



# Color transfer method based on saliency features for color images

Shi Bao<sup>1</sup> · Ye Zhao<sup>1</sup> · Yatu Ji<sup>1</sup> · Nier Wu<sup>1</sup> · Gao Le<sup>1</sup>

Received: 4 October 2023 / Accepted: 5 May 2024 / Published online: 18 May 2024  
© The Optical Society of Japan 2024

## Abstract

With growing demands for higher image quality in the fields of film, video post-production, image restoration, art creation, and computer vision, color transfer between images has become an important research area. Based on previous research on color transfer techniques, this paper proposes a color transfer method for images based on saliency features, aiming at automatic color migration between them. Transferring colors based on the saliency features of the input image can avoid the problem of unnatural color of the output image due to mixing of colors from different regions. First, the local variances of both the original and reference images are calculated, serving as a temporary saliency feature map. This is followed by obtaining a refined saliency feature map after undergoing processes such as minimization filtering, binarization, expansion, and iteration. Subsequently, color is transferred between the saliency and non-saliency regions of the original and reference images. To avoid the generation of pseudo-contours, the image is then refined using base projection. Finally, an output image is obtained by fusing the base-projected image with the outcome from Reinhard's method, ensuring the output retains its naturalness and consistency. We conducted experiments with different types of images such as natural landscapes, buildings, and art paintings. The experimental results show that the method proposed in this paper not only retains the intricacies of the original image but also offers fuller and more realistic color renditions.

**Keywords** Color images · Color transfer · Saliency features · Color components

## 1 Introduction

In the field of visual arts, color is a crucial element that not only brings life and vibrancy to a work, but also a medium to convey emotions, ideas, and messages. Color transfer refers to the process where the color palette of one image (termed

the “original image”) is adapted to that of another, known as the “reference image”. Such techniques for color images are widely used in different fields, enhancing everything from images and videos to the aesthetics of artifacts and artworks, image correction for further processing [1]. For example, in film and TV special effects, color transfer can change the color scheme of a scene, enhancing its emotional expression or setting a special atmosphere. In the domain of advertising, these techniques can help make brand features, leading to more captivating visuals. For image restoration, color transfer can repair damaged or outdated images, rendering them more natural and appealing. Given its multifaceted utility, color transfer is a versatile asset in image processing that can be applied in several fields for solving various image-related problems. Therefore, various color transfer methods have been proposed to enhance the intricacy and accuracy of such effects. Researchers have carried out studies on color transfer methods [2–22], with some based on traditional image processing [2–10], while others leverage the power of deep learning [11–22].

In 2001, Reinhard et al. [2] proposed a global color transfer method based on the statistical information of

---

Ye Zhao, Yatu Ji, Nier Wu and Gao Le have contributed equally to this work.

---

✉ Shi Bao  
kshibao@163.com

Ye Zhao  
zhaoye\_mary@foxmail.com

Yatu Ji  
MLjyt@imut.edu.cn

Nier Wu  
wunier04@imut.edu.cn

Gao Le  
20211800102@imut.edu.cn

<sup>1</sup> College of Information Engineering, Inner Mongolia University of Technology, Aimin Street, Hohhot 010000, China

color spaces. This method adjusts the pixel values of the original image, presenting a simple and effective algorithm. However, when applied to colorful reference images, it can lead to blending of color regions, resulting in unnatural outcomes. Despite these shortcomings, the simplicity of the approach ensured that it was integrated into subsequent studies. Pitie et al. [3] proposed a new idea: matching an  $N$ -dimensional distribution with another by estimating a continuous transformation. Essentially, they constructed a color mapping model to fit the color relationship between the original and reference images. While this method adeptly preserves the structure of the original image, post-processing is necessary for color matching to remove noise and visual artifacts. Revisiting the domain in 2011, Reinhard et al. [4] proposed an incremental color transfer method, catering to images with different dynamic ranges. To address artifacts in the iterative distribution transfer (IDT), Ueda et al. [5] employed the results of Reinhard's method. By adaptively mixing the two methods they obtained color results more similar to the reference image, although issues with pseudo-contours persisted. In 2014, Hwang et al. [6] proposed a color transfer method based on moving least squares, notable for its advantage of being applicable to a wide range of situations. Whether faced with differences due to camera parameters, shooting times, or illumination conditions, this method proved resilient. However, it requires a large number of feature points as control points, implying that the target and source images must share the same scene, which may be limited in practical applications. Researchers such as Grogan [7, 8], leaned into a shape alignment-based approach. While it focuses on a localized approach, its implementation is time-intensive. In recent years, Wu et al. [9] proposed a new color transfer method using saliency feature mapping. This addressed the problem of color region mixing evident in Reinhard's method and achieves certain results. In 2022, Xu et al. [10] proposed a new IDT-based color transfer approach. Relying on color component projection, they adjusted weights to suppress pseudo-colors and improve the color similarity.

As the field of computer vision has advanced, methods based on statistics and machine learning have come into their own. Methods based on probabilistic models [11–13] have been adept at modeling color distributions, yielding more accurate transfer results. However, advent of deep learning has ushered in new possibilities for color transfer. Deep learning methodologies [14] aim to transform the transfer problem into a nonlinear regression problem, seeking to obtain the appearance mapping relationship between original and reference images [15]. Utilizing advanced architectures such as convolutional neural networks [16–19] (CNNs) and generative adversarial networks [20–22] (GANs), researchers have managed to learn more complex color mapping and transfer laws, achieving higher quality results.

While these strategies have achieved some success in dealing with different image types and application scenarios, there are still many directions worth exploring and enhancement. Color transfer faces many technical challenges. A pressing concern is aligning color distributions across different images to ensure naturalness and consistency of the transfer. Further, evading the inadvertent introduction of artifacts and distortions during color transition remains paramount.

In this paper, the image region exhibiting minimal pixel value fluctuations is designated as the “salient region”. Conversely, the region undergoing maximal pixel shifts is termed the “most significant region”. Predominantly, the color of the salient region of the reference image is transferred to the salient region of the original image, while the non-salient region of the reference image is transferred to the non-salient region of the original image. Most of the time, the resultant image post-color transfer has good picture quality. Therefore, we divide the original and reference images into salient and non-salient regions, completing color transfer correspondingly between these salient regions. Our methodology involved calculating the local variances of both images, establishing a temporary saliency feature map. This was followed by refining the saliency feature map through minimization filtering, binarization, expansion, and iterations. Subsequent steps encompassed transferring colors between the salient and non-salient regions of both the images. To suppress the generation of contours, we employed base projection. Finally, the final output image was derived by fusing the base-projected image with the resulting image derived using Reinhard et al.'s method.

## 2 Color transfer method of Reinhard [2]

This section delves into a seminal technique in the field of color transfer: the method proposed by Reinhard et al. [2].

Introduced in 2001, Reinhard et al.'s method seeks to equalize the color distributions of the original and reference images. By mapping them into a common color space, the color distributions of both the images can be better aligned to achieve color transfer. The steps are outlined as follows: (1) Convert the original and reference images to the  $YUV$  color space. (2) Compute the mean and the standard deviation for the original and reference images. (3) For each pixel of the original image across every channel, subtract the mean value of the original image, multiply by the ratio of the standard deviations of the reference to the original image, and then add the mean value of the reference image. (4) Transform the processed original image back to the color space to get the final result. This can be mathematically represented by Eq.(1):

$$I_{i;C}^{(R)} = (I_{i;C}^{(src)} - \mu_C^{(src)}) \frac{\sigma_C^{(ref)}}{\sigma_C^{(src)}} + \mu_C^{(ref)}, \tag{1}$$

where  $C \in \{Y, U, V\}$ ,  $\mu_C^{(src)}$  and  $\mu_C^{(ref)}$  denote the mean of each channel of the original and reference images, and  $\sigma_C^{(src)}$  and  $\sigma_C^{(ref)}$  denote the standard deviation of each channel of both the images. Reinhard et al.’s approach stands out for its simple and intuitive concept, which is easy to understand and implement. Although Reinhard et al.’s method holds a place in the field of color transfer as a classic algorithm, it has certain shortcomings when dealing with complex scenes and preserving details. Therefore, this paper delves deeper into Reinhard et al.’s method, aiming to uncover more efficient and accurate color transfer methods.

