

Received April 6, 2020, accepted April 30, 2020, date of publication May 8, 2020, date of current version May 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2993300

Saliency Consistency-Based Image Re-Colorization for Color Blindness

JINJIANG LI¹, XIAOMEI FENG¹, AND HUI FAN¹

¹School of Computer Science and Technology, Shandong Technology and Business University, Yantai 264005, China

²Co-Innovation Center of Shandong Colleges and Universities: Future Intelligent Computing, Yantai 264005, China

Corresponding author: Jinjiang Li (lijinjiang@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61772319, Grant 61976125, Grant 61976124, and Grant 61773244, and in part by the Shandong Natural Science Foundation of China under Grant ZR2017MF049.

ABSTRACT For patients with color vision defects, owing to the destruction of cone cells and loss of function, some color information is lost, hence changing the originally transmitted information by the image. The purpose of the traditional method of correction is to help patients distinguish between colors, but it does not take into account the problem of the saliency of the image. In this paper, we propose a saliency consistency-based image re-colorization for color blindness. We use image retrieval methods to select a large number of images and use co-saliency methods to detect salient areas of standard color images and color-blind simulated images. According to the detection results, an image with the same detection result was selected as the reference image. We grayscale the significantly changed image, recolor the grayscale image using the reference image, and the color scheme of the recolored image is similar to the reference image. The color matching scheme of the reference image makes the significance of the image basically unchanged in standard vision and color vision defects, thereby making the color blindness patient's perception of the image close to standard vision. We invite green blind patients to evaluate CVD simulation images with different recoloring methods subjectively. In addition, we use different evaluation criteria to evaluate the experimental results objectively. In the subjective evaluation and objective evaluation, the method proposed in this paper has achieved good results, which validates the effectiveness of our method.

INDEX TERMS Saliency, image recoloring, co-saliency detection.

I. INTRODUCTION

People with normal eyesight can correctly perceive the frequency of light reflected on the surface of an object. However, patients with color vision deficiency, or colorblindness, cannot perceive colors correctly owing to genetic factors [1] or physical damage. As color vision defects (CVD) are not usually associated with fatal diseases and the number of patients reporting such defects is relatively small, colorblindness has not attracted much attention; therefore, no effective method for curing colorblindness has been developed. The Ishihara test [2] is used to detect color vision defects, which are of three main types: red-green colorblindness, blue-yellow colorblindness, and full colorblindness. Colorblindness can severely affect patients' daily lives, restricting their work-life balance. For example, people with colorblindness cannot get a driver's license. Many professions in related fields, such as engineering and medicine, also have certain requirements for

color perception ability. Although the inability to distinguish colors does not affect people's learning and cognition, working in color-related industries is a challenge for patients with colorblindness. On screen displays of most devices, patients with colorblindness are not fully considered.

Human color vision is based on three light-sensitive pigments [3], [4], which exist in three dimensions. Three types of cone cells—L-cones, M-cones, and S-cones—on the retina determine the correctness of trichromaticity. When any one of the cone cells is damaged or loses function, people can only experience part of the spectrum, necessarily missing spectral information. People suffering from protanopia lack L-cones, making it impossible for them to distinguish between red and green, and those suffering from deuteranopia lack M-cones and cannot distinguish between pale red and pale green. People suffering from tritanopia lack S-cones and cannot distinguish between light yellow and light blue. Approximately 8% of the population worldwide suffers from color vision defects. Therefore, helping colorblind patients perceive color more effectively must be addressed urgently.

The associate editor coordinating the review of this manuscript and approving it for publication was Gangyi Jiang.

One method for improving patients' ability to differentiate color is to recolor the image, while another method is to improve the color contrast through auxiliary equipment tools to achieve differentiation. However, assistive tools need to be tailored, and not every patient can access such tools. From the image perspective, image recoloring has wider applicability. Using the priority of key colors to color the dichromats image [5]. Lin [6] uses the directional feature vector of the color vision defect simulation result to reverse the color distribution. RGB color space is converted to λ , Y-B, R-G color space for recoloring. Poret [7] developed a filter based on the Ishihara test for color correction of colorblind images. Huang [8] proposed a rapid recoloring method based on the best mapping relationship between colorblind images and standard images. Dody [9] RGB color cluster and graph coloring color blind image color correction algorithm. In addition to the color correction of color-blind images, color correction in multi-viewpoint imaging systems has also been extensively studied [10]–[12].

Most recoloring algorithms are performed on standard images, and the colors of the recolored images are rearranged so that patients with colorblindness can distinguish colors [13]. However, one drawback of this color correction method is that the color distribution of a recolored image looks unnatural, such as producing blue apples and cherries. Besides, the purpose of recoloring for color blind images is only to help patients distinguish colors, but it does not consider the problem of saliency changes.

To solve these problems, we propose a saliency consistency-based image re-colorization method. This method uses image retrieval to filter a large number of similar images to form a collection. CVD simulation is performed on the images in the collection, and co-saliency detection is performed on the colorblind images. The detection result is compared with the ground-truth saliency map of the image, and the image with the salient area unchanged is selected as the reference image. The reference image is used to recolor the grayscale image so that the color distribution of the grayscale image is similar to that of the reference image. Because the salient area of the reference image is basically unchanged, this means that the color loss is less during the process of converting the image into a colorblind image; therefore, the colorblind patient's perception of the image is closer to that of a normal person.

The proposed method has the following contributions:

- We propose saliency consistency-based image re-colorization, which not only realizes color discrimination, but also realizes saliency correction of color blind images.
- We use image retrieval technology to collect images. The collected images are similar in semantics and content, so the image coloring effect is better.
- We recolor according to the color scheme of the reference image, instead of randomly changing the color of the image, the color of the recolored image is natural.

II. RELATED WORK

A. COLOR BLINDNESS SIMULATION

The human retina contains two typical types of cells—cone cells and rod cells. Cone cells are active under normal illumination and are mainly concentrated in the depressed area of the retina facing the lens axis, called the fovea centralis. There are 5 to 7 million cone cells in the region, and each cone cell is connected to one optic nerve. Therefore, cone cells can sense brightness and color information and have enhanced resolving power. In areas other than the fovea centralis, there are between 75 and 150 million rod cells. Because rod cells are more sensitive to light, they are mainly active in low light. As many rod cells are connected to one optic nerve, the color of the rod cells is low, and therefore, the perception of color by the eyes mainly depends on the cone cells.

Human color perception is based on three light-sensitive pigments that are presented in a three-dimensional state, and the power of each wavelength specifies the degree of the color stimulus. The trichromatic accuracy is determined by three types of light-sensitive pigments cells (L-, M-, and S-cones) in the retina. Different wavelengths of light stimulate the receptors differently. For example, yellow-green light stimulates the same degree of L-cones and M-cones, but the degree of stimulation of S-cones is weak. The nerve center combines information from various cone cells in response to light of different wavelengths. Therefore, when any cone cell is destroyed or loses function, color vision defects occur.

In order to enable patients with colorblindness to perceive their surroundings properly, colorblind image simulation is needed (see Figure 1). Many colorblindness simulation methods have been carried out, and some rules have been summarized. For example, Brettel [14] used spatial transformation to perform red-green colorblindness simulation, converted the RGB color space to the LSM color space, and performed simulation based on the response of three cone cells. Machado [15] implemented colorblindness simulation based on human color vision theory. Based on the electrophysiological study, a simulation matrix was developed. For normal color vision, three-color vision and two-color vision can be processed. Okajima [16] used a personal color vision model to improve the brightness difference. Rivera [17] and coworkers used a convolutional neural network, taking Ishihara printing plates as input and adjusting the brightness through the network to achieve real-time colorblind simulation. Aytac [18] designed an integrated colorblind analog user interface. In addition, ready-made applications such as ColorDoctor are also available.

B. IMAGE RETRIEVAL

With the development of screen display technology, users can view many different types of images and videos. In order to achieve effective retrieval of multimedia data systems, image retrieval was developed, enabling users to quickly and accurately retrieve images according to their needs.

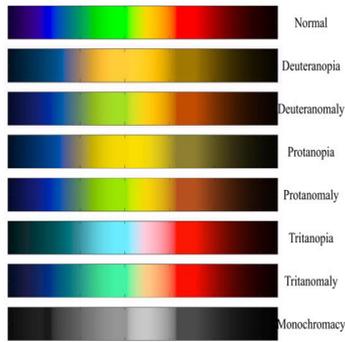


FIGURE 1. Top: Color spectrum that non-colorblind people can see (standard vision). Below: The three primary types of colorblindness paired with their “less-severe” versions.

With the development of deep learning, image retrieval technology has achieved significant development [19] and many applications such as product search [20], face recognition [21], and image geolocation [22] have been developed. Beginning with the text-based image retrieval (TBIR) process, such as the Page-Rank [23] method, image retrieval has undergone a complex development process. Although the TBIR method can locate images quickly and accurately, it also has some disadvantages, such as time-consuming and labor-intensive manual labeling, strong subjectivity of text, and insufficient object description of the image. By 2000, content-based image retrieval (CBIR) technology appeared and mathematical description was performed through low-level features such as color, texture, and shape in the image content. By calculating the similarity of features, the retrieved images are filtered to obtain the image that best meets the needs of the user. Such as the dictionary learning algorithm, FV algorithm, VLAD algorithm mentioned in [24]. In order to overcome the lack of simple visual features, semantic-based image retrieval (SBIR) is proposed. SBIR includes natural language processing and traditional image retrieval techniques. Using of convolutional neural networks to comprehensively consider the low-level and high-level features of images bridges the semantic gap. An image retrieval method based on capsule network is given in [25].

The image retrieval process has developed from text to visual content to semantic-based retrieval wherein the idea is to extract the relevant attributes [26], relative attributes [27], and absolute attributes [28] of images and compare the features to retrieve the desired image.

C. SALIENCY DETECTION

Development of multimedia technology has resulted in a large number of images with varying quality. The key issue is to quickly filter useful images and delete unnecessary images. Computers can be used to simulate human vision, extract a small amount of significant information from a large amount of image information for analysis, and help people to solve detection problems. Currently, visual saliency research has

become a topic of great interest in the field of multimedia technology—the goal of which is to simulate human visual perception and attention mechanism and to realize automatic prediction, localization, and extraction of important information. Saliency detection involves cross application of psychology, neuroscience, and computer vision [29] and can predict the position of the gaze point of the human eye and detect a salient object.

