Contrast-Aware Color Consistency Correction for Multiple Images

Yinxuan Li, Li Li, Jian Yao, Menghan Xia, and Hanyun Wang

***Abstract*—Color consistency optimization for multiple images is a challenging task in remote sensing and computer vision. To ensure that the visual quality of corrected images is satisfactory, not only should the color discrepancies between multiple images be invisible but also the contrast of individual images should be visually appealing. Most color correction approaches focus on eliminating drastic color discrepancies, but ignore the problem of image contrast enhancement. Some color correction approaches even tend to degrade the image contrast to ensure that the image tones are consistent especially when the contrast of input images is low. To solve this problem, we present a contrast-aware color consistency correction approach in this article. We attempt to eliminate the drastic color differences and enhance the contrast of input images simultaneously. We creatively integrate the problems of color consistency correction and image contrast enhancement into the same global energy optimization framework, and we also design a special cost function to minimize the color discrepancies and enhance the image contrast using the original color information. Thus, although the contrast of input images is low, our approach can still generate the corrected images with consistent tones and visually appealing contrast. At last, we select several challenging datasets to evaluate our approach. The experimental results visually and quantitatively demonstrate the effectiveness and superiority of the proposed contrast-aware color consistency correction approach. The results also demonstrate that our approach significantly outperforms the existing approaches, especially when the contrast of the input images is low.**

***Index Terms*—Color consistency, color correction, image contrast, image mosaicking, multiple images.**

Manuscript received November 18, 2021; revised January 21, 2022, April 17, 2022, May 10, 2022, and May 31, 2022; accepted June 12, 2022. Date of publication June 15, 2022; date of current version June 27, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 42101440, in part by the Shenzhen Central Guiding the Local Science and Technology Development Program under Grant 2021Szvup100, and in part by the Open Research Fund Program of Key Laboratory of Digital Mapping and Land Information Application, Ministry of Natural Resources. *(Corresponding author: Li Li.)*

Yinxuan Li and Li Li are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China (e-mail: yinxuanli@whu.edu.cn; li.li@whu.edu.cn).

Jian Yao is with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China, and also with the AI Application and Innovation Research Center, The Open University of Guangdong, Guangzhou 510091, China (e-mail: jian.yao@whu.edu.cn).

Menghan Xia is with the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: menghanxyz@gmail.com).

Hanyun Wang is with the School of Surveying and Mapping, Information Engineering University, Zhengzhou 45000, China (e-mail: why.scholar@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2022.3183188

I. Introduction

In THE process of producing panorama [1], [2], digital orthophoto map (DOM) [3], and textured three-dimensional (3-D) model [4]–[6], we usually need to mosaic several images into a large image as seamlessly as possible. However, there usually are drastic color discrepancies between input images if they are captured by different sensors or at different times. The color inconsistency seriously degrades the visual quality of composite image and textured 3-D model. To solve this problem, we usually apply the image blending [7]–[9] methods to smooth the color differences between adjacent images. However, image blending can only eliminate the local and small color discrepancies, and it may fail when the color differences are global and drastic. Thus, it is necessary to perform color consistency correction for multiple images to ensure that the drastic and global color discrepancies have been eliminated before image mosaicking and texture mapping. Otherwise, the color artifacts will appear on the composite image and textured 3-D model. In this article, we focus on color consistency correction for multiview image mosaicking. It should be noted that the proposed approach can be used for another applications, too.

In recent years, many advanced color consistency correction approaches in the fields of remote sensing [10] and computer vision [11] have been proposed to eliminate drastic color discrepancies. Most of these approaches focus on how to extract the robust color correspondences and how to design the optimal color model and energy function. In most cases, existing approaches can effectively correct color differences and obtain corrected images with global consistent tones. However, their correcting results may not be visually pleasing. Because sometimes the contrast of the corrected images is low, which significantly degrades the visual quality. To visually illustrate this issue, we present an example in Fig. 1. The color discrepancies between input images are large and the contrast of each individual image is low, as shown in Fig. 1(a). We present the corrected images generated by two existing approaches [12], [13] in Fig. 1(b) and (c), respectively. Although the images presented in Fig. 1(b) and (c) have consistent global tones, the composite images are still unpleasant because the contrast of corrected images is still low. The scores of color distance (CD) and measure of enhancement (EME) presented in Fig. 1 also support the above conclusion. The definitions of CD and EME can be found in Section V-A.

However, most existing approaches focus on eliminating drastic color discrepancies, and ignore the issue of image contrast enhancement in their framework. In general, the color consistency

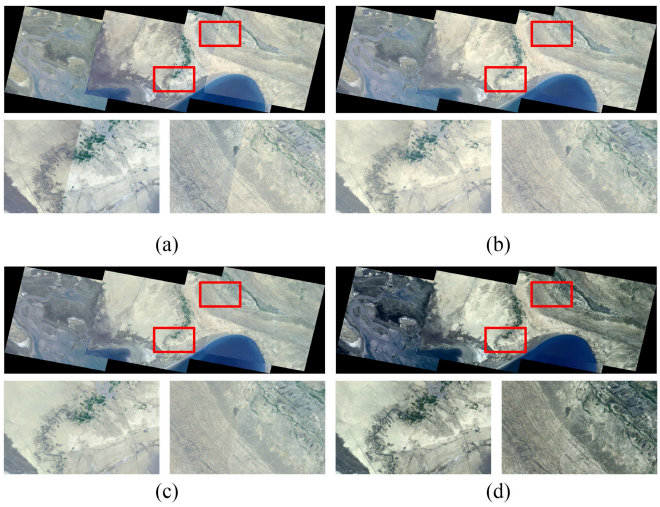

Fig. 1. Visual illustration of our work. The input images are presented in (a). Color correction results generated by Yu *et al.*'s approach [12], Xia *et al.*'s approach [13] and our contrast-aware approach are presented in (b)–(d), respectively. (a) Input images. (b) Yu *et al.*'s approach ($CD = 15.53$, EME $= 5.13$) ($CD = 7.10$, EME $= 4.80$). (c) Xia *et al.*'s approach. (d) Our proposed approach ($CD = 4.34$,EME $= 4.90$) ($CD = 4.05$, EME $= 14.88$).

term used in the existing methods [1], [12]–[14] is defined as the sum of color distances between images. Since blurred images usually lead to small color distances, the color consistency term tends to degrade the image contrast and generate blurred images. For two images, some advanced color transfer approaches [15]–[17] are proposed to preserve the original image details and transfer the color of target image into the reference image simultaneously. These approaches work well for two images, but they are difficult to be used in the case of multiple images. For multiple images, Li *et al.* [3] and Xia *et al.* [13] proposed to design a gradient term so as to preserve the original image details. But the gradient term cannot guarantee high visual quality of corrected images when original images are of low visual quality. To ensure visual quality of corrected images given low quality original images, Xia *et al.* [13] additionally designed a dynamic range term. But the designed dynamic range term only attempts to stretch color distribution ranges of corrected images based on that of original images. Most of the information contained in color distribution is ignored. So the ability of [13] to enhance image contrast is still insufficient, the visual quality of the corrected images may still be unpleasant, as shown in Fig. 1(c). In addition, Shen *et al.* [4] and Li *et al.* [18] proposed to perform image enhancement as preprocessing and postprocessing of global color correction, respectively. Since the enhancement is done after/before global color correction, global color optimization and high image quality cannot be achieved simultaneously. Postprocessing may corrupt global color consistency, while preprocessing cannot guarantee image quality after global color correction. Until now, it is still challenging to automatically generate the corrected images with both consistent tones and visually appealing contrast.

To simultaneously enhance the contrast of individual images and eliminate the color discrepancies, we propose a novel contrast-aware color consistency correction approach. Instead of performing image contrast enhancement and color consistency correction step by step, we creatively integrate these two problems into the same energy optimization framework. So we can simultaneously optimize these two problems together. We also design a special cost function that considers both color consistency term and image contrast term. Thus, we can generate the corrected images with consistent tones and visually appealing image contrast, as shown in Fig. 1(d). In our contrast term, we adaptively use the color information of input images to guide the process of image contrast enhancement. Inspired by the histogram equalization approach [19], we attempt to generate the corrected images with uniform histogram. However, the problems of overenhancement and underenhancement may appear if there are peaks and valleys in the original color histogram [20]–[24]. To alleviate these two problems, we propose to construct the color histogram using the contextual and gradient information of each pixel. In our proposed contrast-aware color correction approach, instead of preserving the image contrast of original images, we propose to enhance the image contrast. Thus, even if the contrast of original images is low, we can still generate the composite image with high quality.