### 3 Proposed method

In this paper, we propose a color transfer method for color images based on saliency features, and the flowchart of the method in this paper is shown in Fig. 1.

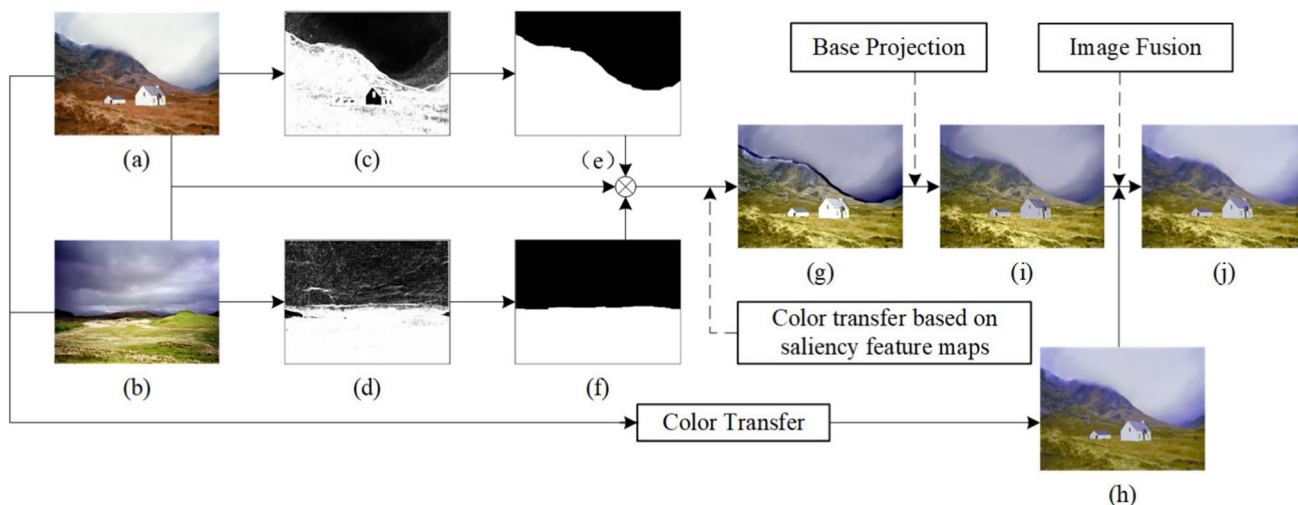
The method proposed in this paper aims at more accurate color migration while preserving the distinct style of the image. First, we obtain a temporary saliency feature map of the input image by calculating pixel variance in the CIEL\**a*\**b*\* color space. Second, an enhanced saliency feature map is obtained based on the temporary map through appropriate methods such as binarization. Third, according to the obtained saliency feature map, both the original and reference images are divided into salient and non-salient regions. Color transfer is then performed in these corresponding

regions, and their results are summed. Fourth, we produce a projection image by projecting the color components of the original image and adjusting weights to process the summed results. Finally, the outputs from Reinhard’s method and the projection are fused. In our experimental section, we apply our method alongside previous methods applied to several images. The efficacy of this method is then verified through objective evaluation.

Figure 1 provides an in-depth view of the framework of the proposed algorithm, elucidating aspects such as the derivation of temporary saliency feature map, the method of the saliency feature map and the algorithm for the region division and color transfer processes, projection result image, and final fusion result image.

#### 3.1 Temporary saliency feature map

Humans often effortlessly determine the significance of image regions, naturally focusing their attention toward significant parts [23]. Therefore, calculating the salient image sections is crucial. The CIEL\**a*\**b*\* color space contains all the colors that the human eye can perceive, and this color space calculates the variance of the image to better handle the differences between different colors in the local space. Therefore, in this paper, we first convert the original and reference images in Fig. 1 to the CIEL\**a*\**b*\* color space. Pixel variance is then computed, as described by Eq. (2), to obtain the temporary saliency feature maps for both images. These variance values are subsequently normalized between 0 and 1 utilizing the tanh function.



**Fig. 1** Flowchart of method proposed in this paper: **a** original image, **b** reference image, **c** temporary saliency feature map of original image, **d** temporary saliency feature map of reference image, **e** saliency feature map of original image, **f** saliency feature map of refer-

ence image, **g** image after sub-region transfer, **h** resulting image of Reinhard et al.’s method, **i** Resulting image of base projection, **j** final output image

$$v_i = \frac{1}{|S_{i,\rho}|} \sum_{j \in S_{i,\rho}} [(L_j^* - \langle L_i^* \rangle)^2 + (a_j^* - \langle a_i^* \rangle)^2 + (b_j^* - \langle b_i^* \rangle)^2], \quad (2)$$

$$M_i^{(v)} = 255 * \tanh\left(\frac{v_i}{\alpha}\right), \quad (3)$$

where  $S_{i,\rho}$  denotes the set of pixels in a window centered at  $i$ th pixel with chessboard distance  $\rho$  of 2, and  $|S_{i,\rho}|$  is the number of pixels in the window. The region with a distance of  $\rho = 2$  from the central pixel refers to the  $5 \times 5$  area centered at that pixel.  $\langle L_i^* \rangle$ ,  $\langle a_i^* \rangle$ , and  $\langle b_i^* \rangle$  represent the average values of each channel within the  $5 \times 5$  window centered on the

$i$ th pixel, and  $L_j^*$ ,  $a_j^*$ , and  $b_j^*$  denote the values of each channel in the  $j$ th pixel in the window.  $\alpha$  is set to 15, and  $M_i^{(v)}$  is taken as the value of the  $i$ -th pixel in the temporary saliency feature map. Figure 3 presents the temporary saliency feature maps,  $M^{(src)}$  and  $M^{(ref)}$ , obtained by converting the original and reference images from Fig. 2.  $M^{(src)}$  and  $M^{(ref)}$  are obtained from Eq. (3).

### 3.2 Saliency feature map

The resulting temporary saliency feature maps for the original and reference images are computationally processed as



