and an affective word. Our method only needs a lightweight training database and does not require necessarily additional training for a new image category, thus it is widely applicable for adjusting general images.

Our major contribution is a complete system, which is able to adjust image appearance with just one affective word. This system includes several indivisible components, each of which has its own benefit, including:

1. *Modeling the relationship between color themes and affective words:* Starting from theoretical and empirical concepts in the art theories, we carefully design a color composition theory guided feature descriptor (CC-features), followed by a thoughtfully selected regression to map the feature vectors to the image-scale space which is commonly used to quantify emotions (see Sect. 2.3 for more details). Experiments show that our CC-features produce smaller errors on the test sets, compared with the traditional features [24]. This model also facilitates the use of rich Internet color themes and allow us to construct an expressive color theme database. Based on the

database, we can select a set of closely related candidate themes for an affective word efficiently.

2. *A new color theme extraction algorithm:* We give a new insight into the extraction and design an optimization balancing the area, contrast, and position rules. Experiments show that our method produces the results which are more consistent with visual perception. The extracted theme is then used to assist a selection process to find the most suitable theme from the candidates.
3. *A new image adjustment algorithm with a given color theme:* Based on the modeling and color theme selection in previous steps, we are convenient to use a relatively lightweight method for the final color adjustment. We use a direct approach to determine the relationships between individual pixels and the colors in a target theme. An RBF-based interpolation method is proposed to ensure the smoothness. Results show that this method can also reduce the original-target color blurring effect.

2 Related work

2.1 Color adjustment

Color harmonization Cohen-Or et al. [6] used classical art theories for image color adjustment and then enhanced the harmony among the colors of a given image for pleasant visual perception. This work only focuses on the color harmonization of an image and normally does not require user interactions. The main highlight of this work is to show the power of classical art theories which summarizes the sense of beauty in a compact and usually computable manner. After that, numerous other works [17, 27, 29] on various subjects arose based upon these art theories.

Color transfer and colorization Reinhard et al. [25] proposed a statistics based approach for color transfer between images. The method was extended to grayscale images in [32]. Alex Yong-Sang Chia et al. [5] proposed a colorization method based on the Internet images with a user given segmentation mask and text labels. Wang et al. [31] also developed an interesting method for color and tone adjustment from examples. Though their adjustment results may convey specific emotions, finding a reference image with both similar appearance and the desired emotion is often a tough job. Even with an Internet based approach [5], it still needs the existence of similar objects with popular names. These seriously limit the methods available for the affective image adjustment.

Edit propagation In edit propagation, users draw a few strokes and these edits are automatically propagated to the rest of the image based on the similarity among pixels [9,

16, 19]. Li et al. [18] formulated the edit propagation as an interpolation problem and greatly improved the efficiency. In their works, users still need to manually select regions and target colors, not as simple and direct as in Wang's method [30].

Color theme enhancement Wang et al. [30] presented an image color theme adjustment method with an image and an target color theme as the inputs. Compared to previous methods, this approach is able to work without a mandatory requirement of user interactions, although the user's tuning can improve the required image segmentation, and thus makes a considerable improvement of the final result. This method also pays much attention to transcendental texture color information and is expected to produce a more natural result if the same kind of texture is fortunately recorded and accurately located in the database.

2.2 Internet based image processing

Recently, Internet based image processing has attracted more attentions, as it has the big advantage of huge amount of media resources. Most of the studies focus on Internet based image composition [2, 10], and some emphasize the construction of attribute databases [3]. In our work, we utilize the Internet resources to build the color theme database, which is expressive enough to meet various affective adjustment requirements.

2.3 Color and emotions

Describing emotions is a core task in affective computing. Generally, there are two means to describe emotions: the emotion class and the emotion dimension. The emotion class is a traditional method for emotion classification, for example, the well-known Ekman's six basic emotions (*happy, sad, surprise, fear, anger* and *disgust*) [7]. In a more recent work, Machajdik et al. used four positive emotions (*amusement, awe, contentment, excitement*) and four negative emotions (*anger, disgust, fear, and sad*) for image classification [21]. The dimensional approach gives specific formulae to quantify the emotions and is more widely used because of its accuracy. The most famous two-dimensional system widely used in art design is the image-scale space with two dimensions warm-cool and hard-soft [12, 13].

Among the studies of the image-scale space, Kobayashi's famous art work [13] maps 490 5-color themes as well as 180 words like *sweet, romantic* to the image-scale space. It is based on the long-term psychophysical investigations and commonly used in various applications about the emotions of images [26, 27]. Although the collection is not large, it is typical enough as training data and can achieve high degree of consensus on this highly subjective task. Figure 1

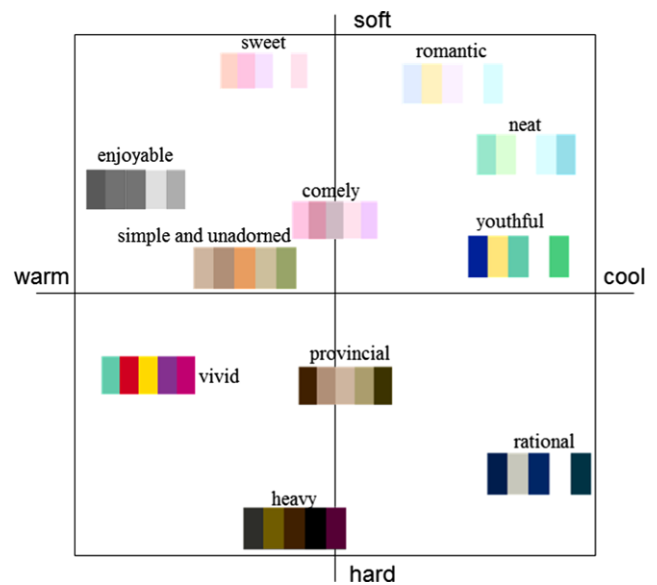


Fig. 1 Examples of 5-color themes located in the image-scale space

illustrates a few examples of 5-color themes and their corresponding affective words located in the image-scale space. In this paper, we adopt this image-scale system to model the relationship between color themes and affective words.