II. Related Works

To solve the problem of color consistency correction, many approaches have been proposed in remote sensing [10] and computer vision [11] fields. Here, we divide these approaches into two categories: 1) path propagation-based and 2) global optimization-based approaches.

The intuitive method is to repeatedly correct the color differences for two images using the existing color transfer approaches. We name this kind of approaches as the path propagation-based approaches [25]–[27]. In general, they first manually or automatically select one or several images as the reference, then repeatedly transfer the colors of the rest of images into the reference along the transferring path. The path of each image usually is determined using a shortest path algorithm. Pan *et al.* [25] proposed a network-based color correction approach for multiple aerial image mosaicking. This method first finds the optimal transferring path for each target image from the graph constructed using the area Voronoi diagrams with overlap. Then, it transfers the color of each target image into the reference using the optimal transferring path. Xie *et al.* [27] proposed to select an image subset instead of one image from the weighted image graph as the reference images. The initial solution was obtained by transferring the colors of remaining images into the reference using a histogram matching method. They further improved the initial solution by optimizing a global energy function. The key advantage of the path propagation-based approaches is that they can fully use the current advanced color transfer approaches [15], [16], [28], [29] to correct the color differences between two adjacent images. However, there are two essential disadvantages in these approaches. The first one is that these approaches usually need to select one or several images from the input images as the reference. However, the automatic selection of the reference is still an open issue. The second one is that it is difficult to avoid the appearance of accumulated

errors, especially when the number of input images is large. Thus, these approaches work well for several images, but may generate inconsistent tones for a large set of images.

The global optimization-based approaches can effectively solve these two disadvantages. Instead of correcting the input images individually, they attempt to correct all images at the same time. In their approaches, they formulate the color consistency correction problem as an energy function optimization problem. They attempt to solve the color model parameters of all images globally and simultaneously. In general, we can divide the global optimization-based approach into two stages. First, the color model is designed to approximate the color discrepancies between input images. The commonly used color models include the linear model [1], [4], [12], [30], [31], gamma model [32], [33] and more flexible models like piecewise spline [13], [34] and combination model [3]. Second, an energy function is designed to globally solve the parameters of the applied color models for all images. The first global color correction approach is presented in [1]. They applied the linear model to correct the global color differences between panoramic images. Xiong and Pulli [32] presented a simple and effective luminance and color correction approach for panorama mosaicking on mobile phones. For luminance and color components, they applied the gamma and linear models to correct the luminance and color differences, respectively. The coefficients of the models are calculated using a global optimization process. Park *et al.* [33] proposed a robust low-rank matrix factorization approach to estimate the parameters of gamma models for input images. Liu *et al.* [14] proposed a robust color consistency optimization approach for multiple satellite images. In this method, inliers of color correspondences are applied to estimate the parameters of linear models. However, these approaches may fail to correct local color differences. To solve this problem, Yu *et al.* [12] proposed to eliminate the color differences globally and locally. This approach first eliminates the global color differences using the linear color model, then the remaining local color differences are smoothed using their proposed local optimization method. Zhang *et al.* [35] applied different linear models to approximate the color differences for different object categories. Similarly, they further applied a local edge optimization to smooth the artifacts along the category boundaries. Actually, the local optimization applied in [12] and [35] is a postprocessing and can be regarded as the variant of image blending approach.

In most cases, existing global optimization-based methods can automatically generate the corrected images with consistent tones. However, the visual quality of the corrected images may be still low because the image quality preservation is not considered in their energy function. To solve this problem, Shen *et al.* [4] proposed to enhance the contrast of the corrected images by linearly stretching the intensities of the corrected images. However, the linear contrast stretching applied in their approach is actually a postprocessing, which is not optimal. Xia *et al.* [13] designed the gradient and contrast terms to avoid the loss of original image details and penalize the dynamic range, respectively. To preserve the gradient information of input images, Li *et al.* [3] directly constrained that the gain of linear color models are equal to 1. However, their ability to enhance the image contrast is still low, the corrected images may be still unpleasant especially when the contrast of input images is low. In addition, the original color information also has not been considered in their quality terms.

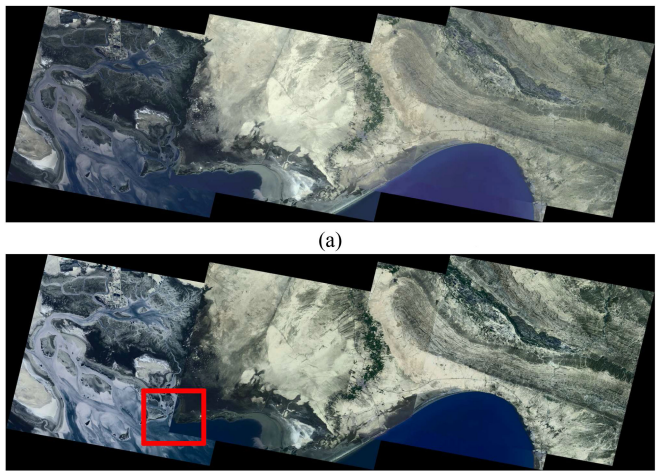

Fig. 2. Results generated by (a) image enhancement + color correction and (b) color correction + image enhancement. The red box highlights the color differences. (a) Image enhancement + Color correction ($CD = 2.25$, EME $=$ 11.80). (b) Color correction + Image enhancement ($CD = 13.70$, EME $= 16.38$).

III. Motivation of Our Work

Given a set of input images $\mathcal{I} = \{\mathbf{I}_i\}_{i=1}^{N}$, the goal of color consistency correction approaches is to generate the visually pleasant corrected images. The global tones of corrected images should be consistent and the contrast of individual images should be visually appealing. However, it is usually a tradeoff between color consistency optimization and image contrast enhancement.

To generate the images with consistent tones and visually appealing contrast, Shen *et al.* [4] proposed to perform image enhancement and color correction step by step. However, this kind of methods may generate unpleasant results. In Fig. 2, we present an example. In Fig. 2(a), we enhance the contrast of input images using our proposed improved histogram equalization method at first. Then, the color discrepancies between enhanced images are eliminated using our color correction approach. Although the global tones of corrected images are consistent, the contrast is low. The value of CD is small but the score of EME is also low. This is because the color consistency term tends to degrade the image contrast to eliminate the color discrepancies. In Fig. 2(b), we optimize the color consistency at first, then enhance the contrast of corrected images. The result is also unpleasant because the color differences between corrected images are visible. The score of EME is significantly improved but the value of CD also drastically increases. This is because the image contrast enhancement method may enlarge the color inconsistency of adjacent images. In addition, the contrast of the leftmost image is so high that the final result is disharmonious. Thus, it is more reasonable if we can simultaneously eliminate color discrepancies and enhance the image contrast. The corrected results will be pleasant if the image enhancement approach can

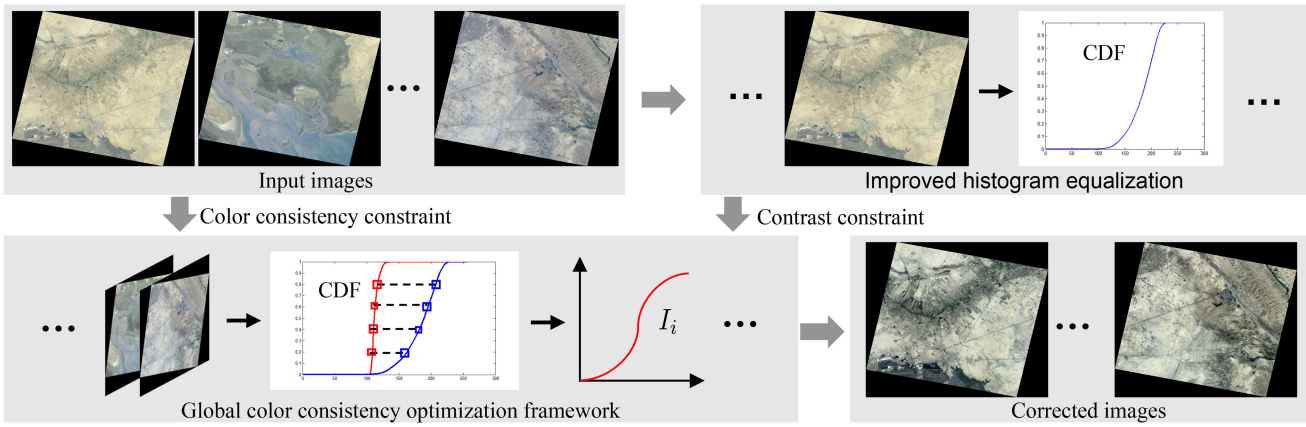

Fig. 3. Overview of our proposed contrast-aware color consistency correction approach.

be integrated into the framework of color consistency correction, especially when the contrast of input images is low.