With the development of deep learning, attention models are introduced into neural networks to mimic human attention mechanisms in visual scenes. A recurrent attention model with hard alignment was proposed by Mnih [30] and coworkers, but model training was more difficult. Later, the image title of the circular attention model to align the words with the image area was studied in [31]. Sermanet [32] used recursive attention model for fine-grained classification by focusing on discriminant regions. In addition, attention models are applied to computer vision [33], [34], which effectively resolve computer vision problems by analyzing context information in the aforementioned applications.

The saliency detection model relies on the local contrast [35], global contrast [36], and background prior [37] cues to detect salient objects. In addition, CNN has achieved great success in saliency detection. In [38], the FCN-based saliency model is combined with the saliency model of multi-scale image regions. Wang [39] repeatedly used FCN to refine the saliency map. Amulet network [40] was used to combine multiscale context information. Hu [41] used guided filtering to refine the saliency map.

In the face of complex images, frequently occurring patterns, or main foregrounds, are used to characterize the main content of the images and batch processing of the images is realized, which promotes the development of co-saliency detection methods. Co-saliency finds the common saliency regions from multiple related images. The image collection can be divided into four parts: common foreground (CF), common background (CB), uncommon foreground (unCF), and uncommon background (unCB). The co-saliency method is divided into three steps: extracting effective features, using information clues to characterize saliency, and calculating common saliency areas. Among the effective feature extractions are mainly the low-level features such as color histograms, Gabor filters, or SIFT descriptors [42]. In addition, there are methods to take the results of the saliency regions in a single image as intermediate features [43] and detect the co-saliency regions in the entire image collection.

D. GRAYSCALE IMAGE COLORIZATION

Grayscale image colorization is the process of recoloring a grayscale image to enhance the color perception of the image [44], and it is widely used in the fields of refurbishment of black-and-white photos and recoloring of medical images. The colorization method can be divided into three categories: user-assisted coloring, automatic coloring, and grayscale coloring based on reference images.

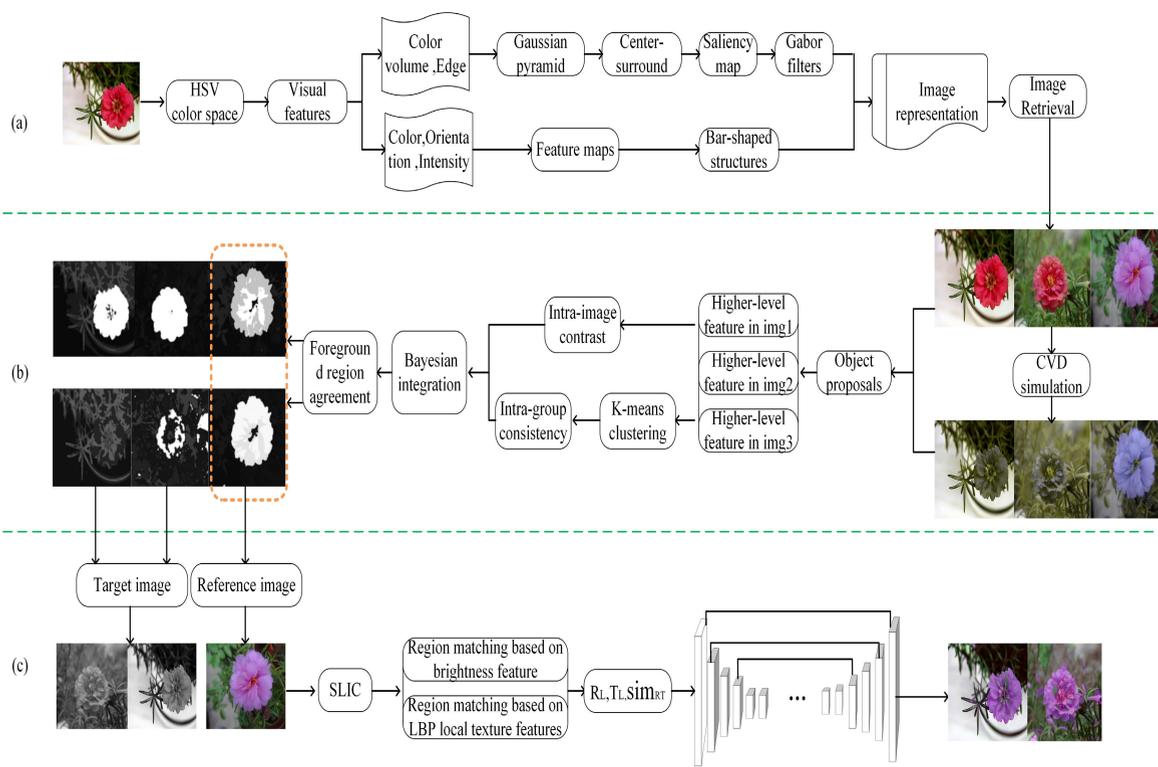


FIGURE 2. Proposed framework for our method including three parts: (a) Generate image collection through image retrieval; (b) Co-saliency method to detect the saliency map of the original image and the color-blind image; (c) Saliency changed images are recolored.

The user-assisted shading method involves manually marking the color [45] or the area where the color is migrated [46] by the user, using the least square method to optimize it based on the relationship between the colored pixels and neighboring pixels, and finally, realizing the colorization of the global image. However, the process of labeling samples is time-consuming and laborious and the coloring effect depends heavily on the color of the label. To simplify the coloring process, a reference image-based coloring method is proposed. Grayscale the target image and recolor it using the reference image. As per the method given in [47], find the pixel points that match the target image in the reference image and recolor the target image with reference to the color distribution in the reference image. The effect of recoloring is different due to use of the different color migration methods.

The purpose of automatic coloring is to realize automatic color migration through machine learning or statistical learning methods, such as histogram statistics [48], and multi-model prediction methods [49]. An increasing number of people are using convolutional neural networks to learn prior information from large-scale data to achieve automatic coloring. However, coloring is an ill-conditioned problem with multimode uncertainty. Therefore, using automatic image coloring may result in problems, such as out-of-bounds color and blurring in the edge area.

Reference-based image coloring provides a basis for coloring by inputting a reference image similar to the content

of a grayscale image. Bugeau [50] super-pixel the image and determines the similarity between the target image and the reference image based on pixel-level attributes. However, when there is a difference between the target image and the reference image, coloring based on the similarity between features is more prone to make mistakes.

III. PROPOSED METHOD

Due to the destruction of cone cells, patients with colorblindness cannot recognize certain colors. If the color scheme of an image contains colors that the colorblind patient cannot recognize, then even after conversion of the image to a colorblind image, the CVD patient still cannot effectively perceive the image. Because the color scheme of the image is different, the loss of color in the conversion is also different. In order to make the perception of images of a colorblind patient as close as possible to that of a normal person, a color scheme with less color loss should be chosen. The algorithm pipeline is shown in Figure 2. First, search for images with similar semantics based on image retrieval techniques (see III. A). Then, conduct co-saliency detection on semantically similar images (see III. B) and select a suitable reference image according to the saliency map of the colorblind image and the standard image. Finally, the reference image is used to color the target image (see III. C) so that the color distribution of the colored image is similar to that of the reference image.



FIGURE 3. Image retrieval.

A. CONTENT-BASED IMAGE RETRIEVAL

Extracting meaningful features from an image is difficult in content-based image retrieval. The human visual system has directional selectivity for visual attributes such as color, direction, and intensity information; therefore, integration of effective information throughout an image is a challenging problem. In this paper, the saliency structure model is used, and the saliency histogram is used to represent the effective features of the image through the saliency area as a visual clue. The image retrieval is shown in Figure 3.

It has been experimentally proven [35] that the human visual system is more sensitive to color, direction, and intensity. In order to briefly represent the main features of an image, feature quantization is required in the HSV space and limited colors are selected to describe the true color image to the greatest extent [51]. In order to obtain the color map, the H, S, and V color channels are uniformly quantized into 6, 3, and 3 bins so that a total of $6 \times 3 \times 3 = 54$ color combinations can be obtained. Using $M_C(x, y) = w, w \in \{0, 1, \dots, N_C - 1\}$ for color combinations, where $N_C = 54$. The intensity information is represented by V. After quantization, the intensity information is represented as $M_I(x, y) = s, s \in \{0, 1, \dots, N_I - 1\}$, where $N_I = 16$. In addition, edge information $g(x, y)$ is detected using intensity information $O(x, y)$ and the Sobel operator is used to detect the image. After uniform quantization, the edge pattern $M_O(x, y) = \theta, \theta \in \{0, 1, \dots, N_O - 1\}$, of the image is obtained, where $N_O = 60$.

The main visual features of images are extracted in the HSV space, but not all features are effective. Based on this situation, a color volume is defined in the HSV space to describe the salient features. Since the shape of the HSV space can be simulated as a cylindrical coordinate system, for the color volume of a random point (h, s, v) , the formula based on the volume of the cylinder can be defined as follows:

$$cv_1(x, y) = \pi \times s(x, y)^2 \times v(x, y) \times \frac{h(x, y)}{360}, \quad (1)$$

where $s(x, y) \in [0, 1], v(x, y) \in [0, 1], h(x, y) \in [0, 360]$.

Convert the three-dimensional (3-D) coordinate system to the two-dimensional (2-D) coordinate system and define the color volume of the point (h, s, v) as follows:

$$cv_2(x, y) = s(x, y) \times \cos(h(x, y)) \times s(x, y) \times \sin(h(x, y)) \times v(x, y), \quad (2)$$

Combine the 3-D coordinates with the 2-D coordinate information to define $cv = \{cv_1, cv_2\}$. Gaussian pyramids $cv(\sigma)$ and $g(\sigma)$ are generated using cv and edge information $g(x, y)$, with scale $\sigma = [0, 1, 2, 3, 4]$. Gaussian pyramid information is used to represent the center-surround receptive fields, as follows:

$$F(c, s, cv) = |cv(c) \ominus cv(s)| \forall cv \in \{cv_1, cv_2\}, \quad (3)$$

$$F(c, s, cv) = |g(c) \ominus g(s)|, \quad (4)$$

where c, s represents two different scales and \ominus represents the subtraction operation.

The feature map is obtained after the center surround, and the 2-D Gabor function is used for modeling. The model uses Gabor energy, selects the appropriate direction to detect the salient structure, and describes the image features. If the direction of the feature is the same as the direction of the Gabor, the feature is considered to be a significant structure. To avoid the repetitiveness of the significant structure space, the content-based saliency structure histogram (SSH) [52] is used for retrieval.