Fig. 2 Input image: **a** original image, **b** reference image

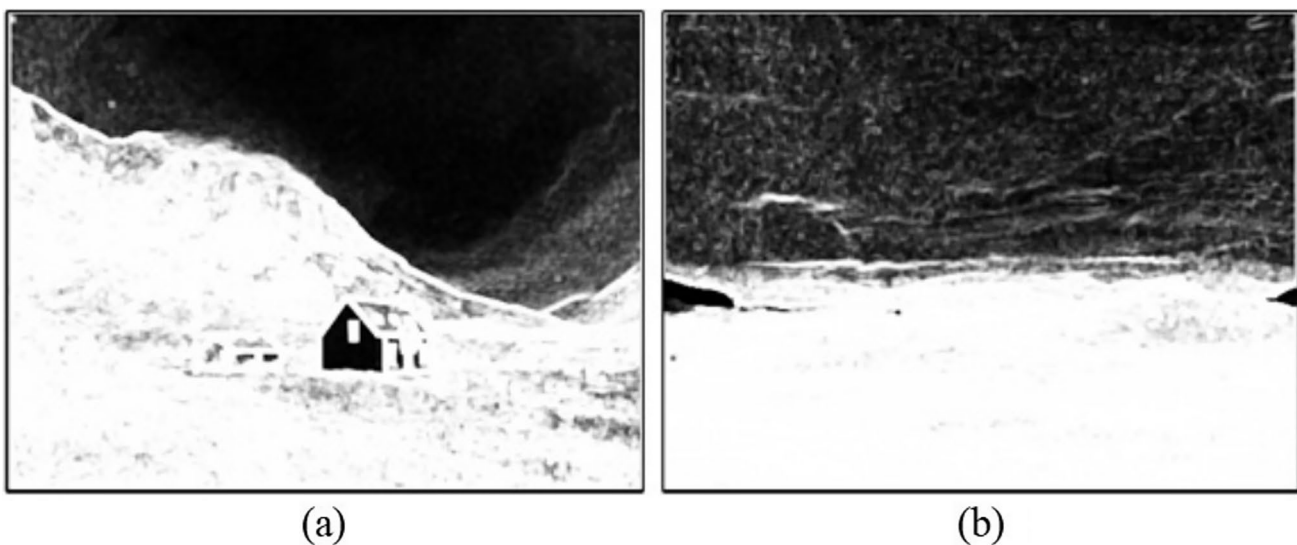


Fig. 3 Temporary saliency feature map: **a** original image, **b** reference image



depicted in the flowchart in Fig. 4 to derive the saliency feature maps. Within our proposed method, the temporary saliency feature maps  $M^{(src)}$  and  $M^{(ref)}$  were initially subjected to a minimum filter. This step ensures that the thresholds do not surge excessively in the following binarization process. The minimum filter is represented as follows:

$$M'_i = \text{MIN}\{M_j \mid j \in S_{i,\rho}\}, \tag{4}$$

where  $\text{MIN}(\cdot)$  returns the minimum value of the set of pixels  $S_{i,\rho}$  within a window by centering pixel  $i$  with chessboard distance  $\rho$  of 2. Substituting  $M^{(src)}$  and  $M^{(ref)}$  yields  $M'^{(src)}$  and  $M'^{(ref)}$ . The results of this minimum filtering process,  $M'^{(src)}$  and  $M'^{(ref)}$ , are then binarized, dividing the results into significant and non-significant regions.

$$M''_i = \text{BF}(M'_i),$$

$$\text{BF}(z) = \begin{cases} 255, & z \geq \langle M' \rangle, \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where  $\text{BF}(\cdot)$  is the binarization function, and  $\langle M' \rangle$  is the average of the temporary saliency feature maps after minimum filtering. As a threshold for binarization,  $M'^{(src)}$  and  $M'^{(ref)}$  are used. The purpose of expansion is to remove non-significant pixels (or black pixels in a white region) from

the more salient regions in the binarization result due to improper delineation. Conversely, erosion serves to remove white pixels from the black regions. In this paper, the window size of both expansion and erosion was set at  $7 \times 7$ . The number of iterations was five for expansion and seven for erosion. The erosion intensity was twice that of expansion to balance the expansion of the black area into the white area during the minimum filtering. This ensures a more accurate color conversion within the relevant regions. The saliency feature maps obtained in this step are  $SM^{(src)}$  and  $SM^{(ref)}$ , as shown in Fig. 5.

### 3.3 Region division and color transfer

This section aims to transfer colors between regions that receive the same level of attention to the human eye and avoid the problem of mixing color regions. Since the color transfer step uses Reinhard’s method, this part operates in the  $YUV$  color space. Figure 6 shows the comparison of the output of Reinhard’s method using  $\text{CIEL}^*a^*b^*$  and  $YUV$  color space. The first and second rows are the original and reference images, and the third and fourth rows are the output images using Reinhard’s method in  $\text{CIEL}^*a^*b^*$  and  $YUV$  color space respectively.

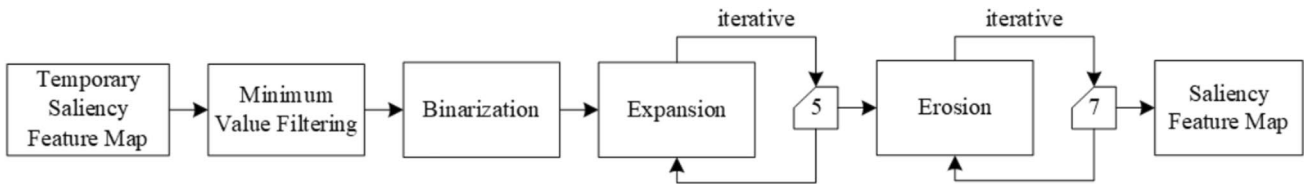
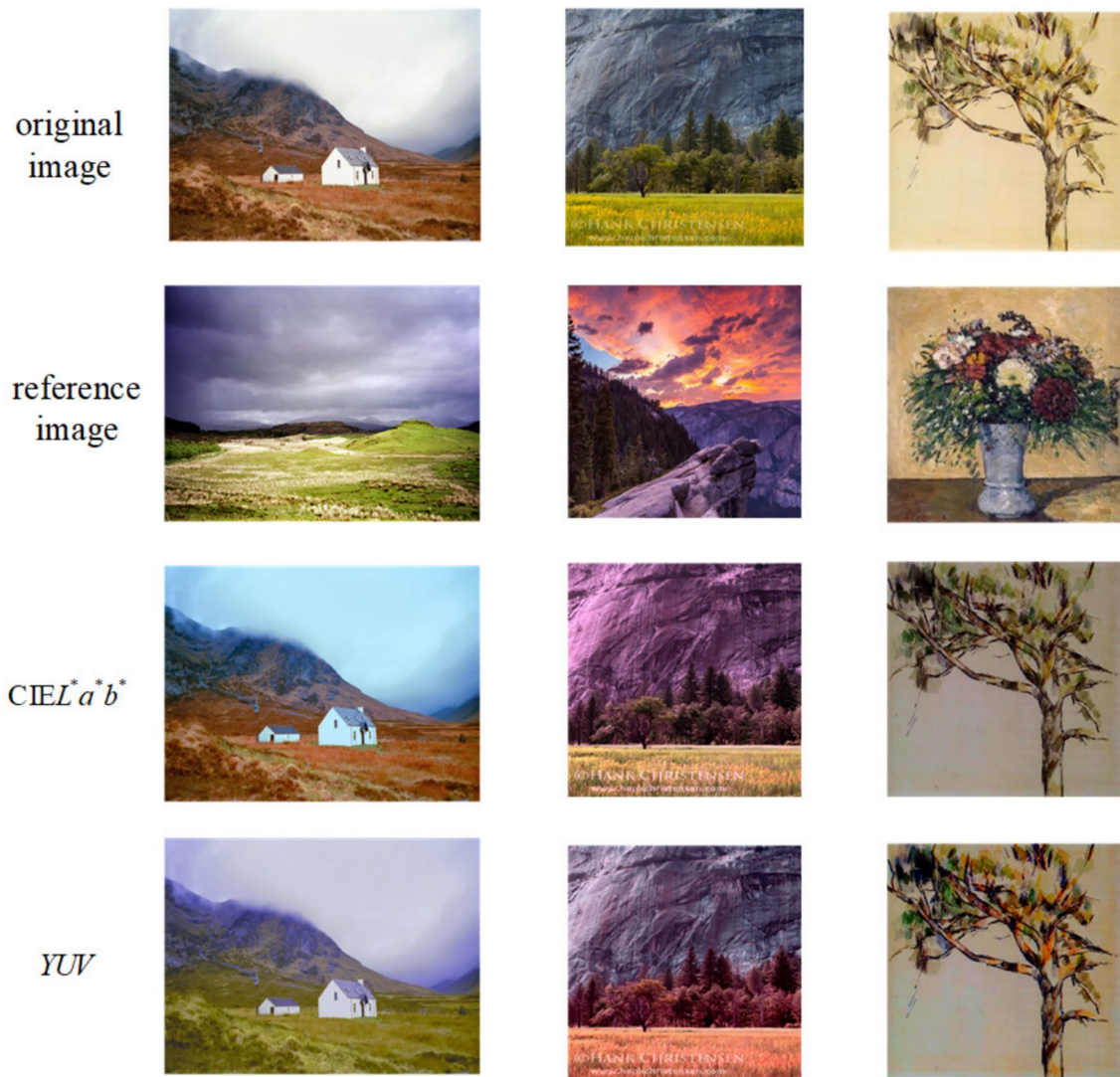