The intuitive idea is to linearly enlarge the dynamic range of original images like the methods presented in [13]. The contrast of individual images will be enhanced using this method. However, this method does not fully consider the color distributions of input images. Histogram equalization approaches make full use of color distribution to enhance image. They believe the contrast of images will be enhanced if the color histograms of images are uniform. Inspired by this, we design a contrast term to encourage the generation of the corrected image with a uniform histogram. However, the problem of overenhancement or underenhancement may appear if there are prominent peaks or valleys in the histogram. The human visual system is very sensitive to the edge pixels, and is relatively insensitive to the smooth regions. For better visual quality, the contrast of edge pixels should be enhanced much more than the pixels in the smooth areas. Thus, the constructed histogram should consider the contextual information of each pixel. In addition, to emphasize the importance of edge pixels, the gradient information also should be considered. In our study, to produce appealing corrected images, we will propose an improved color histogram construction approach using the contextual and gradient information of each pixel.

As we know, the key of global optimization-based approaches [3], [13] is to estimate the optimal remapping function for each input image using the extracted color correspondences. We also observe that the essence of histogram-based image enhancement approaches [19], [20] is also to find an appropriate remapping function for each image using the information of color histogram. Thus, if we can apply the same remapping function to approximate the color discrepancies and the image contrast together, we can easily integrate them into the same global energy optimization framework.

IV. Proposed Color Consistency Correction Approach

As shown in Fig. 3, the proposed approach has two stages. First, for each individual image, we propose an improved histogram equalization approach to extract the weighted color histogram using contextual and gradient information. Then, we integrate the improved histogram equalization approach into the global color consistency optimization framework. Thus, we can simultaneously optimize the color consistency and enhance the image contrast.

A. Proposed Histogram Equalization Approach

In this section, we will introduce an improved histogram equalization approach. Consider an image $\mathbf{I} = \{I(\mathbf{p})\}$, where $I(\mathbf{p})$ is the pixel intensity of $\mathbf{p}$ at location (x, y), and $0 \leq I(\mathbf{p}) \leq R-1$. R denotes the number of intensity levels. In our study, we use the YCbCr color space. The contrast enhancement is only applied in the Y channel. Thus, the value of R is 256 in our study.

1) Traditional Histogram Equalization: In traditional histogram equalization approach, it first constructs the histogram $\mathbf{H} = \{h(r_k)|0 \leq k \leq R-1\}$, where $h(r_k)$ is the number of pixels with the intensity r_k. The probabilistic distribution of an intensity r_k in image $\mathbf{I}$ is calculated as

$$p(r_k) = h(r_k) \Big/ \sum_{k=0}^{R-1} h(r_k). \tag{1}$$

Based on the intensity probabilistic distribution, the corresponding cumulative distribution function (CDF) $C(r_k)$ is calculated according to the definition in [19]

$$C(r_k) = \sum_{j=0}^{k} p(r_j). \tag{2}$$

Last, the enhanced intensity e_k is obtained using the following transformation:

$$e_k = (R-1)C(r_k). \tag{3}$$

Namely, the CDF is regarded as the remapping function for image enhancement. However, the problems of overenhancement and underenhancement may appear if there are prominent peaks and valleys in the image histogram.

In Fig. 4, we present two visual examples. In the first example, the enhanced image generated by the traditional histogram

(a) (b) (c)

Fig. 4. Input images are presented in (a); (b) and (c) are the enhanced images generated by the traditional histogram equalization and our proposed improved approaches, respectively. For the first image, the overenhancement occurs at the smooth region of the ocean because of a large number of similar intensity pixels. For the second images, the underenhancement occurs at the region of the river (marked with red dashed ellipses) because the pixel number of corresponding intensity levels is small. (a) Input images. (b) Traditional histogram equalization. (c) Improved histogram equalization.

equalization is unpleasant, because the problem of overenhancement appears in the region of the ocean, as shown in the first row of Fig. 4. In addition, for the region of land, the detail of image is destroyed. In the second row of Fig. 4, we observe that the problem of underenhancement appears in the enhanced image generated by the traditional histogram equalization, especially in the region of the river. The problem of overenhancement also appears in the region of land.

2) Improved Histogram Equalization: To alleviate these problems, we present an improved histogram equalization approach using the contextual and gradient information. For each pixel $\mathbf{p} = (x, y)$, a $w \times w$ window is constructed at first. Let $\mathcal{N}_{w\times w}(\mathbf{p})$ denote all pixels in the window of pixel $\mathbf{p}$. The mean distance $m(\mathbf{p})$ between the pixel $\mathbf{p}$ and the neighboring pixel $\mathbf{q}$ is calculated as

$$m(\mathbf{p}) = \sum_{\mathbf{q}\in\mathcal{N}_{w\times w}(\mathbf{p})} (I(\mathbf{p}) - I(\mathbf{q}))/|\mathcal{N}_{w\times w}(\mathbf{p})\,| \tag{4}$$

where $|\mathcal{N}_{w\times w}(\mathbf{p})|$ denotes the number of pixels in the region $\mathcal{N}_{w\times w}(\mathbf{p})$. The neighborhood similarity $s(\mathbf{p})$ is defined as

$$s(\mathbf{p}) = 1 - \exp\left(-\frac{m(\mathbf{p})^2}{\sigma}\right) \tag{5}$$

where σ is the standard deviation. For the pixel $\mathbf{p}$ located at the smooth region, the value of $s(\mathbf{p})$ will be very small especially when $m(\mathbf{p}) < \sigma$. In addition, we also observe that the gradient magnitude of the smooth region is small, and the edge pixels have large gradient. Thus, to further increase the proportion of edge pixels, we define the gradient term $g(\mathbf{p})$ as

$$g(\mathbf{p}) = 1 - \exp\left(-\frac{G(\mathbf{p})^2}{\sigma}\right) \tag{6}$$

where $G(\mathbf{p})$ is the gradient magnitude of $\mathbf{p}$. For the edge pixels with large gradient, the value of $g(\mathbf{p})$ will also be large. Here, we find that there is a parameter σ need to be set empirically. As we know, the Gaussian function penalty a lot if $m(\mathbf{p})$ and $G(\mathbf{p})$ are smaller than σ. In most cases, the values of $m(\mathbf{p})$ and $G(\mathbf{p})$ are smaller than 10 in the smooth region. Thus, we set $\sigma = 10$ in our study, and it works well in most cases.

Based on the defined $s(\mathbf{p})$ and $g(\mathbf{p})$, we will redefine the image histogram $\mathbf{H}$. For each intensity r_k, the bin $h(r_k)$ is calculated as

$$h(r_k) = \sum_{\mathbf{p}\in\mathbf{I}} \eta\left(r_k, I(\mathbf{p})\right)\left(s(\mathbf{p}) + g(\mathbf{p})\right) \tag{7}$$

where $\eta(r_k, I(\mathbf{p}))$ is a binary logic function. $\eta(r_k, I(\mathbf{p})) = 1$ if $r_k = I(\mathbf{p})$, otherwise $\eta(r_k, I(\mathbf{p})) = 0$. Next, the probability distribution and CDF are calculated as the traditional histogram equalization does. In our improved histogram-based image enhancement approach, the contextual and gradient information have been considered. The proportion of the pixels in smooth regions will be suppressed, and the proportion of edge pixels will be increased. Thus, we can effectively avoid the appearance of artifacts caused by overenhancement and underenhancement. The enhanced images generated by our improved histogram-based approach are shown in Fig. 4(c). Obviously, the results generated by the improved approach are significantly better than the results of the traditional histogram equalization approach.