B. CO-SALIENCY DETECTION

Co-saliency detection is used to identify common salient objects in a group of related images. It is also widely used in image retrieval, image annotation, and video foreground extraction. To effectively extract the salient features of an image, not only the images in the collection are used, but also the visually similar images outside the collection are used. First, use the BING [53] method to extract 256 object proposal (OP) windows in the image and then use CNN with Restricted Boltzmann Machines (RBMs) to build higher level features [54]. In order to consider information more widely, use K-means to classify $\{x_{m,p}\}_{p=1}^{256}$ as $\{C^k\}_{k=1}^K$, where the center of each class C^k is given by $\{c^k\}_{k=1}^K$, and calculate the consistency between images as follows:

$$\Pr(x_{m,p} | y_{m,p} = 1) = \frac{1}{Z_m^{IMC}} \exp\left(\frac{-Ed(f_{m,p}, \hat{c})^2}{\sigma^2}\right), \quad (5)$$

$$Z_m^{IMC} = \sum_{p=1}^M \sum_{q=1, p \neq q}^{256} \exp \left(\frac{-Ed(f_{m,p}, \hat{c})^2}{\sigma^2} \right), \quad (6)$$

where $Ed(\cdot)$ is the Euclidean distance.

Use visually similar neighborhoods in image groups to expand the field of view. If the neighborhood backgrounds are similar, the risk of significant area division errors can be reduced and the purpose of suppressing background areas can be achieved.

To find significant features, the features are screened by scoring. A Bayesian frame was used to score the consistency between the contrast in the graph and the contrast in the group and judge the significance. We define the significance of $x_{m,p}$ as follows:

$$\begin{aligned} \text{Cosal}(x_{m,p}) &= \Pr(y_{m,p} = 1 | x_{m,p}) \\ &= \frac{\Pr(x_{m,p} | y_{m,p} = 1) \Pr(y_{m,p} = 1)}{\Pr(x_{m,p})} \\ &\propto \frac{1}{\Pr(x_{m,p})} \Pr(x_{m,p} | y_{m,p} = 1), \end{aligned} \quad (7)$$

Among them, use $\{x_{m,p}\}_{p=1}^{256}$ to represent OPs of I_m in the current image group and the 2-D random variable $y_{m,p}$ to indicate whether $x_{m,p}$ is a feature in the co-dominant region. $\frac{1}{\Pr(x_{m,p})}$ represents the contrast inside the image. In information theory, the logarithmic form of $\frac{1}{\Pr(x_{m,p})}$ is $-\log \Pr(x_{m,p})$ —the self-information of the random variable [55]. When the probability of OPs decreases and the self-information increases, it is more likely to be the foreground object. $\Pr(x_{m,p} | y_{m,p} = 1)$ stands for consistency between images.

To obtain a clear boundary, the co-saliency scores of OPs need to be converted into pixel-level saliency maps. According to the method in [56], a foreground region agreement (FRA) is proposed and applied to two phases of intra-image and inter-image. For the FRA in the figure, the image is super-pixel, and the image I_m is represented as $\{sp_i\}_{i=1}^{N_m}$, where sp_i is a single pixel and N_m is the number of super-pixels in the image. Using the image classification pooling method, the co-saliency score of a single pixel sp_i is equal to the sum of the co-saliency scores of all the pixels in the OPs, given by:

$$\text{Cosal}^{rgh}(sp_i) = \sum_{p=1}^{256} \text{Cosal}(x_{m,p}) B(sp_i, x_{m,p}), \quad (8)$$

$$B(sp_i, x_{m,p}) = \begin{cases} 1, & \text{Area}(sp_i \cap x_{m,p}) > \frac{1}{2} \text{Area}(sp_i) \\ 0, & \text{otherwise,} \end{cases} \quad (9)$$

where $\text{Cosal}^{rgh}(sp_i)$ represents the pixel value of a single pixel sp_i and $\text{Area}(\cdot)$ represents the area.

Based on the score of a single pixel, set an adaptive threshold and select the foreground pixels. After obtaining the intra-image FPA, use the popular sort algorithm [57] for the

obtained foreground pixels to calculate the co-saliency score of each pixel.

For the pixel sp_i of the intra-group FRA in I_m , calculate the Euclidean distance between the features of the adjacent pixels and obtain the most similar pixel $\{sp_n\}_{n=1}^{(m-1)N}$, the number of which is N , shown as follows:

$$\text{Cosal}^{FRA_2}(sp_i) = \varphi(sp_i) \cdot \exp(-\phi(sp_i)), \quad (10)$$

where $\varphi(sp_i)$ represents a super-pixel set similar to sp_i and $\exp(-\phi(sp_i))$ represents the similarity between super-pixels. According to the consistency of the super-pixels between the images, the saliency nodes in the image and the image are combined to obtain the saliency map in the image group.

C. IMAGE COLORING

A grayscale image removes the color information of an image, leaving behind only the brightness information, not enough to colorize the image using only gray information. A gray image can be colorized by using prior knowledge or by manually setting rules. Based on the reference image coloring, find the similarity between the target image and the reference image and color by pixel matching.

For different images, due to the pixel mismatching between the target image and the reference image, the shading effect is poor. In this paper, we use support vector machine (SVM) to calculate the initial class probability and perform the initial classification. To ensure the accuracy of the classification, the image is subjected to simple linear iterative clustering (SLIC) to improve the consistency of the classification space. Pixels with similar brightness, color, and texture constitute a super-pixel region, which eliminates duplicate information in the image and improves the processing speed of the image. Through brightness and texture features, combined with classification information, pixel matching is performed on the reference image and the target image. The pixels are colored according to the pixel matching result, and thus, the grayscale image is finally colored.

Owing to some differences in the brightness information between the target image and the reference image, direct region matching does not give satisfactory results; therefore, the pixels are classified. First, the reference image is remapped so that its brightness value is closer to that of the target image, and then area matching is performed. The brightness remapping formula is given in equation (11). Features used for region matching include brightness features and LBP for texture matching, where the mean brightness can better reflect the global information of the image.

$$\bar{Y}_S = \frac{\sigma_T}{\sigma_S} (Y_S - \mu_S) + \mu_T, \quad (11)$$

where \bar{Y}_S and Y_S are the brightness values of the reference image pixels after remapping and before remapping, respectively. μ_T and μ_S are the average brightness values of the target image and the reference image, respectively; and σ_T , σ_S are the standard deviations of the target image and the reference image, respectively.

The matching information of the target image and the reference image is preprocessed through the brightness information to obtain a set of reference image subregions that have similar brightness characteristics as that of the target image subregions. To make the matching more accurate, the local matching (LBP) is used to complete the entire matching process in two image subregions with similar brightness characteristics. The LBP operator is a description of the local neighborhood texture in an image. Any pixel point $g(x_c, y_c)$ in a certain neighborhood in the image, g_c is the gray value of the center point of the local neighborhood of the image, and the local texture $T_c = (g_c, g_0, \dots, g_{p-1})$, $c \in \Omega$, g_p is the gray value of the pixel points p evenly distributed in the neighborhood, such as the equation (12).

$$g_p(x_c, y_c) = \left(x_c + R \cos\left(\frac{2\pi p}{P}\right), y_c - R \sin\left(\frac{2\pi p}{P}\right) \right), \quad (12)$$

The LBP value of the local area centered on according to is given as follows:

$$LBP_{P,R} = 2^p \times \sum_{p=0}^{P-1} s(g_p - g_c), \quad (13)$$

where $P = 8$ and $R = 1$.

According to the brightness feature R_L, T_L and the LBP texture feature of the reference image and the target image, the similarity $sim_{T \leftrightarrow R}$ between the color reference image R and the target grayscale image T is obtained. After obtaining the similarity between the images, the target grayscale image is colored according to the similarity between the target image and the reference image. When there is no similar relationship between the target image and the reference image, how to color the area? This paper refers to the method in [58] and uses an end-to-end network to color the grayscale image. For the chroma branch, consider $T_L, sim_{T \leftrightarrow R}, T'_{qp}$ as the input of the network and output the colored result P'_{ab} , where $T'_{ab}(p) = T_{ab}(\phi_{R \rightarrow T}(\phi_{T \rightarrow R}(P)))$ is the result of reference to the chromaticity of the ground-truth image. Calculation of loss L_1 is performed on the network output result P'_{ab} and the ground-truth chroma value T_{ab} , making the image effects more realistic after coloring.

$$L_{chrom}(P'_{ab}) = \sum_p smooth_{L_1}(P'_{ab}(p), T_{ab}(p)), \quad (14)$$

where p represents a pixel of the image.

For the part where the similar region cannot be found in the reference image, the perceptual loss function is used for training, shown as follows:

$$L_{perc}(P_{ab}) = \sum_p \|F_P(p) - F_T(p)\|^2, \quad (15)$$

where $F_P(P)$ represents the features in the original image P_{Lab} and $F_T(P)$ represents T_{Lab} of the target image.

The traditional grayscale image colorization method has a better coloring scheme for the matching area and a poor coloring effect for the nonmatching area. In this paper, the

TABLE 1. Analysis of co-saliency detection results of colorblind images in Example 1.

| Name Label | AUC | AP | F-measure |
|------------|--------------|--------------|--------------|
| 1 | 0.608 | 0.693 | 0.111 |
| 2 | 0.549 | 0.531 | 0.305 |
| 3 | 0.484 | 0.638 | 0.633 |
| 4 | 0.699 | 0.677 | 0.861 |
| 5 | 0.547 | 0.615 | 0.204 |
| 6 | 0.764 | 0.634 | 0.986 |
| 7 | 0.674 | 0.579 | 0.221 |

perceptual loss function is used to train the nonmatching area to effectively color it.

IV. EXPERIMENT

In this section, we first introduce some details of the algorithm implementation (see IV. A). Afterwards, the experimental process is introduced through several examples, and the experimental results are analyzed (see IV. B). Afterwards, we show a quantitative comparison between the detection result of the recolored image and the detection result of the unprocessed image (see IV. C). Finally, we compare our method with other methods and perform the human subjective evaluation (see IV. D).

A. IMPLEMENTATION DETAILS

During the experiment, image retrieval technology was used to establish a suitable image collection. The computer used in the experiment was configured with i7 3.6 GHz CPU, 16 GB RAM and a GTX 980 GPU, and compilation software MATLAB R2018a-64bit.

We use area under curve (AUC), average precision (AP), F-measure, Mean absolute error (MAE) and Root Mean Squard (RMS) to evaluate co-saliency detection of different image. The AUC is calculated by calculating the area under the ROC curve. AP is the area of the siege of the precision-recall curve and the two axes. PR curve cannot be used to evaluate the saliency picture more comprehensively in many cases. So people use the non-negative parameter β to weigh precision and recall. We set β to 0.3 to emphasize precision.