3 Overview

The framework of our system is illustrated in Fig. 2, consisting of an offline phase and a runtime phase. In the offline phase, we model the color theme— affective word relationship by exploiting theoretical and empirical concepts from the art theories [1, 11, 13, 22]. Both affective words and color themes are mapped to the image-scale space. The 180 affective words in Kobayashi's theory [13] have predetermined image-scale coordinates and are treated as landmark words to interpolate any input word. We use the 5-color theme because of its representation power and rich Internet resources. The art principles lead to a well-designed CC-features (a color composition theory guided feature descriptor), with which we train a model relating color themes to the image-scale space. Finally, we use Internet resources to enlarge the database to more than 400,000 color themes.

During the runtime phase, for an affective word, we adopt a standard semantic similarity computation to get the weight corresponding to each landmark word, and finally get the coordinate in the image-scale space. M nearest candidate color themes are selected. Then we propose a strategy to choose the most appropriate color theme balancing the appearance compatibility with the input image and suitability with the affective word. Finally, we adjust the image appearance according to the selected color theme using an RBF based interpolation approach.

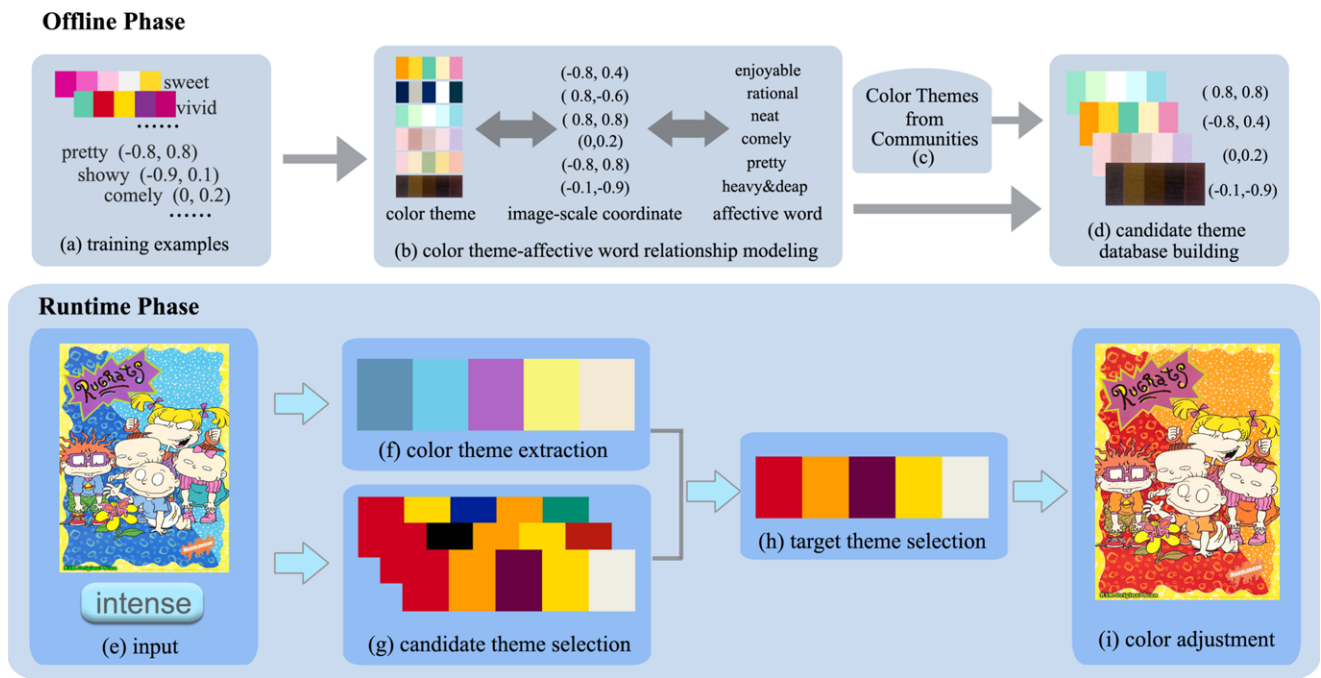


Fig. 2 The overall pipeline of our framework. Our framework can be divided into the offline phase and the runtime phase. In the offline phase, we model the color theme-affective word relationship (b) from the training examples (a). The candidate theme database (d) is built based on this relationship model and on-line communities (c). In the

runtime phase, the input (e) is an image and an affective word. Candidate themes (g) are selected with the affective word from the candidate theme database (d). Using a strategy to match the candidate themes (f) extracted from the input image, we select the most suitable target theme (h) and adjust the image (i) accordingly

4 Modeling color theme— affective word relationship

4.1 CC-features

In the art theories, color pattern is commonly used to describe the relationship among colors in a theme and abstract the color composition. Our CC-features are composed of two parts: the color theme itself and the color pattern related features. Given a color theme, we first convert it to the HSV color space and compute the differences (mean, max, min) between adjacent colors across each single channel (refer to the color contrast rules in [1, 11, 22]). These differences along with theme colors themselves form the first part (the color theme itself) of the CC-features.

According to the key points in [13], color pattern plays more important role than color itself in human visual perception. For instance, no matter what a color theme actually is, one with a strong contrast tends to express the *intense* emotion; on the other hand, that with a weak contrast tends to express the *mild* emotion. From the complex artistic descriptions in [1, 11, 13, 22], we summarize six types of color pattern (Table 1), which are intuitive and most commonly used. We quantify the decision rules of color patterns to form the pattern related part: the number of hues (the MDF (main decision factor) of the hue color pattern), the number of tones (the MDF of the saturation color pattern), the maximum minus the minimum across each single channel in the HSV

Table 1 Color pattern type

Color pattern	Decision rules with the artist experience	Example
Hue	4–5 hues in a 5-color theme; using tone V, S, B, Dp can achieve obvious effect	
Saturation	2–3 hues in a 5-color theme	
Unified	use the same or similar hues/tones	
Sharp	use strong or mild contrast in hue contrast; use tone with enormous difference	
Clear	use clear tone (V, B, P, Vp) matching with white or black	
Dull	use dull tone (S, L, Dl, Lgr, Gr, Dp, Dk, Dgr) matching with grey	

color space (the MDF of the unified and sharp color patterns), and the number of clear colors (the MDF of the clear and dull color patterns). The above numbers are based on the Practical Color Coordinate System (PCCS) [27], which is a widely used discrete color space. In PCCS, chromatic colors are discretized to 10 hues and 12 tones for each hue. The 12 tones can be divided into clear colors (V, B, P, Vp) and

dull colors (S, L, DI, Lgr, Gr, Dp, Dk, Dgr), respectively. All the above features form a 30-dim vector for later learning.