In Fig. 5, we visually present the intensity histograms constructed by the traditional and our proposed improved histogram equalization approaches. For the first image, in the traditional intensity histogram, the peak (marked by the red box) occurs at the intensity levels corresponding to the region of ocean. This is because the ocean region is smooth and the area is large. Thus, the intensity levels of the ocean region will be overstretched

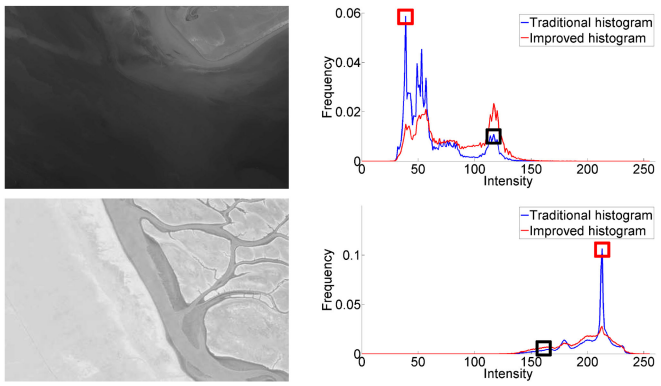

Fig. 5. Y channels of images and the corresponding traditional and improved intensity histograms.

according to (3), and the overenhancement occurs. However, in our improved intensity histogram (the red one), we successfully suppress the proportion of the ocean region as our expect. In addition, in our improved intensity histogram, we also improve the proportion of the land region (marked by the black box). Thus, we avoid destroying the image detail of the land region. For the second image, we observe that the proposed approach successfully suppress the peak (marked by the red box) of the traditional intensity histogram. Thus, we avoid the problem of overenhancement that occurs at the land region. In addition, we also observe that the proportion of the river region (marked by the black box) increases in the improved intensity histogram as our expect. Thus, the proposed approach can alleviate the problem of underenhancement that occurs at the river region.

B. Contrast-Aware Color Correction

In Fig. 2, we illustrate that the final corrected and enhanced images are not optimal if the image enhancement and color correction are performed step by step. In this section, we will introduce how to integrate these two problems into the same energy optimization framework.

1) Color Remapping Function: In essence, both color correction and image enhancement attempt to find an optimal color remapping function for each image. In the histogram-based image enhancement approaches, they regard the CDF as the transformation function. Although the CDF curve is flexible enough, it is nonparametric. It is difficult to globally solve the nonparametric color remapping functions for multiple images. Thus, we propose to apply the quadratic spline curve [13], [34] to approximate the color discrepancies and the image contrast enhancement. This curve is flexible enough and can be parameterized with several anchor points. Let M denote the number of anchors. We set $M = 6$ according to the suggestion presented in [13].

Let $\{(v_k, \widetilde{v}_k)\}_{k=1}^{M}$ denote the anchors of quadratic spline curve. $\{v_k\}_{k=1}^{M}$ are the evenly fixed horizontal coordinates. The horizontal coordinates indicate the intensity value of input image. $\{\widetilde{v}_k\}_{k=1}^{M}$ are the vertical coordinates which control the final shape of the remapping curve. For each input image $\mathbf{I}_i$, the color remapping function can be expressed as $f_i = \{\widetilde{v}_k^i\}_{k=1}^{M}$. The detailed introduction of this color transformation function can be found in our previous work [13].

2) Model Parameter Optimization: To generate the images with consistent tones and appealing contrast, we need to estimate the optimal remapping function f_i for each input image. The global energy function over all images should be designed at first, and then we need to solve it globally. It should be noted that the parameters of color transformation models are estimated in each channel independently. In our study, we apply the YCbCr color space. For each channel, the energy cost function E is formulated as

$$E = \sum_{\mathbf{I}_i \cap \mathbf{I}_j \neq \emptyset} E_{\text{color}}\left(\mathbf{I}_i, \mathbf{I}_j\right) + \sum_{i=1}^{N} \left(\lambda_1 E_{\text{regular}}\left(\mathbf{I}_i\right) + \lambda_2 E_{\text{contrast}}\left(\mathbf{I}_i\right)\right) \tag{8}$$

where $\mathbf{I}_i$ and $\mathbf{I}_j$ are two overlapped input images. λ_1 and λ_2 are the factors that balance the influence of regular and contrast terms. This cost function is simple, it only consists of color, regular, and contrast terms. For two overlapped images $(\mathbf{I}_i, \mathbf{I}_j)$, the color term is defined as

$$E_{\text{color}}\left(\mathbf{I}_i, \mathbf{I}_j\right) = \sum_{k=1}^{K_c} ||f_i\left(c_k^i\right) - f_j\left(c_k^j\right)||_2 \tag{9}$$

where (c_k^i, c_k^j) denotes a color correspondence between $\mathbf{I}_i$ and $\mathbf{I}_j$. K_c is the number of color correspondences. f_i and f_j denote the color remapping functions of $\mathbf{I}_i$ and $\mathbf{I}_j$, respectively. $||\bullet||_2$ denotes the L_2 norm. To extract the color correspondences from each image pair $(\mathbf{I}_i, \mathbf{I}_j)$, we first construct the CDFs for two images in the overlapped regions. Then, the corresponding intensities with the same probability in the two CDFs are regarded as the color correspondences. The probabilities are evenly fixed along the vertical axis of CDF. In our study, for each image pair, the number of color correspondences is 16. Namely, we set $K_c = 16$.

The color consistency term can effectively ensure that the color discrepancies between adjacent images are minimum. However, the color term tends to obtain the invalid solution, since the color discrepancies between multiple images are minimal if the intensities of all pixels are 0. Thus, to avoid the appearance of invalid solution, we define the regular term $E_{\text{regular}}(\mathbf{I}_i)$ of image $\mathbf{I}_i$ as

$$E_{\text{regular}}\left(\mathbf{I}_i\right) = \sum_{k=1}^{M} ||f_i\left(v_k^i\right) - v_k^i||_2. \tag{10}$$

The regular term encourages that the corrected image colors are similar to the original image colors. The valid solution can be obtained when the energy value of color and regular terms is minimum. Although the color differences between corrected images are invisible, the results may still be unpleasant if their contrast is low, as shown in Fig. 6. The scores of CD and EME also indicate that the corrected images have small color differences and low contrast. In addition, the problem of overblurring may appear. The color and regular terms cannot ensure that the visual quality of the corrected images is pleasant.

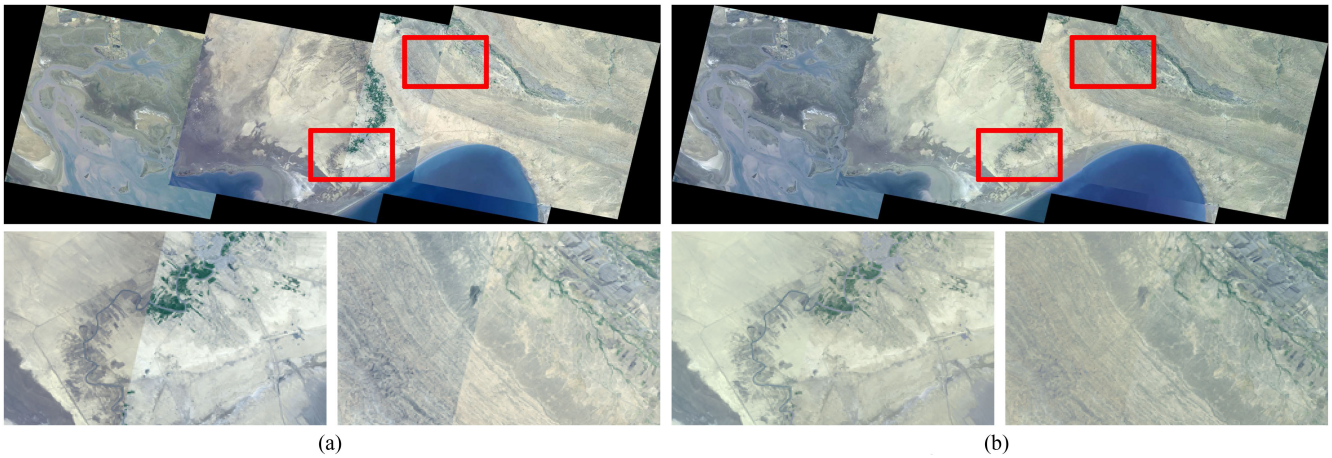

Fig. 6. Corrected images with the use of color and regular terms. The input and corrected images are presented in (a) and (b), respectively. The contrast of the input images is low and the color differences are large. Although the corrected images have the same global tones, the contrast is still low. (a) Input images ($CD = 15.53$, EME $= 5.13$). (b) Corrected images ($CD = 1.57$, EME $= 4.58$).