B. RESULTS AND ANALYSIS

To demonstrate the effectiveness of the proposed method, several images were selected as experimental objects. Example 1, Example 2, and Example 3 were, respectively, selected as the experiments on flowers, puppets, and natural landscapes. While conducting the experiments, the saliency detection was first performed on the standard image and the colorblind image, and the detection result was analyzed. Then, an image with less significant change was selected as a reference image. After graying out the remaining images, the reference image was used to recolor the grayscale image. In this case, the color distribution of the recolored image was more consistent with the reference image; thus, the saliency

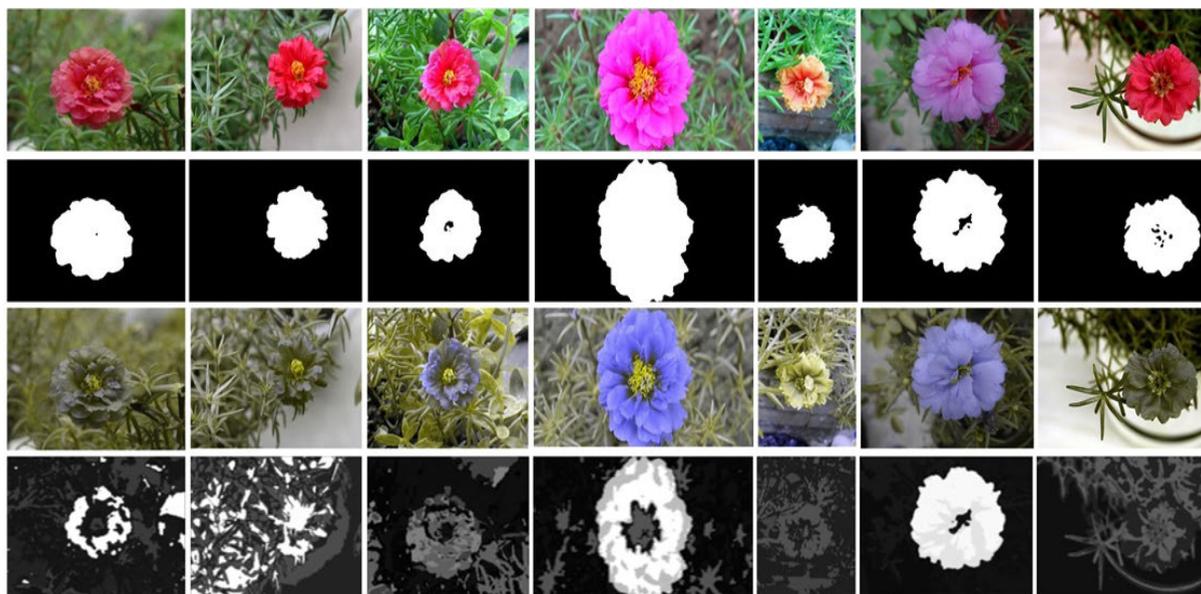


FIGURE 4. Co-saliency detection results of the original image and the colorblind image in Example.

of the image was better retained, and the purpose of saliency correction of the colorblind image is achieved.

1) EXAMPLE 1

The saliency detection results of the original image and the colorblind image of Example 1 are shown in Figure 4, and the images in Example 1 are numbered 1–7 from left to right. The first row is the original image. The second row is the result of the saliency detection in the first row. The third row is the result of the first line image CVD simulation. The fourth row is the result of saliency detection on the third row of the image. It can be seen from the Figure 4 that after a normal image is converted to a colorblind image, the salient areas of some images are reduced, and even in some images, they have changed because with the change in the color distribution of the image, there is a chance of change in the saliency. However, there are also few parts of the image that have not changed in saliency but have been well preserved. This shows that the color-matching scheme of this image has less loss of color information during colorblindness simulation and retains the original state of the image to a large extent, which is suitable for colorblind viewing, because the colorblind patient’s perception of this image is closer to the normal vision.

The saliency detection results of colorblind images and normal images were evaluated by AUC, AP and F-measure values. The higher the AUC score, AP score and F-measure, the more accurate the significant area detected. The analysis results of the color blind image in Example 1 are shown in Table 1, which comprehensively shows the AUC, AP, and F-measure values. Name label6 provides the better performance among all images, with AUC score of 0.764, AP score of 0.634, and F-measure score of 0.986. Meanwhile, name

label 4 achieves AP score of 0.677, and outperforms Name label6 by 0.043. But AUC and F-measure scores of name label 4 are lower than name label6. We compare the saliency maps of the color-blind images name label4 and name label6, and finally decide name label6 as the reference image.

The reference images were used to color the significantly changed target image. The name label images 1, 2, 3, 4, 5, and 7 in Example 1 were gray-scaled, as shown in the first line of Figure 5. The second line is the result of coloring the gray image in the first line. It can be seen from the figure that the color scheme of the image after coloring is mainly concentrated in purple and green, similar to the reference image. The third row is a green blind simulation of the recolored image. The fourth row is the saliency detection of the image after colorblindness simulation. The fifth row is a true saliency map of the name label images 1, 2, 3, 4, 5, and 7.

Figure 5 shows some qualitative results. Among them, it is obvious that the proposed method can better extract the co-saliency regions in different image. In unprocessed images, the leaves in the background shares the same color with the flowers, making the co-saliency detection methods fail to fully detect the flowers, but our method achieve satisfying results because it uses recolorization of target images that can well discriminate target from background. Analysis of the fourth and the fifth lines shows that the saliency of a colorblind image after color correction is closer to the ground-truth image.

2) EXAMPLE 2

In Example 1, experiments are performed on single background image. In this experiment, experiments are performed on complex background image. In Example 2, the image of a teddy bear wearing a red top was used as the experimental

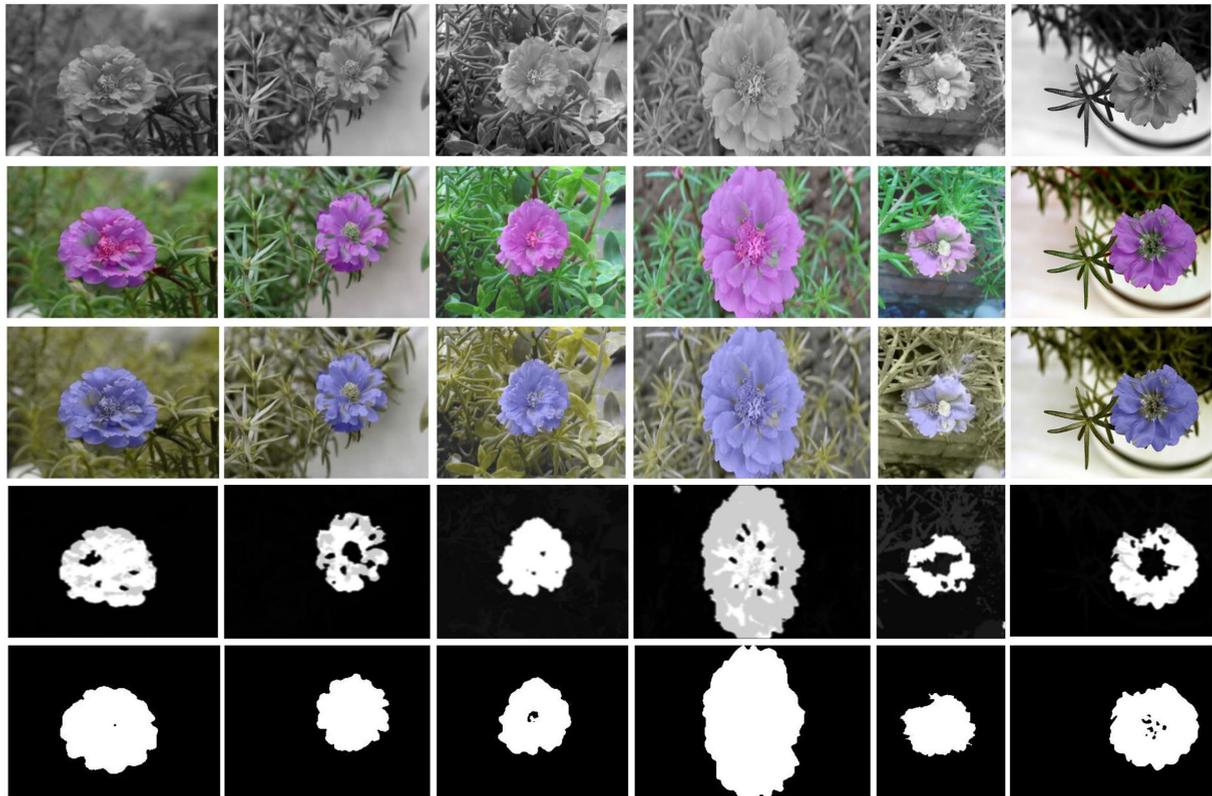


FIGURE 5. Recolorization of the target image using the reference image in Example 1.

object, as shown in Figure 6. The first row is the original image. The second row is the result of the saliency detection in the first row. The third row is the result of the first line image CVD simulation. The fourth row is the result of saliency detection on the third row of the image. However, a colorblind patient whose vision is insensitive to red color cannot effectively identify the color. The third line in Figure 6 is the result of green blindness simulation of the original image, and the fourth line is the result of saliency detection of the simulated image. Comparing the fourth line with the second line shows a change in the salient area. After CVD simulation of the image, the vivid red in the original image is no longer significant but is replaced by the yellow part. But for the whole yellow bear, it can accurately detect the saliency area of the image.

Analyze the saliency detection results of the colorblind image in Example 2, as shown in Table 2. Name label7 provides the best performance among all images, with AUC score of 0.875, AP score of 0.756, and F-measure score of 0.669. Considering and analyzing the detection results comprehensively, choose name label 7 as the reference image and recolor the images with significant changes.

Grayscale the name label images 1, 2, 3, 4, 5, and 6 in Example 2 and recolor them using the reference image name label 7. The coloring results are shown in Figure 7. The second line shows the result of coloring the gray image in the first row. It can be seen from the figure that the color scheme

TABLE 2. Analysis of co-saliency detection results of colorblind images in Example 2.

| Name Label | AUC | AP | F-measure |
|------------|--------------|--------------|--------------|
| 1 | 0.654 | 0.685 | 0.543 |
| 2 | 0.603 | 0.655 | 0.515 |
| 3 | 0.616 | 0.624 | 0.436 |
| 4 | 0.541 | 0.539 | 0.442 |
| 5 | 0.576 | 0.618 | 0.584 |
| 6 | 0.636 | 0.607 | 0.249 |
| 7 | 0.875 | 0.756 | 0.669 |

of the image after coloring is mainly concentrated in yellow, which is similar to the color scheme of the reference image. The third row shows a green blind simulation of the recolored image. The fourth row shows the saliency detection of the image after colorblindness simulation. The fifth row shows the ground-truth image of a saliency image of each image.