Fig. 4 Flowchart for calculating saliency feature map



Fig. 5 Saliency feature map: a original image, b reference image



**Fig. 6** Comparison of effect images for Reinhard et al. method in  $YUV$  and  $CIEL^*a^*b^*$  color spaces

From the comparison result, it can be seen that the output image in  $YUV$  color space is more natural and the colors more closely match the color distribution of the reference image. The criterion for division is the average,  $\langle SM \rangle$ , of all the pixel values of the saliency feature map.

$$\begin{cases} I_{y;i} = I_i, & SM_i \geq \langle SM \rangle, \\ I_{n;i} = I_i, & SM_i < \langle SM \rangle. \end{cases} \quad (6)$$

According to Eq. (6), the original image  $I^{(src)}$  and the reference image  $I^{(ref)}$  are divided into significant and non-significant regions, where  $I^{(src)}$  and  $I^{(ref)}$  are substituted into  $I_i$  in Eq. (6), the original image is then divided into significant,  $I_y^{(src)}$ , and non-significant,  $I_n^{(src)}$ , regions and the reference image is also divided into significant,  $I_y^{(ref)}$ , and non-significant,

$I_n^{(ref)}$ , regions. Next, color transfer is performed in the corresponding region using the Reinhard method, which avoids the problem of mixing the mean values of each region and gives a more natural appearance to the color of the resulting image. The color transfer of the corresponding region was performed according to Eq. (7).

$$I_{y;i}^{(SM;C)} = (I_{y;i}^{(src;C)} - m_y^{(src;C)}) \frac{\sigma_y^{(ref;C)}}{\sigma_y^{(src;C)}} + m_y^{(ref;C)}, \quad (7)$$

where  $C \in \{Y, U, V\}$ ,  $\sigma_y^{(src;C)}$ , and  $\sigma_y^{(ref;C)}$  denote the standard deviation of each channel in regions  $I_y^{(src)}$  and  $I_y^{(ref)}$ .  $m_y^{(src;C)}$  and  $m_y^{(ref;C)}$  denote the means of  $I_y^{(src)}$  and  $I_y^{(ref)}$ , respectively. The formula is calculated as shown in Eq. (8).

$$\begin{cases} m_y^{(src;C)} = \frac{1}{|S_y^{(src)}|} \sum_{i \in S_y^{(src)}} I_{y;i}^{(src;C)}, \\ m_y^{(ref;C)} = \frac{1}{|S_y^{(ref)}|} \sum_{i \in S_y^{(ref)}} I_{y;i}^{(ref;C)}, \end{cases} \quad (8)$$

where  $S_y^{(src)}$  is the set of pixels with  $SM^{(src)} = 255$  in the original image, and  $|S_y^{(src)}|$  represents the total number of this set.  $S_y^{(ref)}$  is the set of pixels with  $SM^{(ref)} = 255$  in the reference image, and  $|S_y^{(ref)}|$  represents the total number of this set.

Similarly, the color transfer in the non-salient region operates as follows:

$$I_{n;i}^{(SM;C)} = (I_{n;i}^{(src;C)} - m_n^{(src;C)}) \frac{\sigma_n^{(ref;C)}}{\sigma_n^{(src;C)}} + m_n^{(ref;C)}, \quad (9)$$

where  $\sigma_n^{(src;C)}$  and  $\sigma_n^{(ref;C)}$  denote the standard deviation of each channel in regions  $I_n^{(src)}$  and  $I_n^{(ref)}$ , respectively.  $m_n^{(src;C)}$  and  $m_n^{(ref;C)}$  denote the mean of each channel in regions  $I_n^{(src)}$  and  $I_n^{(ref)}$ , respectively. The calculation is further detailed in Eq. (10).

$$\begin{cases} m_n^{(src;C)} = \frac{1}{|S_n^{(src)}|} \sum_{i \in S_n^{(src)}} I_{n;i}^{(src;C)}, \\ m_n^{(ref;C)} = \frac{1}{|S_n^{(ref)}|} \sum_{i \in S_n^{(ref)}} I_{n;i}^{(ref;C)}, \end{cases} \quad (10)$$

where  $S_n^{(src)}$  is the set of pixels with  $SM^{(src)} = 0$  in the original image, and  $|S_n^{(src)}|$  represents the total number of this set.  $S_n^{(ref)}$  is the set of pixels with  $SM^{(ref)} = 0$  in the reference image, and  $|S_n^{(ref)}|$  represents the total number of this set. The results of the saliency and non-saliency regions are summed according to Eq. (11). Following this, the color space is reverted to *RGB* to derive the result of  $I^{(SM)}$  after color transfer, as shown in Fig. 7. It is evident that the summed image shows false contours due to color differences between the regions. This problem will be dealt with in Sect. 3.4.

$$I^{(SM;C)} = I_y^{(SM;C)} + I_n^{(SM;C)}. \quad (11)$$

### 3.4 Base projection

False contours are generated between the salient and non-salient regions of the sub-region transferred image. False contours are created when the similar or same color appears as a different color. The base projection generates the same coefficients through the same basis, so the projected colors are also the same. Therefore, in this study, the problem of false contours in images after subregion transfer can be fixed using the base projection method. Concurrently, the base



Fig. 7 Image after sub-region transfer

projection ensures consistent colors in the original image are translated uniformly. Figure 8 shows the flowchart for the base projection.

Within the *RGB* color space, 12 bases  $\text{Base} = \{R, G, B, R^2, G^2, B^2, R^{0.5}, G^{0.5}, B^{0.5}, R * G, R * B, G * B\}$  were obtained based on the *RGB* components of the original image, and all elements of *Base* are multiplied by 100. The derived projection image is represented as  $I^{(SM)}$ . The results of the color component projection were obtained according to the linear equation in Eq. (12).

$$I_i^{(SM)} = \sum_{l \in \text{Base}} k(l) \cdot x_i(l), \quad (12)$$

where  $x_i(l)$  denotes the  $l$ -th base of the  $i$ -th pixel.  $k$  is the projection coefficient that is solved by constructing the objective function from Eq. (13).

$$\tilde{k} = \arg \min_{k \in \mathfrak{R}} \left[ \sum_{i=1}^n (I_i^{(SM)} - I_i^{(SM)})^2 \right]. \quad (13)$$

Here,  $k \in \mathfrak{R}$  means that  $k$  contains 12 real numbers and  $n$  is the number of pixels in the image. Figure 9 shows the results obtained for the base projection section.