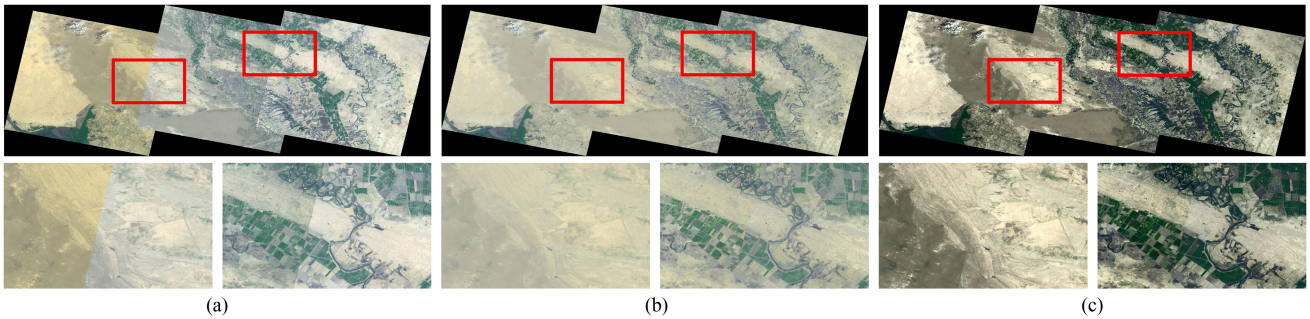

Fig. 7. Illustration of the effectiveness of our proposed contrast term. Input images are presented in (a). The corrected results without and with the use of contrast term are presented in (b) and (c), respectively. (a) Input images ($CD = 12.76$, EME $= 6.20$). (b) Without contrast term ($CD = 0.95$, EME $= 5.64$). (c) With contrast term ($CD = 1.07$, EME $= 13.64$).

To effectively improve the visual quality of the corrected images, some previous works [3], [13] apply the gradient or dynamic range term to avoid destroying the original image details. However, these approaches may fail if the visual quality of input images is low. In many cases, the contrast of input images is relatively low, especially for the satellite images. Thus, we propose to add the contrast term to enhance the image visual quality. For each input image $\mathbf{I}_i$, we first construct the CDF using the image histogram $\mathbf{H}_i$ defined in (7). Then, the corresponding intensities with the evenly fixed probabilities in the CDF are extracted as the contrast control intensities. Let $\{b_k^i\}_{k=1}^{K_b}$ denote these control intensities, where K_b denotes the number of control intensities. Let $\{p_k^i\}_{k=1}^{K_b}$ denote the corresponding probabilities of control intensities. In our study, we set $K_b = 16$. As described in Section IV-A, the histogram-based image enhancement approach tends to generate the image with uniform histogram using the constructed CDF. For each control intensity b_k^i, the enhanced intensity $\widetilde{b}_k^i$ is calculated as

$$\widetilde{b}_k^i = (R-1)\, p_k^i \tag{11}$$

where R is the total number of intensity levels. In the proposed contrast term, we attempt to ensure that the corrected image intensities are uniformly distributed. Thus, we define the contrast term as

$$E_{\text{contrast}}\left(\mathbf{I}_i\right) = \sum_{k=1}^{K_b} ||f_i\left(b_k^i\right) - \widetilde{b}_k^i||_2. \tag{12}$$

We encourage that the corrected image histograms are uniform, so the contrast of the corrected images will be enhanced. In addition, the original image color information is also fused into our contrast term because it is defined using the color histograms extracted from the original images. A visual example is shown in Fig. 7. From Fig. 7, we observe that the color differences of corrected images presented in Fig. 7(c) are slightly larger than the images presented in Fig. 7(b). This is because the contrast term inevitably limits the effectiveness of color consistency term. However, this compromise is worthwhile because the visual quality of the final corrected images is significantly improved. The scores of EME also indicate that the contrast term can significantly enhance the contrast of the input images. It should be noted that the contrast term is applied to luminance channel Y only. The Y represents nonchromatic information of images, and the chromatic information is included in the Cb

and Cr channels. Thus, we can preserve the original chromatic information as much as possible.

In addition, we also add the constraint of monotonic increasing for the spline line as our previous work [13] does. Finally, the color model parameters can be calculated by minimizing the energy function E. E is a quadratic polynomial, which can be transformed to the standard form of constrained quadratic programming. It can be solved using a convex quadratic programming algorithm. Detailed descriptions of the solving process can be found in Appendix A.

V. Experimental Results and Discussion

To evaluate the performance of the proposed contrast-aware color consistency correction approach, we selected four datasets (Coast, River, Land, and Village) captured by different sensors to conduct our experiments. It should be noted that the images included in each dataset have been geometrically aligned before performing color correction. The Coast, River, and Land datasets are orthoimages, so the images included in these three datasets already be aligned into the same coordinate system. For Village dataset, we directly apply the software *PTGui*[1] to perform geometrical alignment. The Coast, River, and Land datasets are multitemporal satellite images. Thus, their contrast is relatively low and the color differences between the original images are large. The Village dataset is captured by the unmanned aerial vehicle (UAV) camera. The Village dataset is a simulated dataset. In this dataset, the tones of images are manually adjusted to increase the challenge of color consistency correction. And the contrast of images are manually turned down to verify the effectiveness of our contrast-aware color correction approach.

A. Evaluation Metrics

In this study, we select two objective metrics to quantitatively evaluate the performance of color correction approaches. The first metric is the color distance (CD) [13]. It is used to calculate the color differences between corrected images. The other is the measure of enhancement (EME) [36]. It is applied to evaluate the contrast of individual images. It should be noted that the CD is computed in YCbCr color space, and the EME is computed using luminance component only.

1) Color Distance: For two overlapped images $\widehat{\mathbf{I}}_i$ and $\widehat{\mathbf{I}}_j$, two color histograms are extracted from the overlapped regions of two images. Then, the value of this metric is defined as the distance between these two color histograms. The smaller value of CD indicates the better color consistency. This metric is calculated as

$$CD = \sum_{\widehat{\mathbf{I}}_i \cap \widehat{\mathbf{I}}_j \neq \emptyset} w_{ij} \frac{\Delta H\left(\widehat{\mathbf{I}}_{ij}, \widehat{\mathbf{I}}_{ji}\right)}{N_b} \tag{13}$$

where $\widehat{\mathbf{I}}_i$ and $\widehat{\mathbf{I}}_j$ are two corrected images with overlap. $\widehat{\mathbf{I}}_{ij}$ represents the region of $\widehat{\mathbf{I}}_i$ overlapped with $\widehat{\mathbf{I}}_j$. There is a similar meaning for $\widehat{\mathbf{I}}_{ji}$. The weight w_{ij} is set proportional to the area of the overlapped region, and $\sum w_{ij} = 1$. $\Delta H(\bullet)$ denotes the bin-to-bin distance between two color histograms extracted from $\widehat{\mathbf{I}}_{ij}$ and $\widehat{\mathbf{I}}_{ji}$. N_b is the bin number of the histogram.

2) Measure of Enhancement: This metric computes the average local contrast in an individual image. The higher value of EME indicates the better contrast. For each corrected image $\widehat{\mathbf{I}}_i$, we divide it into many small nonoverlapping blocks with the size of $w_1 \times w_2$. Let N_p denote the number of blocks. The EME is defined as

$$\text{EME} = \frac{1}{N_p} \sum_{k=1}^{N_p} 20 \times \log\left(\frac{\max\left(\widehat{\mathbf{I}}_{i,k}\right)}{\min\left(\widehat{\mathbf{I}}_{i,k}\right)}\right) \tag{14}$$

where $\widehat{\mathbf{I}}_{i,k}$ denotes the kth block of image $\widehat{\mathbf{I}}_i$. $\max(\bullet)$ and $\min(\bullet)$ denote the maximum and minimum intensities of the block, respectively.

B. Parameter Determination

There are two key parameters (λ_1 and λ_2) that need to be balanced in our proposed approach. λ_1 and λ_2 are applied to balance the influence of regular and contrast terms, respectively. We select four adjacent images from Coast dataset to visually and quantitatively illustrate the influence of λ_1 and λ_2. We named the selected dataset as Coast4. The values of CD and EME of the selected four images are 15.53 and 5.13, respectively.