4.2 Model learning

We relate affective words and color themes by mapping both to the image-scale space. The mapping of affective words involves word similarity computation and interpolation among predetermined landmark words, which is calculated in the runtime phase (Sect. 5.1). We integrate our CC-features into a regression framework to predict the image-scale coordinates of color themes.

The input data comprises color themes $\{T_i\}$ and the image-scale coordinates $\{(wc_i, hs_i)\}$, where $T_i \in R^{15}$ represents 5-color theme; wc_i and $hs_i \in [-1, +1]$ are the values for the warm–cool and hard–soft axes, respectively. Our goal is to predict the image-scale coordinate (wc_{new}, hs_{new}) for a new theme T_{new} . The dataset used to learn the model contains 490 5-color themes and all labeled with the coordinates in the image-scale space.

Then we learn a continuous mapping from the CC-features to the image-scale coordinate (wc, hs) . We adopt the LASSO regression framework, as some recent studies [24] show that it generally gives better results in scoring the quality of 5-color themes compared with SVM regression (RBF kernel) and K-Nearest Neighbors. The learned weights for the mapping can be found in the supplemental materials. Section 6.2 gives experiment results on the CC-features and traditional features [24] using the same regression method.

4.3 Candidate theme database building

After learning the mapping to image-scale space, we build a candidate theme database and expect that there exist enough color themes near every affective word.

Since 5-color theme representation is compatible with the industrial practices [30], we collect the 5-color themes from Adobe Kuler (44,986 themes, kuler.adobe.com) and COLOUR Lovers site (383,938 themes, www.colourlovers.com), which are famous on-line communities to create and share color designs. Using our learned model, we map all the themes to the image-scale space. Thus, our database contains a large amount of color themes to ensure enough candidates for an affective word.

5 Affective image color adjustment

5.1 Candidate color theme selection

Affective words to the image-scale space mapping For a specific input word, we compute its similarities with all the landmark words. The landmark words are described in [13]

and have precise coordinates based on psychological experiments. Currently, we build the similarity model based on HowNet [8], which is a bilingual general knowledge infrastructure describing relations between concepts and attributes. For a given word x , we compute its associativity with every landmark word (a real value in the range [0, 1]) and select k landmarks ($k = 5$ in our experiments) with the largest associativity weights. Suppose the k landmarks have coordinates I_1, I_2, \dots, I_k in the image-scale space with associativity weights w_1, w_2, \dots, w_k , then the weighted coordinate is simply given by $\frac{\sum_{i=1}^k w_i I_i}{\sum_{i=1}^k w_i}$.

Candidate color theme selection M color themes ($M = 100$ in our experiments) nearest to the given word in the image-scale space are selected from the candidate theme database. We use the approximate nearest neighbors algorithm (ANN) for acceleration [23].

5.2 Color theme extraction

We extract the color theme for the original image to evaluate its compatibility with the candidate themes (Sect. 5.3). Here, we propose a new insight into the color theme extraction. In this new approach, we take into account the rules obtained by empirical concepts from the art theories [1, 11, 22] and summarize them as follows. The basic one is the area rule: the colors with large areas determine the basic colors of an image. Although this rule is used in classical color theme extraction [26, 27], we integrate it with another powerful rule—the contrast and position rule: human visual system is sensitive to colors which are contrasted with neighboring colors and often near the center. This rule is verified by the psychological studies [1] and used in the saliency detection methods [4, 20]. However, it greatly improves the color theme extraction results. Compared with previous works, the extracted theme better agrees with human visual perception (see Sect. 6.3 for the comparison results).

Before applying the above rules, we use a segmentation based color quantization to make the later optimization easier. We first segment the image into regions and then iteratively combine regions with distance of average colors (Euclidian distance in the LAB color space with each channel normalized to [0, 1]) smaller than a given threshold (we use 0.15 in our experiments). This simple approach can avoid small fragments with similar color appearance in the final segment results. Then every pixel is assigned to the average color of its region. A detailed illustration of the color quantization is shown in Fig. 3. This quantizing scheme has a property that the same quantified colors often correspond to regions rather than individual pixels, which is more consistent with the human tonal perception system [1].

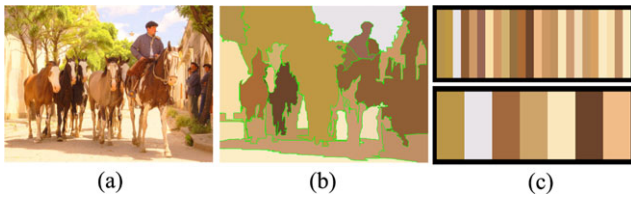


Fig. 3 Color quantization. **(a)** Input image, **(b)** segmentations and average colors, **(c top)** colors before combination, **(c bottom)** quantified colors

By quantizing the rules, we attempt to extract dominant colors by designing the following objective function:

$$\arg \max_T \sum_{i=1}^n [\omega_1 A(t_i) + \omega_2 E(t_i)] \quad (1)$$

where $T = \{t_1, t_2, \dots, t_n\}$ is the extracted color theme, $n = 5$ for a 5-color theme. t_i is a color selected from the quantified colors. $A(t_i)$ is defined as the ratio of pixels with color t_i , which quantizes the area rule (the area probability of t_i). $E(t_i)$ is a term quantizing the position and contrast rule. ω_1, ω_2 are constants used to balance the importance of these two rules and set to 1 and 30 respectively in our experiments. ω_2 is much larger to make the two energy terms in similar order of magnitudes.