3) EXAMPLE 3

In Example 3, difficult natural landscape images were used as experimental objects, such as red maple leaves, pink cherry blossoms, and purple lavender. In Figure 8, the standard image was subjected to green blind simulation. The first row is the original image. The second row is the result of the saliency detection in the first row. The third row is the result of the first line image CVD simulation. The fourth row is



FIGURE 6. Co-saliency detection results of the original image and the colorblind image in Example 2.

TABLE 3. Analysis of the co-saliency detection results of the colorblind images in Example 3.

| Name Label | AUC | AP | F-measure |
|------------|--------------|--------------|--------------|
| 1 | 0.608 | 0.512 | 0.473 |
| 2 | 0.706 | 0.683 | 0.416 |
| 3 | 0.721 | 0.548 | 0.397 |
| 4 | 0.446 | 0.614 | 0.005 |
| 5 | 0.731 | 0.513 | 0.252 |
| 6 | 0.722 | 0.563 | 0.573 |
| 7 | 0.815 | 0.758 | 0.568 |

the result of saliency detection on the third row of the image. The simulation result is shown in the third line of the figure. The fourth line shows the result of saliency detection on the simulated image.

The saliency detection results of the colorblind image in Example 3, as given in Table 3. Name label7 provides the better performance among all images, with AUC score of 0.815, AP score of 0.758, and F-measure score of 0.568. Meanwhile, name label 6 provides the best performance on F-measure. Name label7 AUC score and AP score outperforms name label 6 by 12.8% and 34.6%, respectively. After analyzing the data in the table, the name label 7 with better saliency was selected as the reference image and correct the rest of the images with saliency changes.

Using the reference image name label 7 to recolor the remaining images in Example 3, the coloring results are

shown in Figure 9. The first line shows the result of graying out the target image. The second line shows the result of coloring the gray image in the first line. It can be seen from the figure that the color scheme of the colored image is mainly concentrated in yellow and green. The third line shows the result of green blind simulation of the images in the second line. The fourth line shows the result of saliency detection of the images in the third line. The fifth line shows the ground-truth image of the saliency maps of the name label images 1, 2, 3, 4, 5, and 6. These results verify the effectiveness of our saliency correction for color blind image.

C. COMPARISONS WITH UNPROCESSED IMAGE

In the previous section, the conducted qualitative analysis of the experimental results is detailed. The detection result after converting the original image into a colorblind image is compared with the saliency detection result of the recolored image by RMS, MAE and F-measure. For example, Figure 10 compares the RMS values of the two. As can be seen from the figure, the RMS value of the image after color correction is significantly reduced. This shows that the saliency detection result of the recolored image is more accurate, proving the effectiveness of the method proposed in this paper.

In addition, the MAE values of the recolored and unprocessed images were also compared. The comparison results are shown in Figure 11. It can be seen from the figure that the MAE curve of the recolored image is wrapped by the curve of the unprocessed image and the MAE value of the recolored

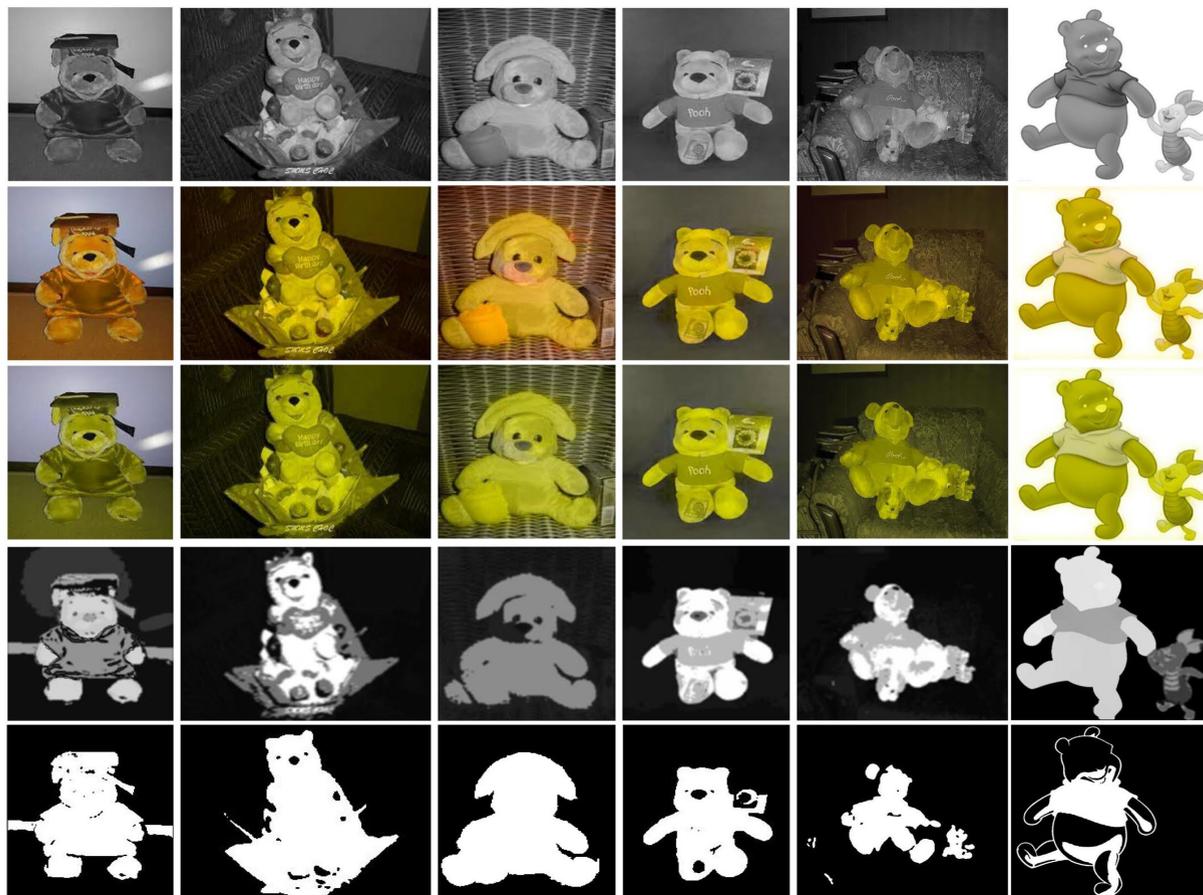


FIGURE 7. Recolorization of the target image using the reference image in Example 2.

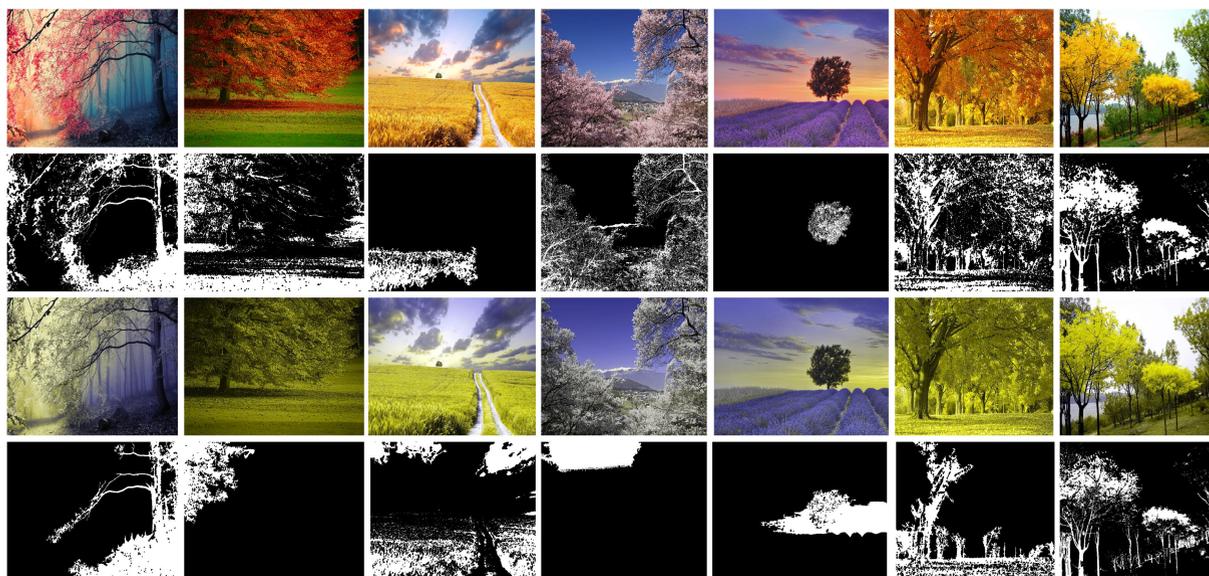


FIGURE 8. Co-saliency detection results of the original image and the colorblind image in Example 3.

image is significantly reduced; hence, the detection result is more accurate.