In Fig. 8, we presented the corrected results with the use of different λ_1. Here, to effectively illustrate the influence of λ_1, we set λ_2 to 0. As the value of λ_1 increases, the contrast of the corrected images increases, but the color differences also increase. This is because the regular term will limit the flexibility of the applied color transformation model. Since the color consistency term tends to generate the blurred images, the contrast of corrected images generated with small λ_1 is low, as shown in Fig. 8(a). However, even if the value of λ_1 is large, the contrast of the corrected images is still lower than the original images, as shown in Fig. 8(d). Although the small λ_1 will generate the images with small color differences, the corrected results may lose the original color information. Thus, to preserve the original chromatic information as much as possible, we set $\lambda_1 = 0.5$ for Cb and Cr channels. For the Y channel, we suggest to set λ_1 as a small value. Although the small λ_1 will drastically degrade the contrast of Y channel, the contrast term presented in our approach will significantly enhance the image contrast of Y channel. Thus, in order to eliminate the luminance differences between images as much as possible, we set $\lambda_1 = 0.1$ for Y channel.

In Fig. 9, we presented the corrected results with the use of different λ_2. We fixed $\lambda_1 = 0.5$ for Cb and Cr channels, and fixed $\lambda_1 = 0.1$ for Y channel. Although the small value of λ_2 is applied, the contrast of corrected images is still better than the original contrast, as shown in Fig. 9(a). In addition, the color differences between corrected images are still very small. It effectively illustrates that our proposed contrast-aware color consistency correction approach can eliminate the color discrepancies and enhance the image contrast simultaneously. As the value of λ_2 increases, the image contrast increases drastically,

[1] [Online]. Available: https://www.ptgui.com/

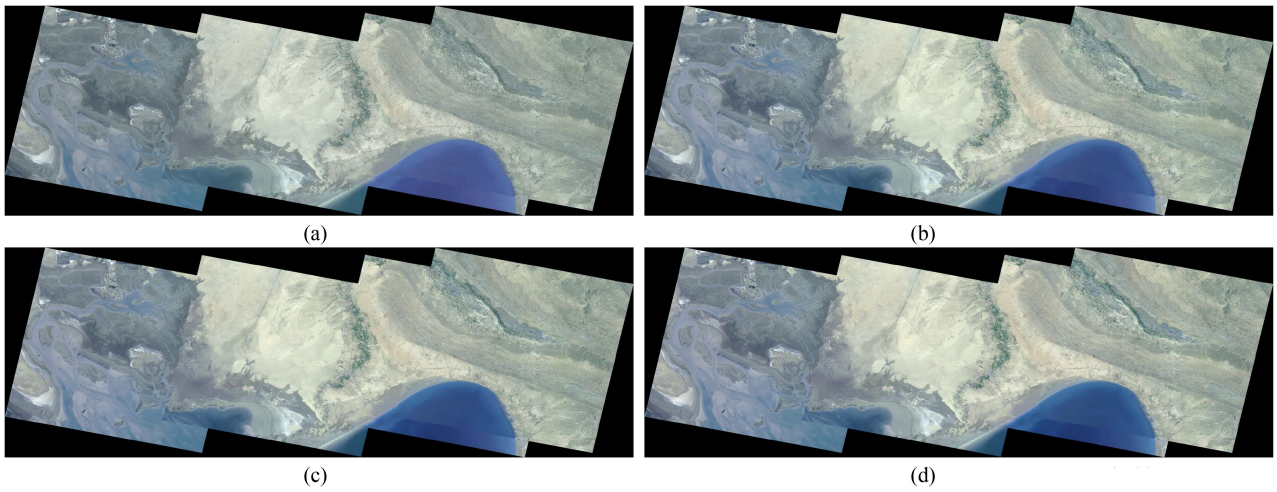

Fig. 8. Corrected results of our approach with the use of different balance weight λ_1. (a)–(d) are the results generated with the use of different λ_1. (a) $\lambda_1 = 0.1$ ($CD = 0.94$, EME $= 4.37$). (b) $\lambda_1 = 0.3$ ($CD = 1.27$, EME $= 4.52$). (c) $\lambda_1 = 0.5$ ($CD = 1.57$, EME $= 4.58$). (d) $\lambda_1 = 0.7$ ($CD = 1.95$, EME $= 4.67$).

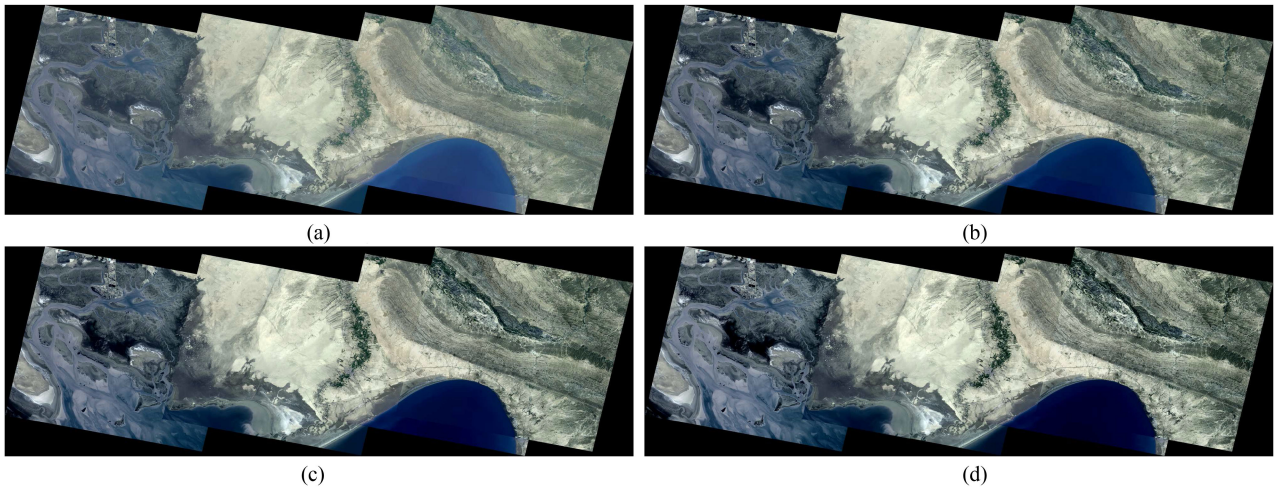

Fig. 9. Corrected results of our approach with the use of different balance weight λ_2. (a)–(d) are the results generated with the use of different λ_2. (a) $\lambda_2 = 0.1$ ($CD = 1.41$, EME $= 6.99$). (b) $\lambda_2 = 0.3$ ($CD = 2.43$, EME $= 11.51$). (c) $\lambda_2 = 0.5$ ($CD = 4.05$, EME $= 14.88$). (d) $\lambda_2 = 0.7$ ($CD = 5.77$, EME $= 19.18$).

but the color differences also increase. This is because the contrast term applied in our approach also limits the effectiveness of color term. We observed that $\lambda_2 \in [0.3, 0.5]$ can balance the color consistency correction and image contrast enhancement well. Thus, we suggest to set $\lambda_2 \in [0.3, 0.5]$. In our experiments, we set $\lambda_2 = 0.5$ in default.

To further clearly illustrate how these two parameters influence the correction results of the proposed approach, we presented the curves of CD and EME offered by different λ_1 and λ_2 in Fig. 10. In this experiment, we also applied the Coast4 dataset. From Fig. 10, we can obtain the similar conclusion with the qualitative evaluation presented in Figs. 8 and 9.

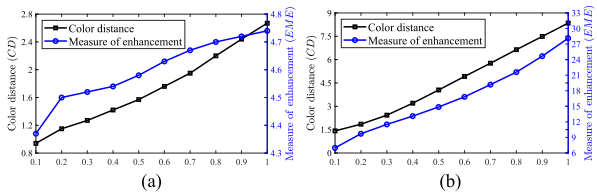

Fig. 10. Curves of CD and EME offered by different (a) λ_1 and (b) λ_2. The Coast4 dataset is applied in this experiment. (a) λ_1. (b) λ_2.

C. Ablation Study

To illustrate the effectiveness of our proposed contrast term, we performed an ablation study to illustrate how the proposed contrast term influences the correction results. We selected Coast, River, Land, and Village datasets to perform the ablation study. We reported the quantitative ablation study in Table I. From Table I, we found that the color differences between corrected images slightly increase after using the contrast term. However, the contrast of corrected images is significantly improved with the use of the proposed contrast term.