The E term is defined as

$$E(t_i) = P(t_i)S(t_i) \quad (2)$$

where

$$P(t_i) = 1 - \frac{\|L(t_i) - L_m\|_2}{\|L_m\|_2} \quad (3)$$

$$S(t_i) = \frac{1}{m} \sum_{j=1}^m A(c_j) D(t_i, c_j) \quad (4)$$

where the P term measures centrality of a color and S is a contrast based saliency measure. m is the number of quantified colors, iterated with c_j . $L(t_i)$ is the barycenter of pixels with color t_i and L_m is the image center. $D(t_i, c_j)$ is the Euclidian distance between two colors.

Arrangement of color themes The order of colors in a theme is important for human visual perception and is also crucial to match two themes. We analyzed hundreds of color themes in [13] (totally 490, more detailed statistics can be found in the supplemental materials) and found that most of them followed the two interesting rules: (1) colors are arranged in the order of hue wheel; (2) colors change smoothly, that is, the sum of distances between adjacent colors is as small as possible.

Based on these observations, we design an objective function to judge the comfortable feeling of an arrangement:

$$\arg \min_{T' \in \text{Perm}(T)} \alpha_1 \sum_{i=1}^{n-1} D(t'_i, t'_{i+1}) + \alpha_2 K(T') \quad (5)$$

where $T, T' = \{t'_1, t'_2, \dots, t'_n\}, n, D$ is the same as those in Eq. (1) and Eq. (4). $\text{Perm}(T)$ denotes the 120 permutations of the 5-color theme T . $K(T')$ is the number of colors which do not meet the first arrangement rule. α_1, α_2 are the weights balancing the importance of the two rules and typically set to 0.7 and 0.3, respectively.

Note that our color theme— affective word model also relies on the order of colors. To make them obey the same arrangement, we use this rule to adjust the order for every color theme in our database.

5.3 Target theme selection

Since an emotion can be expressed by multiple color themes, we select the most suitable color theme by the following formula:

$$\arg \min_{T \in C_T} \lambda_1 D_I(I(T), I_w) + \lambda_2 \sum_{i=1}^{n-1} |D(t_i, t_{i+1}) - D(o_i, o_{i+1})| \quad (6)$$

The first term determines the coincidence between the color theme and the input word, and the second measures the compatibility of a candidate theme with the extracted theme. λ_1 and λ_2 are weights balancing the two terms (we choose 1 and 0.25, respectively, in our implementation). $I(T)$ and I_w are the image-scale coordinates of the color theme and the word, respectively, and D_I is the Euclidian distance metric in the image-scale space. C_T is the set of candidate themes. o is the color theme of the input image. t_i and o_i are the i th color in the theme T and theme o respectively. D is the same as that in Eq. (4) and $n = 5$.

Note that the second term actually measures the difference of color changing between adjacent colors in a color theme. This clearly follows the intuition that we want to keep the color contrast, that is, for regions with strong color contrast, the adjusted image also keeps such differences.

5.4 Image color adjustment

Given the chosen target color theme and the input image, we formulate the color adjustment as an interpolation problem and tend to use a Radial Basis Function (RBF) approach to ensure the smoothness.

Our image color adjustment is typically guided by two principles. First, the original and target color theme are both ordered using our arrangement rule, as we expect a one-to-one correspondence. Second, pixels with similar colors are more likely to receive similar amount of adjustment.

Intuitively, for the colors sufficiently close to one of the extracted theme colors, they should be assigned to the corresponding color in the target theme. Instead, we keep the colors which are far from the extracted theme colors similar to their originals.

Formulation We consider image color adjustment as an interpolation problem. For each pixel i , the adjusted color A_i is given by

$$A_i = \alpha_0^i C_i^o + \sum_{m=1}^n \alpha_m^i C_m^t \tag{7}$$

where C_i^o is the color of the i th pixel in the extracted theme of the input image. C_m^t is the m th color in the target theme, and $n = 5$ as before. α_m^i is the normalized interpolation weight, and $\sum_{m=0}^n \alpha_m^i = 1$ is guaranteed.

Furthermore, α_m^i is normalized by $\alpha_m^i = \beta_m^i / \sum_{k=0}^n \beta_k^i$, $m = 0, 1 \dots n$, where β_k^i is represented by $h_k(\mathbf{f}_i)$ with \mathbf{f}_i the feature vector of the pixel i . $h_k(\mathbf{f})$ is defined as

$$h_k(\mathbf{f}) = \sum_{j \in R_k} \omega_j \phi(\|\mathbf{f} - \mathbf{f}_j\|) \tag{8}$$

where ϕ is the predefined radial basis and is set to Gaussian function as in [18], i.e., $\phi(r) = \exp(-r^2)$. \mathbf{f}_i is a 5-dimensional vector consisting of both the color (in LAB color space) and the location of the i th pixel.

For $k = 1..n$, R_k is a randomly selected subset of pixels corresponding to the k th color in the extracted color theme, and R_0 is a randomly selected subset of pixels not close to any color in the extracted theme (see *Construction of h* for details).

Intuitively, h_k measures the similarities of a pixel to a color in the extracted theme or colors beyond the extracted theme.

Construction of h We build the function h by solving the coefficients ω_j s according to the extracted color theme. For each pixel in the image, if the distance to the nearest color (suppose the k th color) in the extracted theme is smaller than a given threshold (set to 0.4 in our paper), it is labeled l_k , otherwise it is labeled l_0 (see Fig. 4 for an illustration). Let S_k be the set of pixels labeled l_k , and R_k are subset of S_k randomly selected (we select 20 pixels for each R_k). Then we solve ω_j s as a nonnegative least-squares minimization problem [14]:

$$\arg \min_{\omega_j} \sum_{i \in S_k} \left(1 - \sum_{j \in R_k} \omega_j \phi(\|\mathbf{f}_i - \mathbf{f}_j\|) \right)^2 \quad \text{s.t. } \omega_j \geq 0 \tag{9}$$

Generally, we use a simple rule to solve a tough problem in the image color adjustment. We need to consider what adjustment every pixel should receive for a given color theme. Based on the previous selection process, the chosen target theme practically shares similar color contrasts as the original image, and this further improves the appearance consistency. Our comprehensive experiments also convince the reliability of this approach.