To analyze the experimental results more comprehensively, the F-measure values of the recolored image and the unpro-

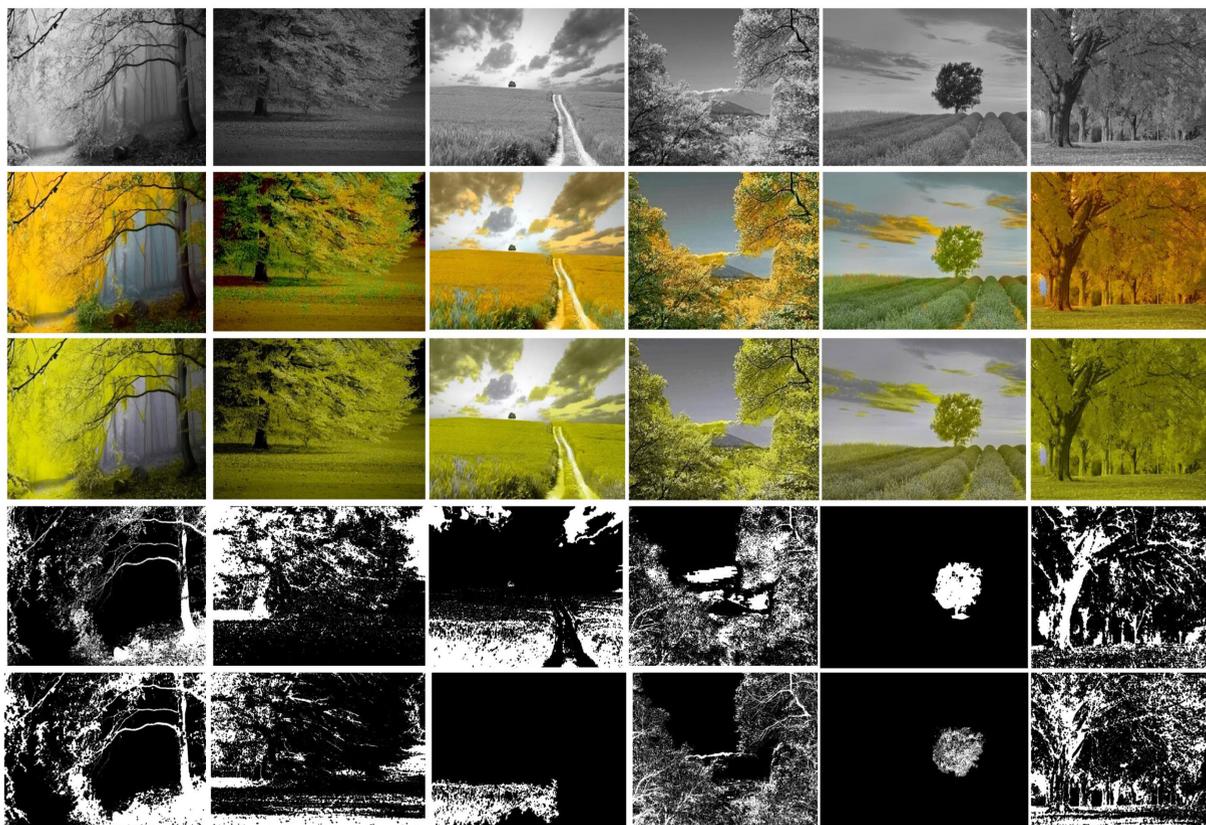


FIGURE 9. Recolorization of the target image using the reference image in Example 3.

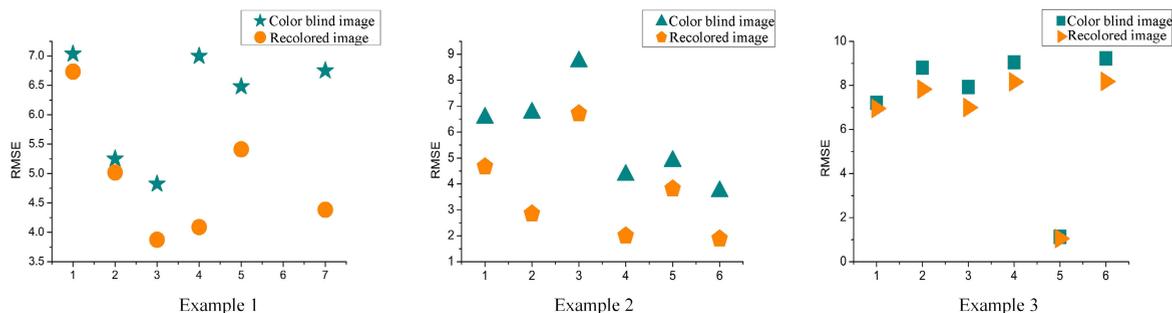


FIGURE 10. Comparison of the RMS value of the recolored image to that the unprocessed image. The color blind image represents the CVD simulation image detection result of the original image. The recolored image represents the detection results of the CVD simulation image processed by our method.

cessed image are compared, such as the histogram shown in Figure 12. It can be seen from the Figure 12 that the F-measure values of the recolored images are higher than those of the unprocessed images. After an image is recolored, the saliency detection result of the image is more accurate, thereby reaching the goal of saliency correction of the colorblind image. Through the above three evaluation metrics, test the effectiveness and feasibility of the method proposed in this article.

D. HUMAN SUBJECTIVE EVALUATION

Color-blind image correction should follow some principles, not just to distinguish colors for correction. For example,

to make the patients better distinguish the colors, the oranges are recolored to blue. But the perception experience of patients with color vision deficiency tells themselves that oranges are not a blue fruit. Therefore, we should maintain perceptual learning while performing color correction[12]. Perceptual learning consists of the following parts: 1) Color naturalness. We want to make the perceived difference between the recolored image and the source image as small as possible so that the color of the image is natural and not obtrusive. 2) Colour consistency. We hope that the color mapping of the recolored image is regular, and the color distribution of the recolored image can be predicted. 3) Colour contrast. The contrast of different colors in the

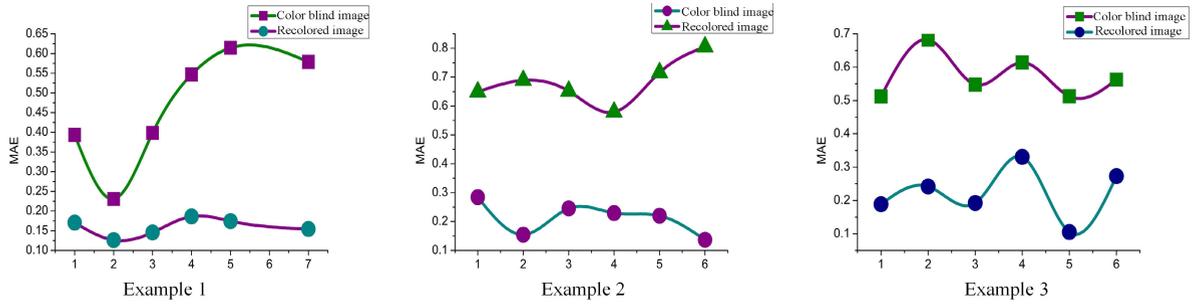


FIGURE 11. Comparison of the MAE value of the recolored image to that of the unprocessed image. The color blind image represents the CVD simulation image detection result of the original image. The recolored image represents the detection results of the CVD simulation image processed by our method.

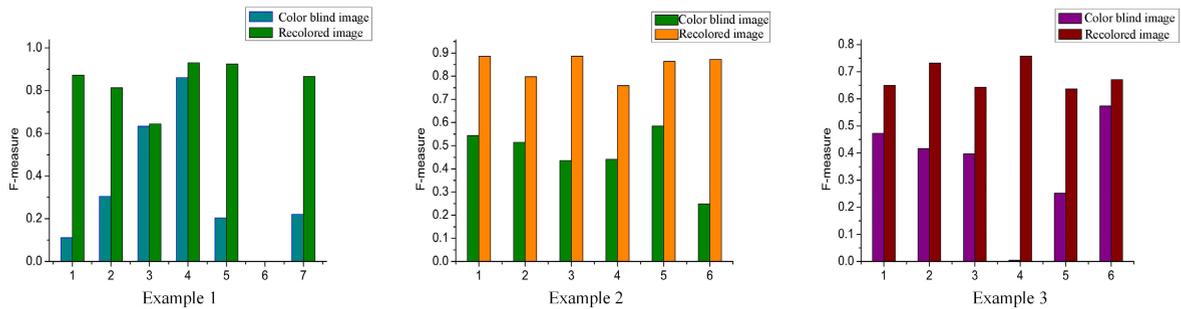


FIGURE 12. Comparison of the F-measure value of the recolored image to that of the unprocessed image. The color blind image represents the CVD simulation image detection result of the original image. The recolored image represents the detection results of the CVD simulation image processed by our method.

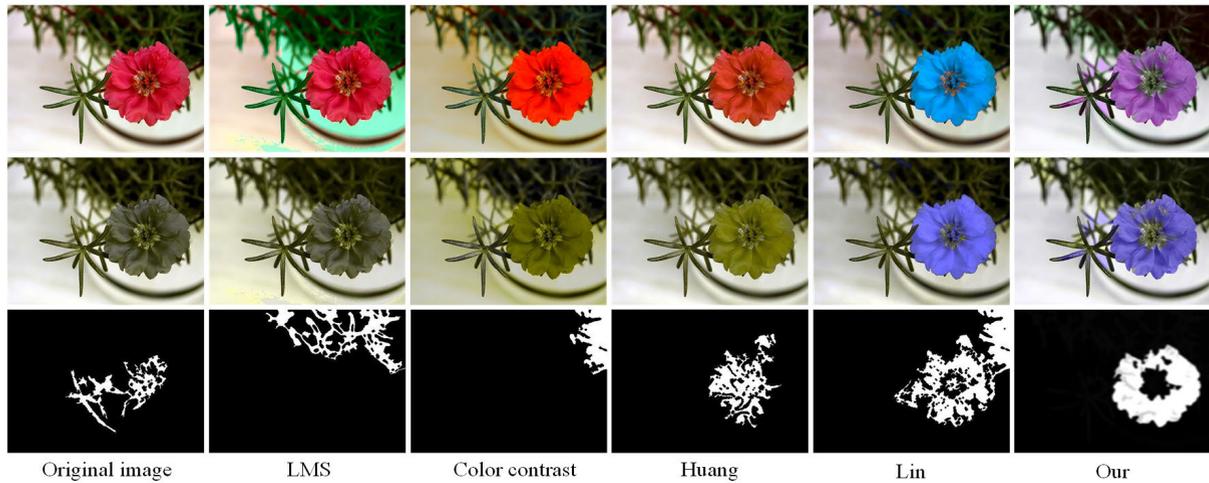


FIGURE 13. Example of comparative experiment.

image is enhanced, which helps CVD patients to distinguish colors.

Based on the above principles, we compared the method in this paper with the previous way. As shown in Figure 13 and Figure 14. The first row is a recolored image with different methods. The second row is the result of a CVD simulation of the first row of images. The third row is the result of saliency detection on the second row of images. We can see

from the Figure 13 that the color distribution of the images recolored by different methods is different. The correction of some methods can help color blind patients to achieve color differentiation, but it does not pay attention to the problem of saliency. CVD simulation was performed on the LMS recolored image, and the petals were blacker. The petals and leaves are similar in color, which is difficult to distinguish. Color contrast and Huang’s methods [8] recolored the images

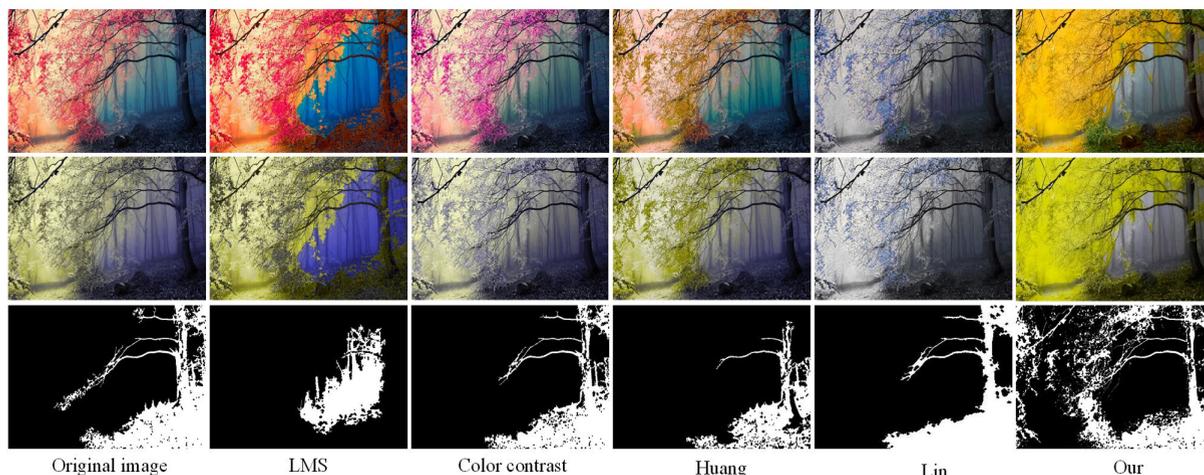


FIGURE 14. Examples of complex image comparison experiments.