### 3.5 Image fusion

Occasionally, the resultant image after base projection exhibits artifactual colors in specific regions. To suppress these pseudo-colors, this study employed the saturation of the original image as a weight in the fusion between the base projection result image and the resultant image from Reinhard et al. Lower color saturation in the original image indicates that the altered color is close to neutral; thus, there

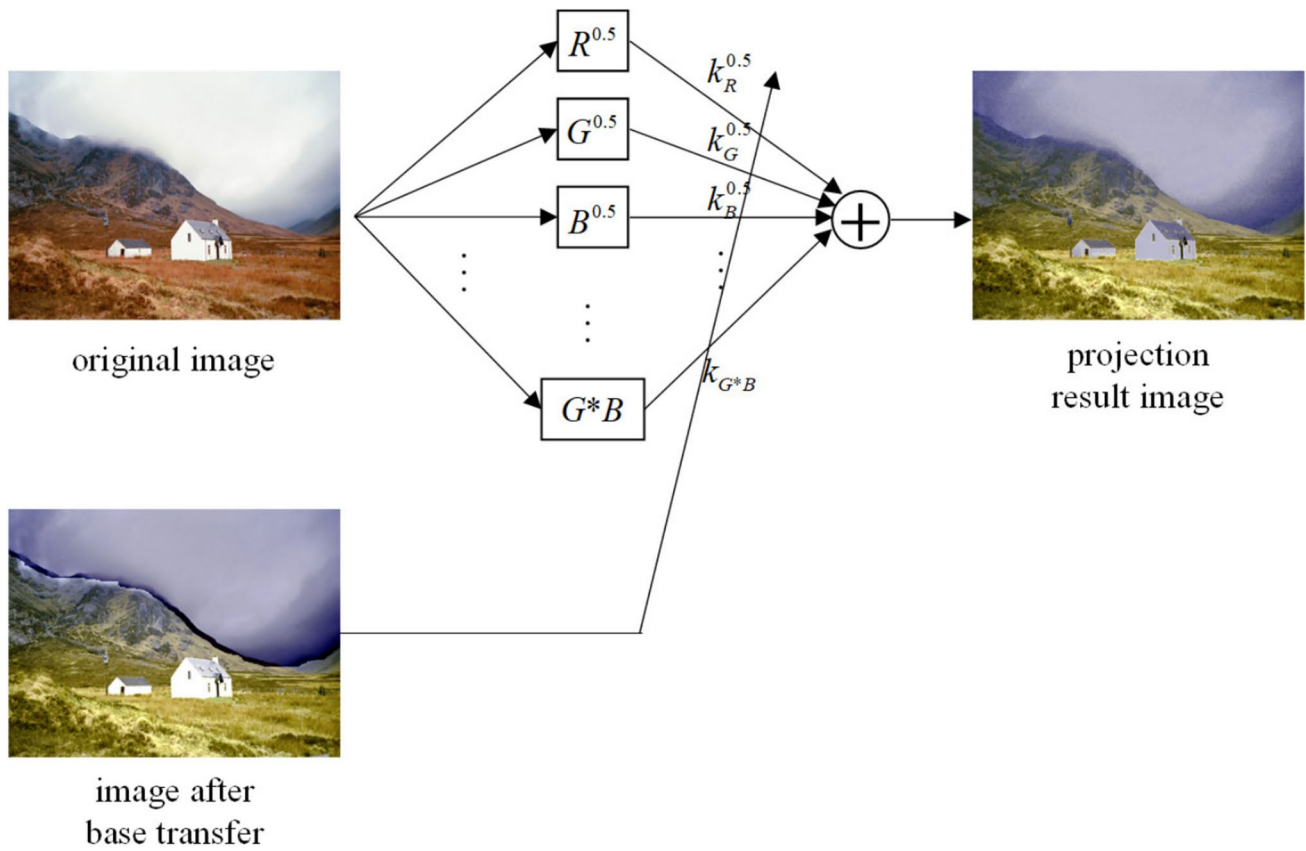


Fig. 8 Flowchart of base projection



Fig. 9 Base projection results

is no need for color transfer, focusing on regions with peak saturation.

For fusion, the color information from the resultant image  $I^{(R)}$  of the Reinhard et al. technique and the resultant image  $I^{(SM)}$  obtained in Sect. 3.4 are combined, utilizing the saturation information of the original image  $I^{(src)}$ , to

obtain the final image  $I^{(out)}$ . The original image is converted to CIEL\*a\*b\* color space, and the saturation information of the original image was calculated as a weighted value according to Eq. (14) [24] for the compounded weighting of the  $I^{(R)}$  and  $I^{(SM)}$  images.

$$w_i = \tanh(\sqrt{a^2 + b^2/\beta}), \tag{14}$$

$$I_i^{(out)} = [(1 - w_i)I_i^{(R)} + w_i I_i^{(SM)}]. \tag{15}$$

Equation (15) aims to prioritize the use of Reinhard’s results in low-saturation regions and the projected image  $I^{(SM)}$  results in high-saturation regions.

### 4 Experimental results

In this section, we delve into the experimental design and provide an in-depth analysis of the evaluation metrics for our proposed method. For the experiments, we selected some photos commonly used in the field of color transfer. This assortment encompassed a diverse range of images, from natural landscapes and architectural buildings to artistic



paintings. In the tuning experiments of the parameters, we carried out tuning experiments for the parameter  $\alpha$  in Sect. 3.1, the number of iterations for erosion and expansion in Sect. 3.2, and the parameter  $\beta$  involved in Sect. 3.5, respectively.

Firstly, Fig. 10 illustrates the effect of the parameter  $\alpha$  on the temporary saliency feature map at different values. Taking the input image in Fig. 10 as an example, the temporary saliency feature map should have the parts of mountains and grass as salient regions and the parts of sky as non-salient regions. Observing Fig. 10, it can be noticed that when the parameter  $\alpha$  is 5, the part of the sky in subfigure (b) will have too many white pixels. When the parameter  $\alpha$  is 25, the part of grass in subfigure (f) will have too many black pixels. The same trend is observed for most of the images; therefore, by way of compromise, we choose to set the parameter  $\alpha$  to 15, which can be applied to all the experimental images.

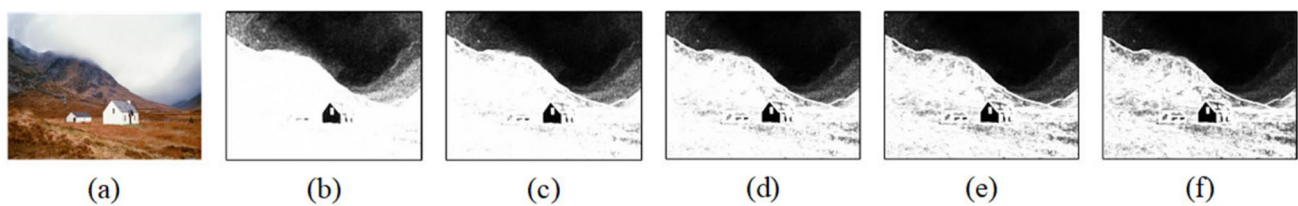
Next, Fig. 11 shows the effect of the number of iterations of erosion versus expansion on the saliency feature maps, with subfigures (b) to (e) showing the saliency feature maps after 2, 3, 4 and 5 iterations of erosion and expansion, in that order. The experimental results show that for most of the images, the features of the input image can be extracted and highlighted more effectively when the number of iterations is set to 5. Therefore, considering the generality and robustness of the method, we choose an iteration number of 5.

Finally, we conducted parameter  $\beta$  tuning experiments during the fusion phase to investigate the effect of different parameter values on the results. The graphical result generated at different parameter values are plotted for a side-by-side comparison in Fig. 12. From our experiments,

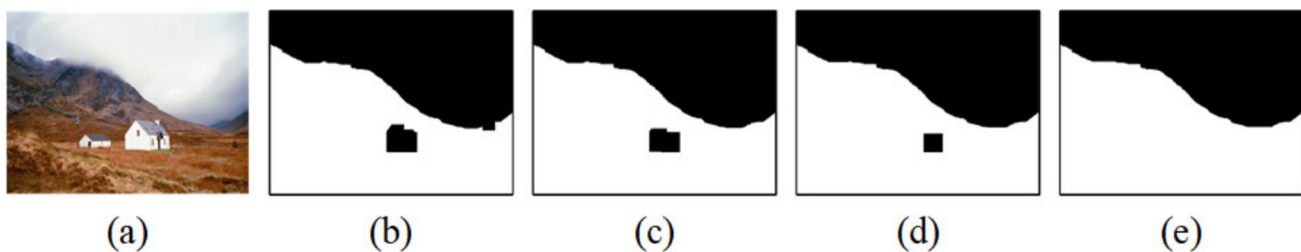
it became evident that the setting for parameter  $\beta$  will have some effect on the results for a given image set. In Fig. 12, the first column is the output image and the second to sixth columns represent a portion of the output image, labelled with dashed boxes, showing the different output effects for parameter  $\beta$  values of 10, 20, 30, 40, and 50, respectively, in that order. Observations indicate that clarity improves incrementally with values of  $\beta$  ranging from 10 to 30. Pixelation is not seen in this range and is most noticeable in the first and second rows. However, as  $\beta$  gradually increases from 30 to 50, the entire image progressively darkens, which is most obvious in the third row. Given these observations, we choose  $\beta = 30$  as the optimal value, producing the most favorable result.