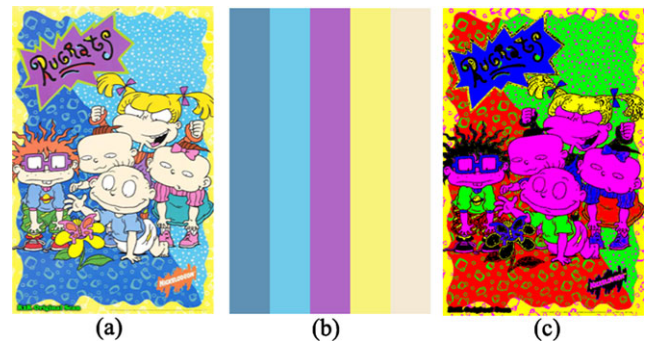


Fig. 4 Construction of h . (a) Input image, (b) extracted theme of input image, (c) regions labeled with l_1, l_2, l_3, l_4, l_5 , and l_0 , represented by the colors red, green, blue, yellow, pink, and black

Table 2 Comparison results between our CC-features and the traditional features

Feature	Warm–Cool		Hard–Soft	
	MAE	MSE	MAE	MSE
Traditional features	0.2529	0.1243	0.1689	0.0440
CC-features	0.2402	0.1111	0.1251	0.0269

6 Experiments

6.1 Results and performance

We have implemented our system and successfully tested it on a variety of images. The results are demonstrated in Fig. 13, and more results can be found in supplemental materials. On a PC with two quad-core 2.27 GHz CPUs and 4 GB RAM, our method takes less than 1 second for a 1024×768 image, allowing interactive feedbacks.

6.2 Evaluation on modeling color theme— affective word relationship

Visual features are important in the relationship model learning. We compare the accuracy between our CC-features and the traditional features. The traditional features are commonly constructed by utilizing a miscellaneous mass of simple color statistics. We use a state-of-art feature descriptor in [24], which integrates comprehensive traditional features and has 334 dimensions.

Table 2 shows the comparison results between our CC-features and the traditional features using a 0.6/0.4 train/test split on the database [13]. The results indicate that the model trained with our CC-features decreases the error in both MAE (mean average error) and MSE (mean squared error) measurements.

6.3 Evaluation on color theme extraction

We compare our color theme extraction algorithm with various previous works. Figure 5 shows the comparison results.



Fig. 5 Comparison of color theme extraction. (a) Theme extraction with the compatibility model [24], (b) K-means clustering [30], (c) segmentation based method [26], (d) color histogram based method [27], (e) our method without theme arrangement, (f) our method

K-means clustering [30] tends to blur similar colors in different regions (Fig. 5(b) loses the blue color of the sky as it is similar to the large white areas). O'Donovan et al. [24] combine the modified K-means clustering and the compatibility model, and this method shares the same disadvantage as the K-means method (Fig. 5(a)). Segmentation based method [26] and color histogram based method [27] can partially avoid the problem of blurring colors, but they may lose small regions. When the small regions contain important colors with strong contrast with others, these methods may lose them in the final theme (Fig. 5(c) and Fig. 5(d) lose the dark brown color of a horse).

Our method is also segmentation based. It is often more appropriate as it considers both the area and color contrast factor (Fig. 5(e)). Furthermore, our method gives more comfortable colors by taking into account the color arrangement rule (Fig. 5(f)). Figure 6 shows more color theme extraction results.

6.4 Evaluation on target theme selection

To evaluate the benefits of our target theme selection, we compare the results with (Fig. 7(c)) and without (Fig. 7(b)) our target theme selection algorithm. Both color themes can reflect an *elegant* feeling. However, the first 3 colors in the original theme describe a gradual changing blue in the image, while the greedily selected theme shows stronger contrast. In Fig. 7(b), the adjusted image tries to show the gradual changing with the 3 distinct colors which have sharp contrast, and this leads to the loss of the purple color, which causes a *less elegant* feeling. Our selection algorithm chooses a more compatible theme, and this avoids the artifacts (Fig. 7(c)).

6.5 Comparison with the 3-color themes

While 3-color themes are also used in practice, however, we focus on the use of 5-color themes in this paper. 5-

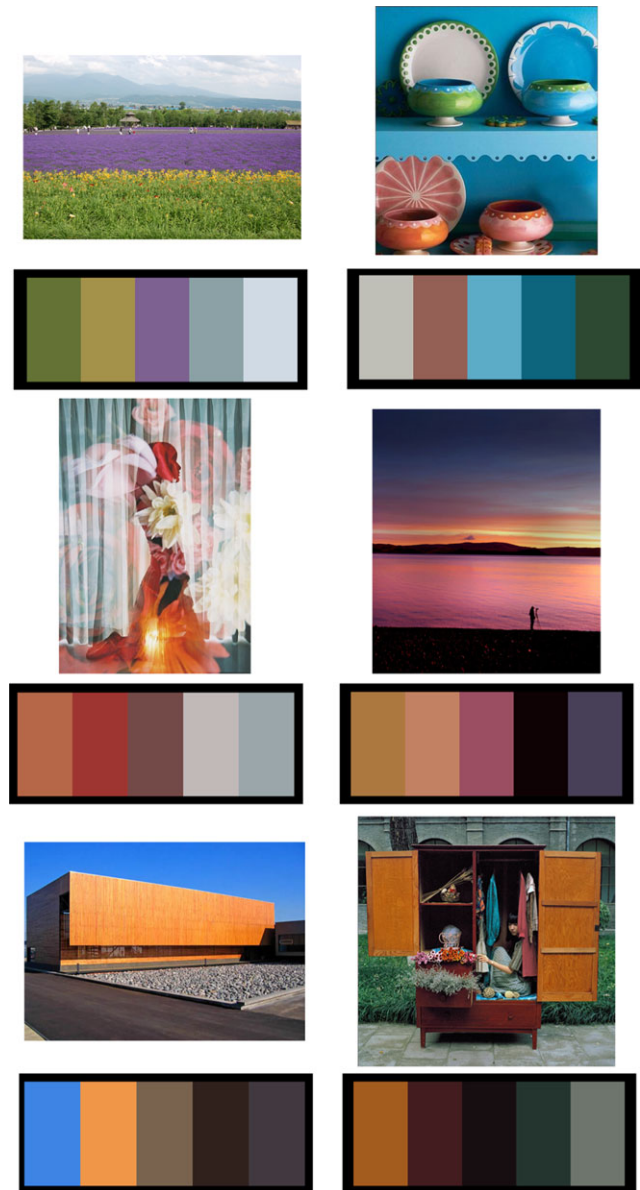


Fig. 6 More color theme extraction results

color themes are generally more representative than 3-color themes, especially when the images contain abundant colors. Figure 8 shows a comparison example. The input in the first row contains fewer colors and the two results look the same. However, when the input image in the second row contains plenty of colors, the result with the 3-color theme seems duller than that with the 5-color theme.