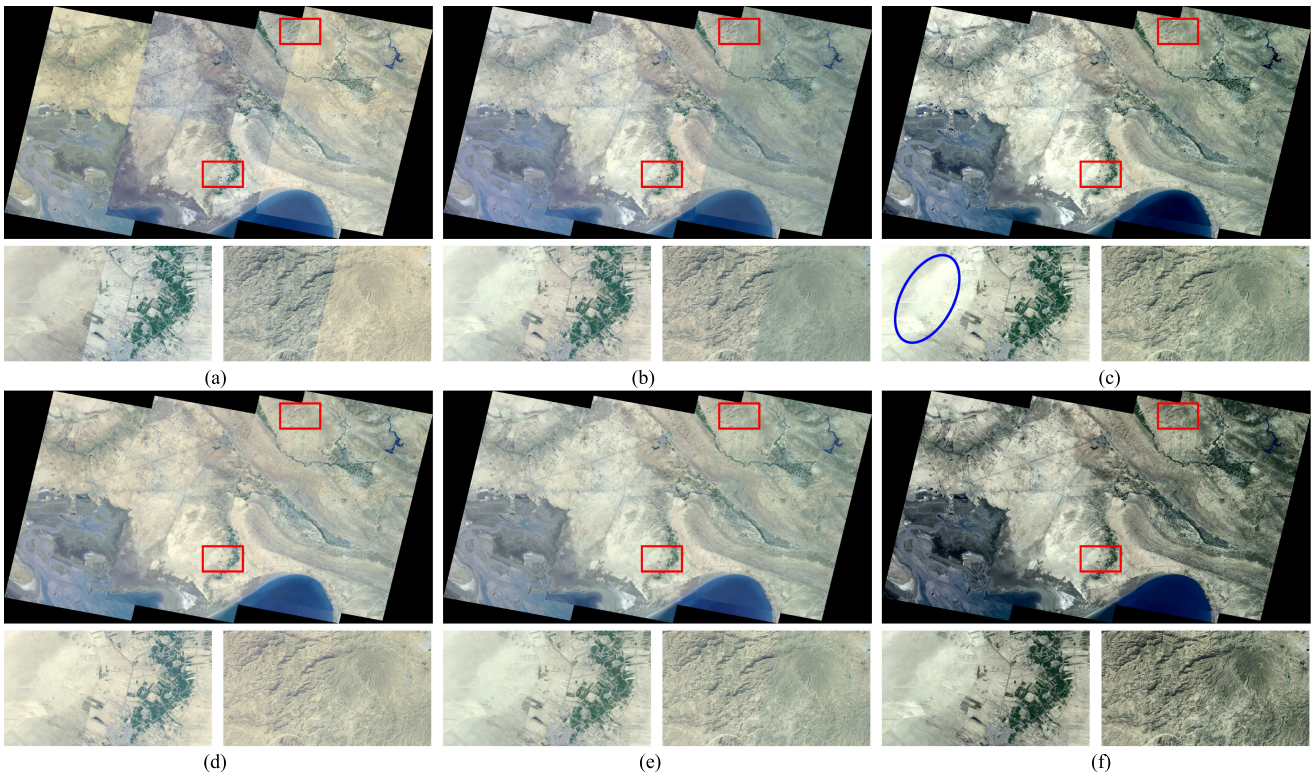

Fig. 11. Comparative experiment on Coast dataset. (a) is the input images. (b)–(f) are the color consistency correction results generated by [1], [4], [12], [13] and our approach, respectively. (a) Input images. (b) Brown and Lowe's approach [1]. (c) Shen *et al.*'s approach [4]. (d) Yu *et al.*'s approach [12]. (e) Xia *et al.*'s approach [13]. (f) Our approach.

TABLE I
QUANTITATIVE EVALUATION RESULTS OF OUR APPROACH WITH AND WITHOUT THE USE OF CONTRAST TERM

Dataset	Input		Without contrast term		With contrast term	
	CD	EME	CD	EME	CD	EME
Coast	12.62	6.72	1.66	5.50	2.82	16.12
River	9.81	6.12	1.75	4.80	2.84	13.27
Land	20.40	14.87	1.98	8.87	2.98	24.84
Village	30.89	12.11	1.29	6.96	2.12	17.30

D. Comparative Experiments

We selected four existing color correction approaches [1], [4], [12], [13] to conduct the comparative experiments. In these four approaches, two approaches [4], [12] have a postprocessing. Shen *et al.* [4] proposed to linearly stretch the intensities of the images after the color consistency correction to enhance the image contrast. In the comparative experiments, we keep this postprocessing to emphasize that our contrast-aware color correction approach can generate the corrected images with consistent tones and visually appealing contrast. Yu *et al.* [12] proposed to perform a local color optimization after global color correction. However, their local optimization actually is a variant of image blending approach. It is not fair to other approaches if the local optimization is applied, because the image blending is not performed in other approaches. Thus, we abandon this postprocessing in the next experiments. The Coast, River, Land, and Village datasets are selected to conduct the comparative experiments.

In Fig. 11, we conducted the comparative experiment on Coast dataset. We observed that the corrected images presented in Fig. 11(b) still have large global color differences between adjacent images, especially between the third and the last strips. Although Yu *et al.*'s [12] and Xia *et al.*'s [13] approaches successfully eliminate the large color differences between adjacent strips, the contrast of the corrected images is low, as shown in Fig. 11(d) and (e). Thus, the visual quality of their results is still unpleasant. The contrast of the corrected images presented in Fig. 11(c) is relatively high. However, there are two problems in their results. First, the problem of overenhancement appears in the final images, as shown in Fig. 11(c). The area of overenhancement is highlighted by the blue ellipse. This area is so bright that some image details are lost. Second, the color differences between the third and the last strips are still visible, as shown in the second enlarged region of Fig. 11(c). In addition, the contrast of the corrected images presented in Fig. 11(c) is also lower than that of our proposed approach. For example, there is a region of mountain in the second enlarged region. The texture of this region is rich. From the comparison between 11(c) and 11(f), we observed that the texture of our corrected images is more clear. Fig. 11(f) presents the corrected images generated by our proposed approach. The global tones of corrected images are consistent while the contrast of individual images is appealing.

Then, we tested all approaches on River dataset, as shown in Fig. 12. Similar to the Coast dataset, it consists of several