TABLE 4. In 1 images, the accuracy rate and MEAN ± SEM of the tests were performed by 7 ordinary people and 7 visually impaired patients.

| Image index | Standard vision | Color vision defects | | | | |
|-------------|-----------------|----------------------|---------------|---------------|---------------|---------------|
| | | M1 | M2 | M3 | M4 | M5 |
| Image 1 | 85.71% | 14.29% | 42.86% | 57.14% | 71.43% | 85.71% |
| Image 2 | 85.71% | 28.57% | 28.57% | 28.57% | 42.86% | 71.43% |
| Image 3 | 85.71% | 28.57% | 28.57% | 71.43% | 85.71% | 85.71% |
| Image 4 | 100.00% | 14.29% | 14.29% | 42.86% | 42.86% | 85.71% |
| Image 5 | 100.00% | 42.86% | 42.86% | 57.14% | 42.86% | 85.71% |
| Image 6 | 100.00% | 28.57% | 28.57% | 57.14% | 57.14% | 71.43% |
| Image 7 | 85.71% | 14.29% | 28.57% | 57.14% | 57.14% | 85.71% |
| MEAN±SEM | 0.9184±0.0289 | 0.2449±0.0408 | 0.3061±0.0373 | 0.5306±0.0514 | 0.5714±0.0623 | 0.8163±0.0263 |

for CVD simulation. Although the petals are contrasted with the green leaves, the significance is not consistent. Lin’s algorithm [6] recolors the petals to sky blue, and the color contrast of the image is enhanced. From Figure 13, we observe the salient areas detected by Lin’s method. In addition to flowers, some green leaves are also treated as salient regions, and the salient regions have changed. Unlike Lin’s recoloring method, this article recolors the petals to purple. This color is consistent with the actual situation, and after the CVD simulation, the salient area of the image is substantially unchanged.

Figure 14 is a landscape map. There are more details in the image, and the content is more complicated. It is challenging to recolor for this kind of image. It can be seen from Figure 14 that the recoloring of the LMS method is incomplete, and the visual experience is poor. The color contrast method recolors the red maple leaf to pink, which is not desirable in perceptual cognition, and it violates the natural principle of color. The recoloring by Huang’s method is still standard, but after the color blindness simulation, the saliency of the image has changed. Lin’s approach recolors the maple tree to blue-purple, which is extremely rare in nature, and the salient areas of the recolored image have changed. In the method proposed in this paper, we select the image whose saliency does not change from the standard color image

collection as the reference image. We perform recoloring on the color-blind image according to the color scheme of the reference image, which meets the principles of natural color and color consistency.

Color feelings are more inclined to personal subjective opinions. To verify the effectiveness of the proposed method, we invite CVD patients to evaluate the recolored image. In the evaluation process, we tested images recolored by different methods and recorded the judgment results of the subjects. Among them, there are five recolor methods, named M1, M2, M3, M4, M5. These five methods are:

M1: LMS method. This method converts RGB space to LMS space. In LMS space, color information lost during color blindness simulation is supplemented, and the lost information is mapped to visible long (red) and short (blue) waves.

M2: RGB Color contrasting method. This algorithm halves the RGB pixel values to provide space for the increase of red and green pixel values, thereby achieving contrast enhancement.

M3: Huang’s color-blind image recoloring method [8]. This method uses a Gaussian Mixture Model (GMM) to represent color information and weigh key colors to ensure smoothness of the color of the recolored image.

TABLE 5. In II images, the accuracy rate and MEAN \pm SEM of the tests were performed by 7 ordinary people and 7 visually impaired patients.

| Image index | Standard vision | Color vision defects | | | | |
|----------------|---------------------|----------------------|---------------------|---------------------|---------------------|---------------------|
| | | M1 | M2 | M3 | M4 | M5 |
| Image 1 | 85.71% | 14.29% | 28.57% | 42.86% | 28.57% | 71.43% |
| Image 2 | 71.43% | 14.29% | 14.29% | 57.14% | 57.14% | 71.43% |
| Image 3 | 85.71% | 14.29% | 28.57% | 42.86% | 42.86% | 71.43% |
| Image 4 | 85.71% | 28.57% | 28.57% | 57.14% | 57.14% | 57.14% |
| Image 5 | 71.43% | 28.57% | 42.86% | 28.57% | 42.86% | 85.71% |
| Image 6 | 85.71% | 14.29% | 14.29% | 42.86% | 57.14% | 71.43% |
| Image 7 | 85.71% | 14.29% | 42.86% | 57.14% | 57.14% | 57.14% |
| MEAN \pm SEM | 0.8163 \pm 0.0263 | 0.1837 \pm 0.0263 | 0.2857 \pm 0.0441 | 0.4694 \pm 0.0408 | 0.4898 \pm 0.0425 | 0.6939 \pm 0.0374 |

M4: Lin's color-blind image recoloring method [6]. This method uses the directional feature vector of the color vision defect simulation result to reverse the color distribution. RGB color space is converted to λ , Y-B, R-G color space for recoloring.

M5: The method proposed in this paper. We use the saliency-unchanged image as the reference image to recolor the saliency-changed image to achieve the purpose of recoloring.

By explaining the purpose and process of the test, we found 7 green blind patients in the school who agreed to participate in the trial. Participants were teachers and classmates from the university, including 5 males and 2 females, aged 36.4 ± 15.2 years. Besides, we also found 7 standard vision subjects. Among them, there were 4 males and 3 females, with an age of 37.6 ± 14.3 . During the test, we need the subjects to point out the salient areas of the image within the specified time. We compare the salient area pointed out by the subjects with the real salient area. If the indication area is within a valid domain, the indication is considered correct. 1 score for right instruction and 0 score for false.

We selected 14 images and used different methods to recolor the color-blind images. The recolored images are divided into two groups, namely I and II. The background of I is relatively simple, and the background of II is relatively complicated. We record the judgment results of each image under each method and count the percentage of correct answers. We use the accuracy rate to evaluate the effect of single image correction. Besides, we calculated the judgment results of a group of images by all subjects. We use MEAN \pm SEM to calculate the accuracy of each method. The MEAN \pm SEM corresponds to the performance of the M_i method. The larger the number of right images, the higher the percentage and the better the correction effect.

To reduce the influence of other factors, keep the indoor light constant during the experiment. Firstly, testers enter gender and age into the computer. Immediately after, the next instruction message appears on the screen to ensure that the examiner understands the program correctly. To ensure the accuracy of the experimental results, we specified that the subjects observe each image for 3 seconds. The monitor is

color corrected, and the screen resolution is set to 1366×768 pixels.

Table 4 and Table 5 record the tests. Table 4 is the statistics for the group I, and Table 5 is the statistics for group II. Among them, M_i , Image $_j$ represents the accuracy rate of the j th image under the i th method. We use MEAN \pm SEM to calculate the accuracy of each method. We can see from the data in Table 4 that the accuracy of judging the salient area under standard vision is high, but there are also cases of misjudgment. The recoloring effect of M1 and M2 methods is relatively weak, and the highest accuracy of a single image does not exceed 50%. The MEAN \pm SEM of the M1 and M2 methods are 0.2449 ± 0.0408 and 0.3061 ± 0.0373 , respectively. The accuracy of the M3 and M4 methods has been improved. The accuracy of a single image has exceeded 20%, and some even exceeded 50%. M5 has high accuracy and MEAN \pm SEM, second only to standard vision. Compared with the previous four methods, the performance of M5 is the best, MEAN \pm SEM is 0.8163 ± 0.0263 .

Table 5 shows the statistics of group II images. Comparing the two sets of data in Table 4 and Table 5, the accuracy rate in Table 5 under standard vision is much lower than that in Table 4. This is because the background of the first group of images is relatively simple, and the context of the second group of images is relatively complicated. It can be seen from the MEAN \pm SEM of the two sets of data that the performance of the M1 and M2 methods is relatively low, and the MEAN \pm SEM is not high. The performance of M3 and M4 methods is better than M1 and M2, and the MEAN \pm SEM of the M4 method is more elevated than M3. M5 achieved excellent performance in all two groups of experiments. The accuracy of a single image and the MEAN \pm SEM are high. The performance of the five methods is sorted, and the order of performance from best to worst is M5, M4, M3, M2, M1.

V. CONCLUSION

In order to facilitate color difference detection in patients with color vision deficiency for salient areas at a level comparable to people having normal vision, this paper proposes saliency consistency-based image re-colorization. First, a content-based image retrieval method was used to find many images to form an image collection. CVD simulation was performed

on the images in the collection, and co-saliency detection was performed on the initial image and the CVD simulation image to compare the saliency maps of the two. Then, the image whose salient area remained almost unchanged was selected as the reference image and used to color the image whose salient area changed so that the color distribution of the colorized image is similar to that of the reference image. Because the color scheme of the recolored image has less impact on a colorblind patient, the colorblind patient's perception of the image is closer to that of a person with normal vision. The result not only achieves the purpose of significant correction but also meets the requirements of the required color correction. In different evaluation metrics show promising detection performances on various image. In future work, we want to explore the recolorization of color blind images in videos.