To maintain a standard of comparison, we applied our method to an identical set of photographs as those utilized by Reinhard et al., Pitie et al., Ueda et al., Wu et al., and Xu et al. The comparative results of these methods, for a total of eight distinct image sets, are exhibited in Fig. 13.

In the illustration provided in Fig. 13, the original image (displayed in the first row) and the reference image (in the third row) exhibit different color distributions. The second row is the fusion weight maps with the weights of the fused parts of Sect. 3.5 as the image output, these weight maps adeptly elucidate the specific image regions we focused on during the color fusion process. Brighter sections of the image focus on the results of the projected image  $I^{(SM)}$ , whereas the darker areas incorporate the outcomes from the method of Reinhardt et al. However, our proposed method manages to produce a color distribution that not only appears more natural but also mirrors the colors of the reference



**Fig. 10** Temporary saliency feature maps of the output at different values of parameter  $\alpha$ : **a** input image, **b**  $\alpha=5$ , **c**  $\alpha=10$ , **d**  $\alpha=15$ , **e**  $\alpha=20$ , **f**  $\alpha=25$



**Fig. 11** Saliency feature maps for different number of iterations: **a** input image, **b** 2, **c** 3, **d** 4, **e** 5

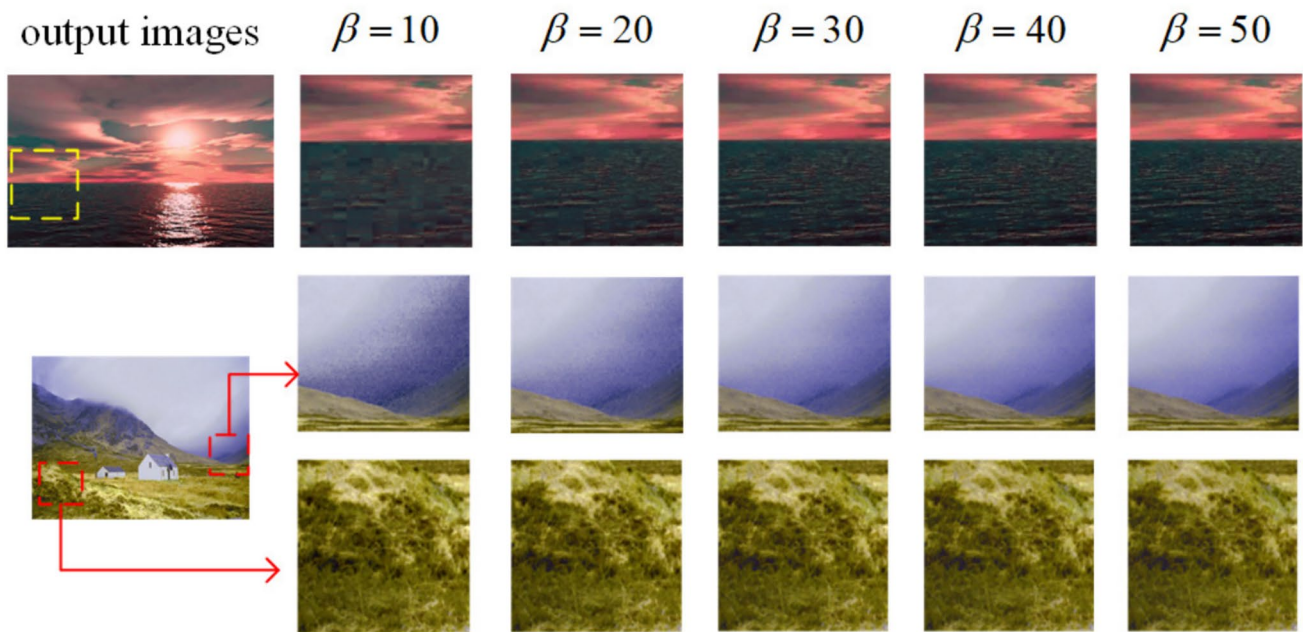


Fig. 12 Detailed images of output results for different values of parameter  $\beta$

image. At the same time, the intricacies and details within the images are preserved maintained.

We have focused on preserving the color of the original image, which has historically been susceptible to distortion in previous methods. Taking image group 4 as an example, Fig. 14 shows a zoomed-in perspective. This close-up reveals that the output images from methods by Pitie et al., Ueda et al., and Xu et al. undergo noticeable color shifts, leading to an unnatural appearance of the wall segment. While Wu et al.'s method exhibits improved color, a pronounced black area is still evident. This recurring appearance of black pixels in the output is attributed to the unsuitability of parameter criteria chosen by Wu et al. for demarcating regions across diverse image types. The results from Wu et al. bear similarities to those by Reinhard et al. Figure 15 shows a detail view of the results of the method herein compared to the results of Reinhard. The first column is the output image, and the second and third columns are the portions of the output image in which the method of this paper and the method of Reinhard et al. mark dashed boxes, respectively. The results of the method in this article are similar to those of the method of Reinhard et al. However, a detailed comparison indicates a difference in brightness: Reinhard et al.'s outcomes are visibly dimmer. Our method, on the other hand, produces vibrant, authentic colors and faithfully conserves the colors in the reference image.

Another important aspect is the ability to retain details. The second image group focuses on the texture and minute details of the original image. Figure 16 displays detailed

views from various methods. Compared to others, both Pitie et al. and Ueda et al. exhibited severe distortions in the sun part of their output images. Furthermore, while the method by Wu et al. has discernible black areas, Xu et al.'s approach performs somewhat better, though it still yields a slightly blurred sky region. In stark contrast, our method is able to better preserve the detailed features from the original image, such as texture and sharp edges, making the output image visually clearer. To evaluate the performance of the different methods, we used objective evaluation metrics: Kullback–Leibler Divergence (KLD) [25], False Color Index (FCI) [10], and Structural Similarity Index Measure (SSIM) [26]. The KLD quantifies the difference in color distribution between reference and output images.

$$\text{KLD}(P||Q) = \sum_i P_i \ln \frac{P_i}{Q_i}, \quad (16)$$

where  $P$  and  $Q$  are the color histograms of the reference image and the output image, respectively. FCI evaluates the prevalence of pseudo-colors in the transferred output image.