6.6 Comparison with color theme enhancement

We compare our method with Wang's method [30] on a given image with a predefined target color theme. Figure 9(a) is a textureless input, and we ignore the transcendent texture color constraints. The user provides 5 scribbled

Fig. 7 Target theme selection (with the affective word *elegant*). (a) Input image and its color theme, (b) result image with greedy selection (selecting the nearest theme in the image-scale space) and corresponding color theme, (c) result image with our target theme selection and corresponding color theme

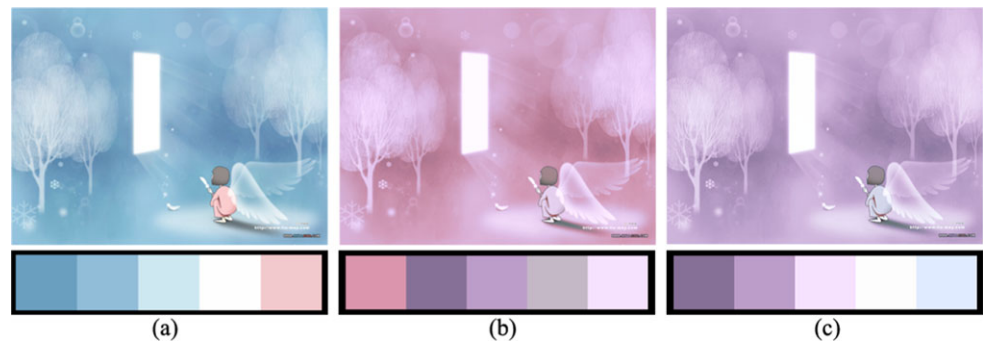


Fig. 8 Comparison with 3-color themes. (a) Input images, (b) 3-color themes and corresponding results, (c) 5-color themes and corresponding results



seeds. The final result of [30] (Fig. 9(b)) appears to be purple, which is the blurred color of pink (in the target) and blue (in the original image). The main reasons are the balance consideration of both the original colors and the target colors and the optimization deficiency for the highly nonlinear energy function. Similar issues are also addressed in [30], as they mentioned “color scribbles are often necessary over textureless regions.” However, this means elaborate color control and violates the simple and convenient goal. Our method with the same target theme is shown in the first row of Fig. 13 (the *romantic* label). We show another case in Fig. 10. Wang’s method [30] may introduce a nonoptimal color correspondence and causes a further color blur (the large grey background areas are the mixed result of the last two colors in the predefined theme), which mainly because of incorrect texture information in these cartoon styles. However the theme color mapping induced by our arrangement can practically avoid these problems.

6.7 User study

We have carried out a comprehensive user study to evaluate the results.

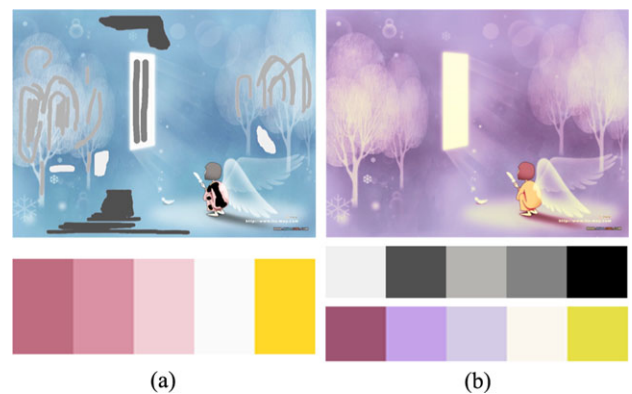


Fig. 9 A failure case of color theme enhancement [30]. (a) Input image with scribbled seeds and the target color theme, (b) result and optimized color (*bottom*) for each seed (*middle*, indicated by seed colors) by color theme enhancement [30]

We invited 15 participants (5 females and 10 males) with ages ranging from 22 to 35, each of which had normal vision. We randomly chose 20 test cases (in the supplemental materials), each of which contains the images adjusted by our method and by an artist using Photoshop, as well as the corresponding affective word. The artist was given 15 minutes to process an image. This is reason-

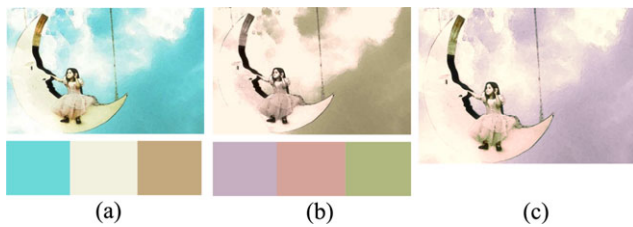


Fig. 10 Comparison with color theme enhancement [30]. (a) Input image and extracted color theme, (b) result generated by color theme enhancement [30] and the predefined color theme, (c) result by our method

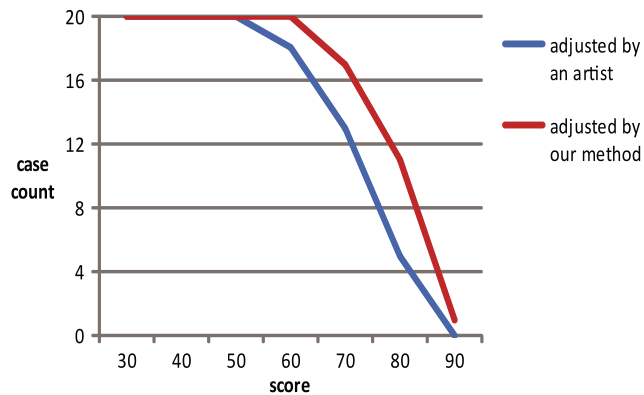


Fig. 11 Comparison with the number of cases against the scores

able as editing for a long time is tiresome and not practical. Participants were given a test case including an affective word and the images adjusted by our method and the artist one time in a random order. For each test case, participants were asked to rate the two images by two criteria similar to [30]: (1) how good an image corresponds with the given affective word; and (2) the perceptual image quality (naturalness). That is, the more natural an image looked and the more consistent it was with the affective word, the higher score he/she gave. In our tests, the score ranged from 0 to 100, and we collected them for adjusted images by our method and by the artist, respectively.