REFERENCES

- [1] B. Wong, "Points of view: Color blindness," *Nature Methods*, vol. 8, no. 6, p. 441, Jun. 2011.
- [2] H. V. W., "Tests for color blindness," *Amer. J. Ophthalmol.*, vol. 1, no. 6, p. 457, Jun. 1918.
- [3] R. W. G. Hunt, *Colour Standards and Calculations*. Hoboken, NJ, USA: Wiley, 2004, pp. 92–125.
- [4] J. Nathans, D. Thomas, and D. Hogness, "Molecular genetics of human color vision: The genes encoding blue, green, and red pigments," *Science*, vol. 232, no. 4747, pp. 193–202, Apr. 1986.
- [5] C. R. Huang, K. C. Chiu, and C. S. Chen, "Key color priority based image recoloring for dichromats," in *Proc. Pacific-Rim Conf. Multimedia*, 2010, p. 637–647.
- [6] H.-Y. Lin, L.-Q. Chen, and M.-L. Wang, "Improving discrimination in color vision deficiency by image re-coloring," *Sensors*, vol. 19, no. 10, pp. 2250–2269, 2019.
- [7] S. Poret, R. D. Dony, and S. Gregori, "Image processing for colour blindness correction," in *Proc. IEEE Toronto Int. Conf. Sci. Technol. Hum. (TIC-STH)*, Sep. 2009, pp. 539–544.
- [8] J. B. Huang, S. Y. Wu, and C. S. Chen, "Enhancing color representation for the color vision impaired," in *Proc. Workshop Comput. Vis. Appl. for Visually Impaired*, 2008, pp. 1–13.
- [9] Q. U. Dody, L. R. M. Tati, M. Richard, P. G. Andika, and N. A. Tauhid, "RGB color cluster and graph coloring algorithm for partial color blind correction," *MATEC Web Conf.*, vol. 255, Oct. 2019, Art. no. 01002.
- [10] F. Shao, G. Jiang, M. Yu, and Y.-S. Ho, "Highlight-detection-based color correction method for multiview images," *ETRI J.*, vol. 31, no. 4, pp. 448–450, Aug. 2009.
- [11] F. Shao, M. Yu, G. Jiang, K. Chen, and Z. Peng, "Colour correction pre-processing and chrominance reconstruction post-processing for multi-view video coding," *IET Image Process.*, vol. 6, no. 2, pp. 129–138, 2012.
- [12] F. Shao, G.-Y. Jiang, M. Yu, and Y.-S. Ho, "Fast color correction for multi-view video by modeling spatio-temporal variation," *J. Vis. Commun. Image Represent.*, vol. 21, nos. 5–6, pp. 392–403, Jul. 2010.
- [13] M. Ribeiro and A. J. P. Gomes, "Recoloring algorithms for colorblind people: A survey," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–37, Sep. 2019.
- [14] H. Brettel, F. Viénot, and J. D. Mollon, "Computerized simulation of color appearance for dichromats," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 14, no. 10, pp. 2647–2655, Oct. 1997.
- [15] G. M. Machado, M. M. Oliveira, and L. A. F. Fernandes, "Physiologically-based model for simulation of color vision deficiency," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 1291–1298, Oct. 2009.
- [16] K. Okajima and S. Kanbe, "A real-time color simulation of dichromats," *IEICE Tech. Report*, vol. 107, no. 117, p. 107–110, 2007.
- [17] M. M. Rivera, A. Padilla, and J. Canul-Reich, *Real-Time Recoloring Ishihara Plates Using Artificial Neural Networks for Helping Colorblind People*. Hershey, PA, USA: IGI Global, 2020, pp. 138–156.
- [18] Aytac, Sakire, "Using color blindness simulator during user interface development for accelerator control room applications," in *Proc. Int. Conf. Accel. Large Experim. Control Syst.*, 2017, pp. 1958–1968.
- [19] F. Radenović, G. Toliás, and O. Chum, "CNN image retrieval learns from bow: Unsupervised fine-tuning with hard examples," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 3–20.
- [20] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "DeepFashion: Powering robust clothes recognition and retrieval with rich annotations," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1096–1104.
- [21] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823.
- [22] J. Hays and A. A. Efros, "IM2GPS: Estimating geographic information from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
- [23] A. Cheng and E. Friedman, "Manipulability of page rank under Sybil strategies," in *Proc. 1st Workshop Netw. Syst.*, 2006, p. 75–82.
- [24] L. Zheng, Y. Yang, and Q. Tian, "SIFT meets CNN: A decade survey of instance retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1224–1244, May 2018.
- [25] B. Davis, F. Gutierrez-Barragan, and P. Praveena, "Content-based image retrieval using capsule networks," in *Proc. Int. Conf. Real-Time Intell. Syst.*, vol. 756, 2018, pp. 463–476.
- [26] L. Jiang, A. G. Hauptmann, and G. Xiang, "Leveraging high-level and low-level features for multimedia event detection," in *Proc. 20th ACM Int. Conf. Multimedia*, 2012, pp. 449–458.
- [27] A. Kovashka, D. Parikh, and K. Grauman, "WhittleSearch: Image search with relative attribute feedback," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2973–2980.
- [28] X. Han, Z. Wu, P. X. Huang, X. Zhang, M. Zhu, Y. Li, Y. Zhao, and L. S. Davis, "Automatic spatially-aware fashion concept discovery," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1463–1471.
- [29] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [30] V. Mnih, N. Heess, and A. Graves, "Recurrent models of visual attention," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2204–2212.
- [31] Xu K, Ba J, Kiros R, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2048–2057.
- [32] P. Sermanet, A. Frome, and E. Real, "Attention for fine-grained categorization," 2014, *arXiv:1412.7054*. [Online]. Available: <http://arxiv.org/abs/1412.7054>
- [33] H. Xu and K. Saenko, "Ask, attend and answer: Exploring question-guided spatial attention for visual question answering," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 451–466.
- [34] Z. Yang, X. He, J. Gao, L. Deng, and A. Smola, "Stacked attention networks for image question answering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 21–29.
- [35] D. A. Klein and S. Frintrop, "Center-surround divergence of feature statistics for salient object detection," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2214–2219.
- [36] M. M. Cheng, N. J. Mitra, and X. Huang, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, p. 569–582, Mar. 2015.
- [37] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3166–3173.
- [38] G. Li and Y. Yu, "Deep contrast learning for salient object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 478–487.
- [39] L. Wang, L. Wang, and H. Lu, "Saliency detection with recurrent fully convolutional networks," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 825–841.
- [40] P. Zhang, D. Wang, H. Lu, H. Wang, and X. Ruan, "Amulet: Aggregating multi-level convolutional features for salient object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 202–211.
- [41] P. Hu, B. Shuai, J. Liu, and G. Wang, "Deep level sets for salient object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2300–2309.
- [42] H. Fu, X. Cao, and Z. Tu, "Cluster-based co-saliency detection," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3766–3778, Oct. 2013.
- [43] X. Cao, Z. Tao, B. Zhang, H. Fu, and W. Feng, "Self-adaptively weighted co-saliency detection via rank constraint," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 4175–4186, Sep. 2014.
- [44] R. K. Gupta, A. Y. S. Chia, and D. Rajan, "Image colorization using similar images," in *Proc. ACM Int. Conf. Multimedia*, 2012, pp. 369–378.

- [45] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 689–698, Aug. 2004.
- [46] D. Kuzovkin, C. Chamaret, and T. Pouli, "Descriptor-based image colorization and regularization," in *Proc. Int. Workshop Comput. Color Imag.*, 2015, pp. 59–68.
- [47] R. C. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, 2016, pp. 649–666.
- [48] S. G. Liu and X. Zhang, "Automatic grayscale image colorization using histogram regression," *Pattern Recognit. Lett.*, vol. 33, no. 13, pp. 1673–1678, 2012.
- [49] G. Charpiat, M. Hofmann, and B. Schlkopf, "Automatic image colorization via multimodal predictions," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 126–139.
- [50] A. Bugeau, V.-T. Ta, and N. Papadakis, "Variational exemplar-based image colorization," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 298–307, Jan. 2014.
- [51] B. Wilhelm and J. B. Mark, *Principles of Digital Image Processing: Core Algorithms*. London, U.K.: Springer, 2009, pp. 231–232.
- [52] G.-H. Liu, J.-Y. Yang, and Z. Li, "Content-based image retrieval using computational visual attention model," *Pattern Recognit.*, vol. 48, no. 8, pp. 2554–2566, Aug. 2015.
- [53] M.-M. Cheng, Y. Liu, W.-Y. Lin, Z. Zhang, P. L. Rosin, and P. H. S. Torr, "BING: Binarized normed gradients for objectness estimation at 300 FPS," *Comput. Vis. Media*, vol. 5, no. 1, pp. 3–20, Mar. 2019.
- [54] D. Zhang, J. Han, C. Li, and J. Wang, "Co-saliency detection via looking deep and wide," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 2994–3002.
- [55] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *J. Vis.*, vol. 8, no. 7, p. 32, Dec. 2008.
- [56] Y. Jia and M. Han, "Category-independent object-level saliency detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1761–1768.
- [57] D. Zhou, J. Weston, and A. Gretton, "Ranking on data manifolds," in *Proc. Adv. Neural Inf. Process. Syst.*, 2004, pp. 169–176.
- [58] M. He, D. Chen, J. Liao, P. V. Sander, and L. Yuan, "Deep exemplar-based colorization," *ACM Trans. Graph.*, vol. 37, no. 4, pp. 1–16, Aug. 2018.



JINJIANG LI received the B.S. and M.S. degrees in computer science from the Taiyuan University of Technology, Taiyuan, China, in 2001 and 2004, respectively, and the Ph.D. degree in computer science from Shandong University, Jinan, China, in 2010. From 2004 to 2006, he was an Assistant Research Fellow with the Institute of Computer Science and Technology, Peking University, Beijing, China. From 2012 to 2014, he was a Post-doctoral Fellow with Tsinghua University, Beijing.

He is currently a Professor with the School of Computer Science and Technology, Shandong Technology and Business University. His research interests include image processing, computer graphics, computer vision, and machine learning.



XIAOMEI FENG received the B.S. degree from the School of Computer Science and Technology, Qilu Normal University, Jinan, China, in 2018. She is currently pursuing the M.S. degree with the School of Information and Electronic Engineering, Shandong Technology and Business University, Yantai, China. Her research interests include computer graphics, computer vision, and image processing.



HUI FAN received the B.S. degree in computer science from Shandong University, Jinan, China, in 1984, and the Ph.D. degree in computer science from the Taiyuan University of Technology, Taiyuan, China, in 2007. From 1984 to 2001, he was a Professor with the Computer Department, Taiyuan University Technology. He is currently a Professor with Shandong Technology and Business University. His research interests include computer aided geometric design,

computer graphics, information visualization, virtual reality, and image processing.

• • •