$$\text{FCI} = \sum_{i=1}^n \sum_{j=i+1}^n \Phi_{ij}, \quad (17)$$

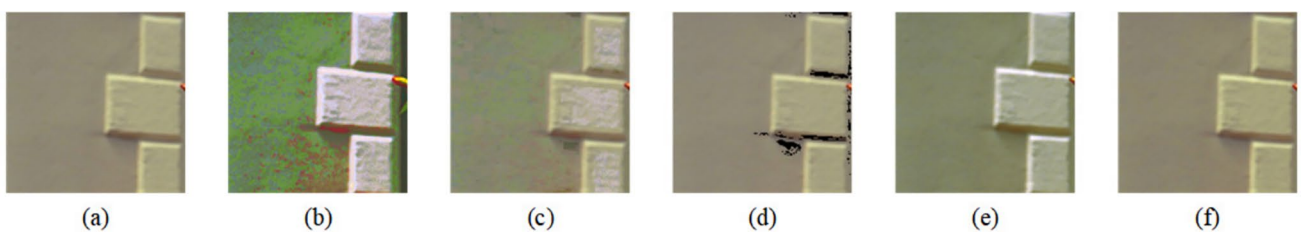
$$\Phi_{ij} = \begin{cases} 1, & \Delta E_{ij}^{(src)} < \xi \text{ and } \Delta E_{ij}^{(out)} > \eta \\ 0, & \text{otherwise} \end{cases},$$

where  $\Delta E_{ij} = \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}$ , and the values of  $\xi$  and  $\eta$  are 5 and 15. SSIM measures the structural



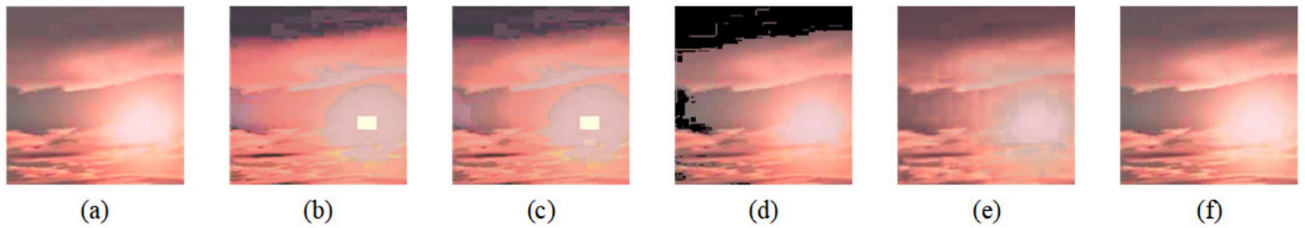
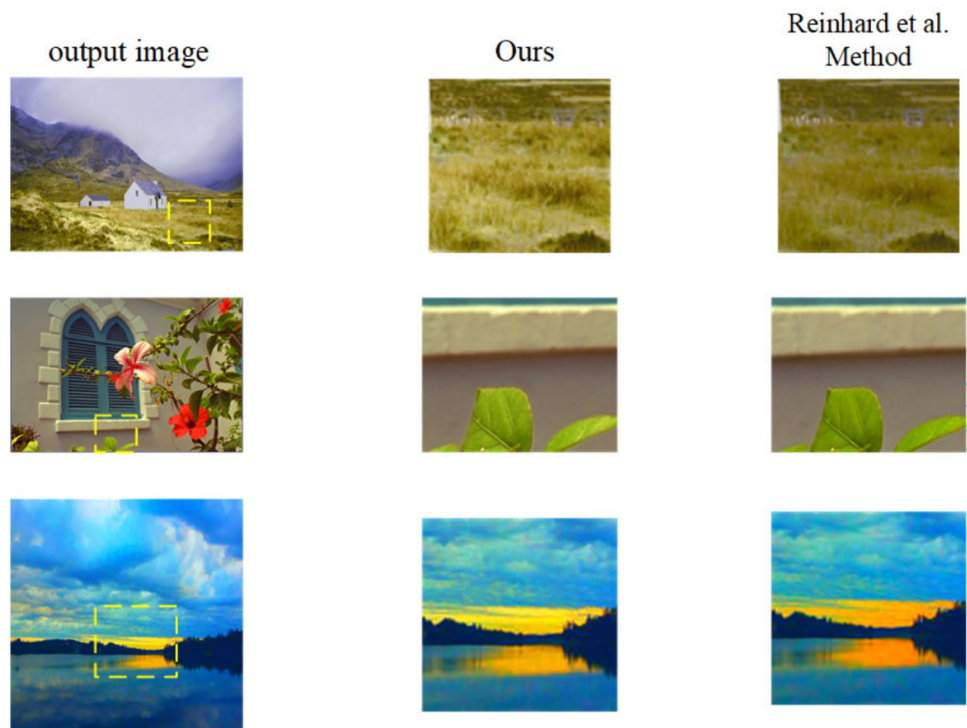


**Fig. 13** Plot of the results of method proposed in this paper alongside Reinhard et al., Pitie et al., Ueda et al., Wu et al., and Xu et al. The first row is the original image, the second row is the fusion weight image and the third row is the reference image



**Fig. 14** Diagrams representing individual details of red-boxed portion of Fig. 13: **a** Reinhard et al. method, **b** Pitie et al. method, **c** Ueda et al. method, **d** Wu et al. method, **e** Xu et al. method, **f** our method

**Fig. 15** Detailed images of the results of this paper’s method compared to Reinhard’s results



**Fig. 16** Diagrams representing various details of yellow boxed portion of Fig. 13: **a** Reinhard et al. method, **b** Pitie et al. method, **c** Ueda et al. method, **d** Wu et al. method, **e** Xu et al. method, **f** Our method

similarity between the transferred output and the original images.

$$SSIM(src, out) = \frac{(2\mu^{(src)}\mu^{(out)} + C_1)(2\sigma^{(src,out)} + C_2)}{((\mu^{(src)})^2 + (\mu^{(out)})^2 + C_1)((\sigma^{(src)})^2 + (\sigma^{(out)})^2 + C_2)}, \tag{18}$$

where  $\mu^{(src)}$  and  $\mu^{(out)}$  denote the means the original and output images,  $(\sigma^{(src)})^2$  and  $(\sigma^{(out)})^2$  denote the variances of the original and output images,  $\sigma^{(src,out)}$  is the covariance between the original and output images, and  $C_1$  and  $C_2$  are constants to avoid the case where the denominator is zero.

Table 1 shows the KLD values for output images produced by different methods. The data suggest that the KLD values of the proposed method are intermediate in comparison to other methods. Specifically, Pitie et al., Ueda et al., and Xu et al.’s methods possess KLD values smaller than ours, while the KLD values from Reinhard et al. and Wu et al. are larger than that of the KLD method of the method

proposed in this paper. However, it is worth noting that the KLD metric does not fully reflect the color distribution similarities between output and reference images, because the KLD metrics do not take into account factors such as color saturation and luminance, and therefore a combination of other metrics is needed to assess the quality of the output image in a more comprehensive way.

Table 2 shows the pseudo-colors calculated from the output images of different methods. We have highlighted the lowest value for each row. Our method consistently registered lower FCI values compared to other methods and have the smallest average value. This indicates that our proposed method produces fewer pseudo-colors in its output images, thus enhancing the resultant image’s quality.