We analyzed the statistics on the collected data. Interestingly, we observed that results by our method were even better than those by the artist.

1. 69.7 % of the results take the adjusted images by our method as better consistent ones, while 30.3 % take the images by the artist. It indicates that in most cases of the highly subjective tests, our method can adjust the images to better meet the desired emotions.
2. For each case, we calculated the average score and added up cases with higher score than a given value. The result is shown in Fig. 11, indicating a per-score improvement of our method.

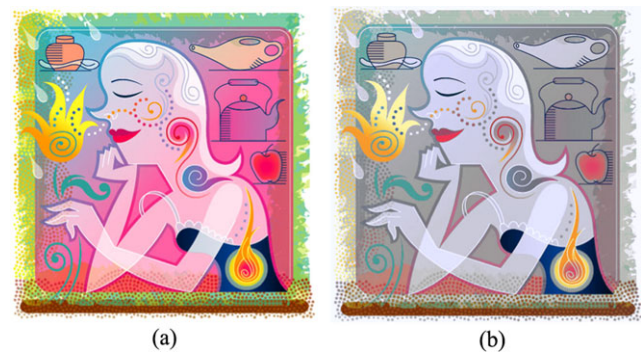


Fig. 12 A failure case containing more than 5 hues. (a) Input image, (b) the adjusted result with *deep*

3. The overall mean score of adjusted images by our method is 80.2 while that of images by the artist is 73.3, indicating our method can even beat the artist given a limited time.

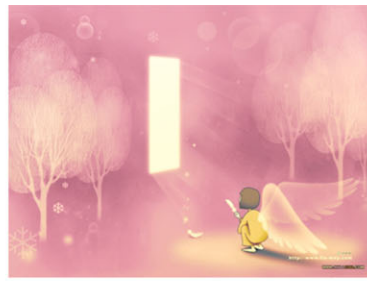
However, the reasons behind the amazing results are understandable. First, manual color adjustment is a boring task and the artist is unlikely to notice every detail of the image within a limited time. Second, we have a color theme database support which is able to efficiently relate a word to corresponding candidate color themes. And statistically, the database is more universal than a single artist's feelings.

7 Conclusions

In this paper, we have presented an automatic image color adjustment system for a desired emotion. With a single affective word input, the system does not need any user interactions, which is easily accessible for nonprofessionals. With our system, common users can now choose words to express their initial thoughts directly on various images and get adjusted ones.

In different stages of the system, including color theme – affective word relationship building, color theme extraction, and color theme guided image adjustment, we have made significant improvements over the stage-of-art methods.

Limitation In this paper, we have adopted a commonly used 5-color theme to describe the image color composition [13, 30], which is often more expressive than 2/3-color themes. And a recent study [24] convinces that there is a preference to color themes with 2–3 hues, so a 5-color theme is enough. However, sometimes an image contains more than 5 hues, and thus the extracted theme cannot fully represent the image color composition, which makes the adjusted result unnatural and incompatible with the desired emotion. Figure 12 shows such a case, where



romantic



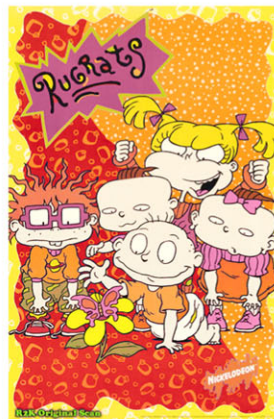
depressed



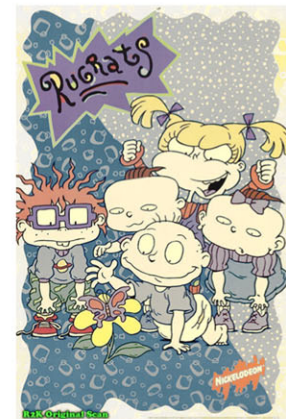
twilight



dim



intense



gentle



deserted



quiet

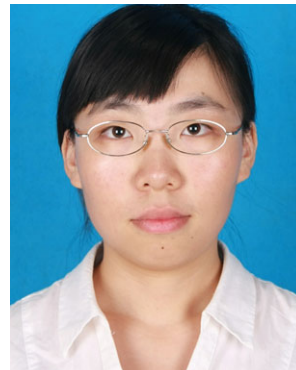
Fig. 13 More results with different affective words

the adjusted image with a *deep* feeling still contains the bright colors. Though such images are unusual and often not good from the artistic perspective, they are still demanded for various reasons. This may be solved using adaptive color theme extraction, and we leave it as a future work.

Acknowledgements This work is supported by the National Basic Research Program (973 Program) of China (2011CB302201), National Natural, and Science Foundation of China (61003094, 60931160443). This work is also funded by Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation, and supported by the Innovation Fund of Tsinghua-Tencent Joint Laboratory.

References

1. Arnheim, R.: *Art and Visual Perception: A Psychology of the Creative Eye*. Univ of California Press, Berkeley (1954)
2. Chen, T., Cheng, M.M., Tan, P., Shamir, A., Hu, S.M.: Sketch2photo: internet image montage. *ACM Trans. Graph.* **28**(5), 124:1–124:10 (2009)
3. Chen, T., Tan, P., Ma, L.Q., Cheng, M.M., Shamir, A., Hu, S.M.: Poseshop: human image database construction and personalized content synthesis. *IEEE Trans. Vis. Comput. Graph.* (2012)
4. Cheng, M.M., Zhang, G.X., Mitra, N.J., Huang, X.L., Hu, S.M.: Global contrast based salient region detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 409–416. IEEE Press, New York (2011)
5. Chia, A.Y.S., Zhuo, S.J., Gupta, R.K., Tai, Y.W., Cho, S.Y., Tan, P., Lin, S.: Semantic colorization with internet images. *ACM Trans. Graph.* **30**(6), 156:1–156:8 (2011)
6. Cohen-Or, D., Sorkine, O., Gal, R., Leyvand, T., Xu, Y.Q.: Color harmonization. *ACM Trans. Graph.* **25**(3), 624–630 (2006)
7. Darwin, C., Ekman, P., Prodger, P.: *The Expression of the Emotions in Man and Animals*. Oxford University Press, New York (2002)
8. Dong, Z.D., Dong, Q.: *HowNet and the Computation of Meaning*. World Scientific, Singapore (2006)
9. Huang, Y.C., Tung, Y.C., Chen, J.C., Wang, S.W., Wu, J.L.: An adaptive edge detection based colorization algorithm and its applications. In: *Proceedings of ACM International Conference on Multimedia, MULTIMEDIA'05*, pp. 351–354. ACM Press, New York (2005)
10. Huang, H., Zhang, L., Zhang, H.C.: Arcimboldo-like collage using internet images. *ACM Trans. Graph.* **30**(6), 155:1–155:8 (2011)
11. Itten, J.: *The Art of Color: the Subjective Experience and Objective Rationale of Color*. Wiley, New York (1974)
12. Kobayashi, S.: *Color Image Scale*. Kodansha International, Tokyo (1991)
13. Kobayashi, S.: *Art of Color Combinations*. Kodansha International, Tokyo (1995)
14. Lawson, C., Hanson, R.: *Solving Least Squares Problems*, vol. 15. Society for Industrial Mathematics, Philadelphia (1995)
15. Lazebnik, S., Schmid, C., Ponce, J.: Affine-invariant local descriptors and neighborhood statistics for texture recognition. In: *Proceedings of 9th IEEE International Conference on Computer Vision*, pp. 649–655. IEEE Press, New York (2003)
16. Levin, A., Lischinski, D., Weiss, Y.: Colorization using optimization. *ACM Trans. Graph.* **23**, 689–694 (2004)
17. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. *IEEE J. Sel. Top. Signal Process.* **3**(2), 236–252 (2009)
18. Li, Y., Ju, T., Hu, S.M.: Instant propagation of sparse edits on images and videos. *Comput. Graph. Forum* **29**(7), 2049–2054 (2010)
19. Lischinski, D., Farbman, Z., Uyttendaele, M., Szeliski, R.: Interactive local adjustment of tonal values. *ACM Trans. Graph.* **25**(3), 646–653 (2006)
20. Liu, T., Yuan, Z.J., Sun, J.D., Wang, J.D., Zheng, N.N., Tang, X.O., Shum, H.Y.: Learning to detect a salient object. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(2), 353–367 (2011)
21. Machajdik, J., Hanbury, A.: Affective image classification using features inspired by psychology and art theory. In: *Proceedings of the International Conference on Multimedia*, pp. 83–92. ACM Press, New York (2010)
22. Matsuda, Y.: *Color Design*. Asakura Shoten, Tokyo (1995) (in Japanese)
23. Mount, D.M., Arya, S.: ANN: a library for approximate nearest neighbor searching. In: *Proc. Center for Geometric Computing Second Ann. Fall Workshop Computational Geometry* (1997)
24. O'Donovan, P., Agarwala, A., Hertzmann, A.: Color compatibility from large datasets. *ACM Trans. Graph.* **30**(4) (2011)
25. Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. *IEEE Comput. Graph. Appl.* **21**(5), 34–41 (2001)
26. Shin, Y., Kim, E.Y.: Affective prediction in photographic images using probabilistic affective model. In: *Proceedings of the ACM International Conference on Image and Video Retrieval*, Xi'an, China, pp. 390–397 (2010)
27. Solli, M., Lenz, R.: Color semantics for image indexing. In: *Proceedings of 5th European Conference on Colour in Graphics, Imaging, and Vision*, pp. 353–358 (2010)
28. Sootitiantawat, S., Auwatanamongkol, S.: Texture classification using an invariant texture representation and a tree matching kernel. *Int. J. Comput. Sci.* **8** (2011)
29. Wang, L., Giesen, J., McDonnell, K., Zolliker, P., Mueller, K.: Color design for illustrative visualization. *IEEE Trans. Vis. Comput. Graph.* **14**(6), 1739–1754 (2008)
30. Wang, B.Y., Yu, Y.Z., Wong, T.T., Chen, C., Xu, Y.Q.: Data-driven image color theme enhancement. *ACM Trans. Graph.* **29**(6), 146:1–146:10 (2010)
31. Wang, B.Y., Yu, Y.Z., Xu, Y.Q.: Example-based image color and tone style enhancement. *ACM Trans. Graph.* **30**(4) 64:1–64:12 (2011)
32. Welsh, T., Ashikhmin, M., Mueller, K.: Transferring color to greyscale images. *ACM Trans. Graph.* **21**(3), 277–280 (2002)
33. Xu, K., Li, Y., Ju, T., Hu, S.M., Liu, T.Q.: Efficient affinity-based edit propagation using k-d tree. *ACM Trans. Graph.* **28**(5), 118:1–118:6 (2009)



Xiaohui Wang Beijing, China. Birthday: May, 1987. She is B.S., graduated from the Department of Computer Science and Technology, Beijing Jiaotong University, China. She is a graduate student of Department of Computer Science and Technology, Tsinghua University. Research interests include image processing and pattern recognition.



Jia Jia Beijing China. Birthday: May, 1981. She is a Ph.D., graduated from Department of Computer Science and Technology, Tsinghua University, China. Her research interests are affective computing, and computational speech perception. She is now an assistant professor of the Department of Computer Science and Technology, Tsinghua University.



Lianhong Cai Beijing, China. Birthday: August, 1945. She received the B.E. degree from Tsinghua University, Beijing, China. She is now a Professor with the Department of Computer Science and Technology, Tsinghua University. She directs the Human-Computer Speech Interaction Laboratory. She has been awarded Scientific Progress Prizes and the Invention Prizes from the Ministry of Mechanism and Electronics, and the Ministry of Education, P.R. China.