To assess image generation quality, we employed the SSIM, which measures structural similarities between the output images of different methods and the original images. Table 3 shows the SSIM values associated with each method



**Table 1** KLD of output images for each method

	Reinhard et al. method [2]	Pitie et al. method [3]	Ueda et al. method [5]	Wu et al. method [9]	Xu et al. method [10]	Ours
Image Group 1	25.1	22.6	<b>22.1</b>	30.6	25.1	23.9
Image Group 2	33.2	<b>19.8</b>	22.7	30.6	36.9	32.7
Image Group 3	38.8	37.5	37.0	<b>30.6</b>	38.6	38.6
Image Group 4	39.4	38.7	37.7	39.2	<b>36.1</b>	39.1
Image Group 5	35.4	<b>33.9</b>	34.7	34.9	35.4	35.7
Image Group 6	37.5	<b>32.1</b>	36.7	33.5	37.5	37.2
Image Group 7	35.7	34.1	<b>33.8</b>	34.8	35.7	35.2
Image Group 8	37.6	32.7	<b>31.7</b>	35.2	36.3	37.5
avg	35.3	<b>31.4</b>	32.1	33.7	35.2	35.0

The smallest value is shown in bold in each row

**Table 2** FCI of output images for each method

	Reinhard et al. method [2]	Pitie et al. method [3]	Ueda et al. method [5]	Wu et al. method [9]	Xu et al. method [10]	Ours
Image Group 1	13787	534160436	134844930	3762619	13787	<b>13722</b>
Image Group 2	491171797	2266459150	2204284375	<b>5162283</b>	1640147810	223378811
Image Group 3	3637337	377120595	228296368	18453842	17303511	<b>2485067</b>
Image Group 4	<b>0</b>	4895980523	1706383907	219633138	1245892799	3
Image Group 5	5043	12723255	4260822	785513	10406	<b>2388</b>
Image Group 6	21	2899817	70461	<b>0</b>	71715	<b>0</b>
Image Group 7	59496	129088879	97608246	4393680	9002494	<b>20</b>
Image Group 8	43424916	35819525	31832496	<b>28461521</b>	47208987	42389217
avg	67289049.6	1031781523.0	550947700.6	35081574.5	369956438.6	<b>33533653.5</b>

The smallest value is shown in bold in each row

**Table 3** SSIM of output images for each method

	Reinhard et al. method [2]	Pitie et al. method [3]	Ueda et al. method [5]	Wu et al. method [9]	Xu et al. method [10]	Ours
Image Group 1	0.921	0.718	0.855	0.857	0.921	<b>0.929</b>
Image Group 2	0.805	0.663	0.686	0.735	0.578	<b>0.826</b>
Image Group 3	0.829	0.679	0.796	0.802	0.806	<b>0.831</b>
Image Group 4	0.903	0.609	0.801	0.832	0.758	<b>0.934</b>
Image Group 5	0.837	0.753	0.822	0.795	0.836	<b>0.839</b>
Image Group 6	0.792	0.825	0.791	<b>0.842</b>	0.778	0.811
Image Group 7	0.862	0.686	0.749	0.758	0.77	<b>0.907</b>
Image Group 8	0.574	0.512	0.542	0.557	0.5	<b>0.624</b>
avg	0.815	0.681	0.755	0.772	0.743	<b>0.838</b>

The largest value is shown in bold in each row

for different image sets, with the maximum value in each row duly marked.

From the table data, it is evident that the proposed method exhibits higher SSIM scores in seven image sets, with improvements of 2%, 15%, 8%, 6%, and 9% over the mean values from Reinhard et al., Pitie et al., Ueda et al., Xu et al., and Wu et al., respectively. This implies the superiority of our method in preserving image structure and details.

Notably, our approach excels at retaining edge information and offers higher fidelity in areas with complex textures. This adherence to the original image's structural similarity contributes to the elevated SSIM scores.

In conclusion, the experimental results clearly demonstrate our method's supremacy, especially when assessed using the SSIM metric. Additionally, our method is more computationally efficient and easy to employ than

deep-learning-based methods. These findings reaffirm our method's prowess in upholding image structure and intricacies, marking a significant advancement in the realm of image generation.

## 5 Conclusion

This paper proposes a method based on the idea of dividing the region for color transfer using saliency feature maps. Our experimental results show that this approach adeptly maintains the detailed features of the original image. Simultaneously, it imparts more vibrant and authentic colors and displays superior fidelity in complex texture regions.

Looking ahead, we aim to refine the current process, seeking to optimize steps without losing the existing accuracy. We are also invested in enhancing the precision of reference image color transfers through improved color space matching.

**Data availability** The code, data, and materials used or analysed during the current study are available from the corresponding author on reasonable request.

## References

- Xie, B., Zhang, K., Liu, K.: Color transfer based on color classification. *J. Phys.: Conf. Ser.* **1601**, 032028 (2020)
- Reinhard, E., Adhikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. *IEEE Comput. Graph. Appl.* **21**(5), 34–41 (2001)
- Pitié, F., Kokaram, A.C., Dahyot, R.: Automated colour grading using colour distribution transfer. *Comput. Vis. Image Understand.* **107**(1–2), 123–137 (2007)
- Pouli, T., Reinhard, E.: Progressive color transfer for images of arbitrary dynamic range. *Comput. Graph.* **35**(1), 67–80 (2011)
- Ueda, Y., Misawa, H., Koga, T., Suetake, N., Uchino, E.: Idt and color transfer-based color calibration for images taken by different cameras. *J. Adv. Comput. Intell. Intell. Inform.* **24**(1), 123–133 (2020)
- Hwang, Y., Lee, J.-Y., So Kweon, I., Joo Kim, S.: Color transfer using probabilistic moving least squares. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3342–3349 (2014)
- Grogan, M., Prasad, M., Dahyot, R.: L2 registration for colour transfer. In: *2015 23rd European Signal Processing Conference (EUSIPCO)*, IEEE, pp. 1–5 (2015)
- Grogan, M., Dahyot, R.: L2 divergence for robust colour transfer. *Comput. Vis. Image Understand.* **181**, 39–49 (2019)
- Wu, Z., Xue, R.: Color transfer with salient features mapping via attention maps between images. *IEEE Access* **8**, 104884–104892 (2020)
- Xu, D., Bao, S., Tanaka, G., Yang, C., Zuo, F.: A color calibration method based on color component projection for suppression of false color caused by iterative distribution transfer. *J. Adv. Comput. Intell. Intell. Inform.* **26**(1), 88–96 (2022)
- Gu, C., Lu, X., Zhang, C.: Example-based color transfer with gaussian mixture modeling. *Pattern Recogn.* **129**, 108716 (2022)
- Tai, Y.-W., Jia, J., Tang, C.-K.: Local color transfer via probabilistic segmentation by expectation-maximization. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, IEEE, vol. 1, pp. 747–754 (2005)
- Xiang, Y., Zou, B., Li, H.: Selective color transfer with multi-source images. *Pattern Recogn. Lett.* **30**(7), 682–689 (2009)
- Liu, D., Jiang, Y., Pei, M., Liu, S.: Emotional image color transfer via deep learning. *Pattern Recogn. Lett.* **110**, 16–22 (2018)
- Liu, S.: An overview of color transfer and style transfer for images and videos. *arXiv preprint arXiv:2204.13339* (2022)
- Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Commun. ACM* **60**(6), 84–90 (2017). <https://doi.org/10.1145/3065386>
- He, M., Liao, J., Yuan, L., Sander, P.V.: Neural color transfer between images. *arXiv preprint arXiv:1710.007562* (2017)
- Chen, T.Q., Schmidt, M.: Fast patch-based style transfer of arbitrary style. *arXiv preprint arXiv:1612.04337* (2016)
- Ulyanov, D., Vedaldi, A., Lempitsky, V.: Instance normalization: the missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022* (2016)
- Zhu, J.-Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2223–2232 (2017)
- Choi, Y., Choi, M., Kim, M., Ha, J.-W., Kim, S., Choo, J.: Stargan: unified generative adversarial networks for multi-domain image-to-image translation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8789–8797 (2018)
- Luan, F., Paris, S., Shechtman, E., Bala, K.: Deep photo style transfer. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4990–4998 (2017)
- Cheng, M.-M., Mitra, N.J., Huang, X., Torr, P.H., Hu, S.-M.: Global contrast based salient region detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(3), 569–582 (2014)
- Zou, A., Shen, X., Zhang, X., Wu, Z.: Neutral color correction algorithm for color transfer between multicolor images. In: *Advances in Graphic Communication, Printing and Packaging Technology and Materials: Proceedings of 2020 11th China Academic Conference on Printing and Packaging*. Springer, pp. 176–182 (2021)
- Kullback, S., Leibler, R.A.: On information and sufficiency. *Ann. Math. Stat.* **22**(1), 79–86 (1951)
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.