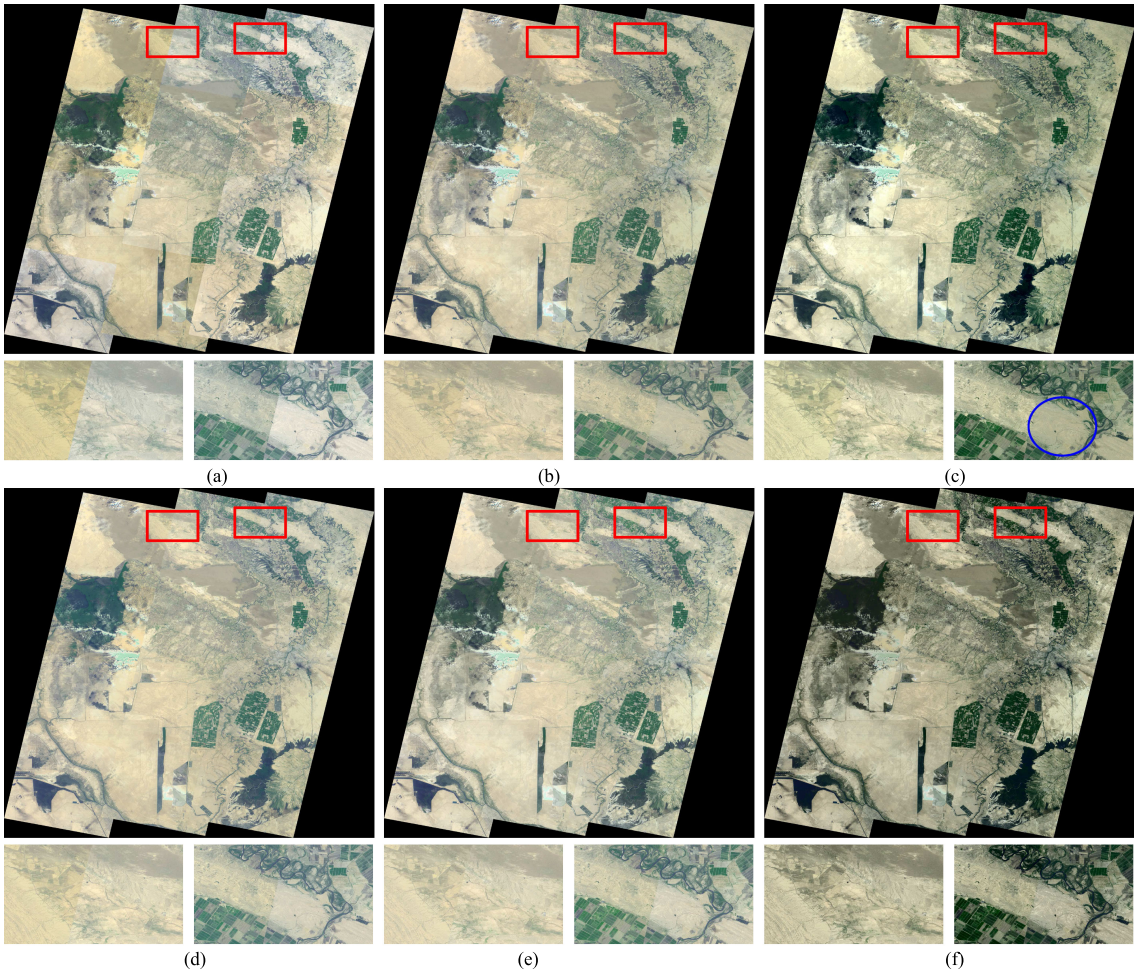

Fig. 12. Comparative experiment on River dataset. The input images are shown in (a). (b)–(f) are the color consistency correction results generated by [1], [4], [12], [13] and our approach, respectively. (a) Input images. (b) Brown and Lowe's approach [1]. (c) Shen *et al.*'s approach [4]. (d) Yu *et al.*'s approach [12]. (e) Xia *et al.*'s approach [13]. (f) Our approach.

multitemporal satellite images. The color differences of this dataset is relatively small. We observed that all approaches can effectively eliminate the global color differences of the input images. However, for some local regions, we found that Xia *et al.*'s approach [13] and the proposed approach outperform other approaches. In the first enlarged region, the color differences of images presented in 12(e) and (f) are smaller than that of images presented in Fig. 12(b)–(d). Similar with the Coast dataset, the contrast of the images presented in Fig. 12(b)–(e) is low. The corrected images of these approaches are blurred in some areas. Shen *et al.*'s approach [4] effectively enhances the contrast of corrected images. However, it still fails to enhance some local details. In Fig. 12(c), the image region highlighted by the blue ellipse is still blurred. In contrast, our approach effectively enhances the local details in this highlighted region, as shown in the Fig. 12(f).

Next, we evaluated all approaches on Land dataset, as shown in Fig. 13. From Fig. 13(b)–(d), we observed that the corrected images still have large local color differences, especially in the second enlarged region. In addition, the contrast of images presented in Fig. 13(b) and (d) is low. We also observed that the corrected images offered by Shen *et al.*'s approach have the highest contrast. However, the problem of overenhancement appears again, as shown in the regions marked by the blue ellipses. Xia *et al.*'s approach [13] offers the better color consistency optimization results, but the contrast of the corrected images is relatively low. The proposed approach performs well on this dataset. We effectively eliminate the large color differences between the input images. Although the image contrast of our results is not the highest, we successfully avoid the problem of overenhancement.

Finally, we tested all approaches on Village dataset. Both tone and contrast of each image in this dataset have been adjusted manually beforehand to increase the challenge. As a result, the color discrepancies are large and the contrast is low, as shown in Fig. 14(a). In addition, the geometrical misalignment between images is large. From Fig. 14, we observed that our approach offers the best result, and significantly outperforms the other approaches in both color consistency and image visual quality. The corrected images presented in Fig. 14(b) suffer from severe image blurring and color inconsistency. Although Shen *et al.*'s approach [4] generates the images with high contrast, the color differences of the corrected images are still large. In addition, the problem of overenhancement appears, as shown in

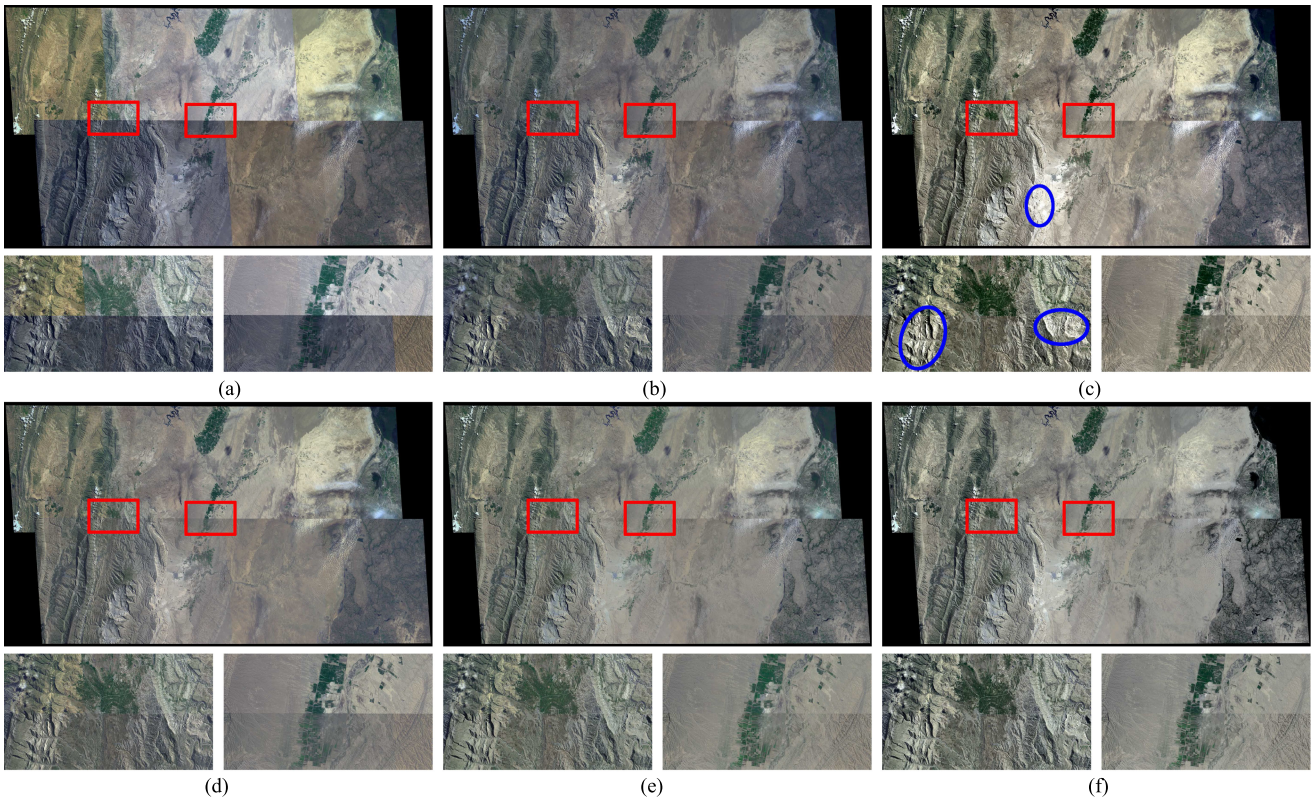

Fig. 13. Comparative experiment on Land dataset. The input images are shown in (a). (b)–(f) are the color consistency correction results generated by [1], [4], [12], [13] and our approach, respectively. (a) Input images. (b) Brown and Lowe's approach [1]. (c) Shen *et al.*'s approach [4]. (d) Yu *et al.*'s approach [12]. (e) Xia *et al.*'s approach [13]. (f) Our approach.

TABLE II
QUANTITATIVE EVALUATION OF THE DIFFERENT APPROACHES

Dataset	Input			Approach 1			Approach 2			Approach 3			Approach 4			Our approach		
	CD	EME	$T(s)$	CD	EME	$T(s)$	CD	EME	$T(s)$	CD	EME	$T(s)$	CD	EME	$T(s)$	CD	EME	$T(s)$
Coast	12.62	6.72	–	4.91	6.73	8.34	3.71	11.68	10.02	4.37	5.91	7.88	2.96	6.66	10.26	**2.82**	**16.12**	18.26
River	9.81	6.12	–	3.96	6.14	12.73	4.41	10.60	15.85	4.01	5.64	12.98	**2.73**	6.25	20.85	2.84	**13.27**	24.85
Land	20.40	14.87	–	6.92	15.01	8.54	9.64	**29.65**	11.80	6.82	11.87	9.13	3.09	13.32	12.31	**2.98**	24.84	12.89
Village	30.89	12.11	–	10.10	12.26	2.67	10.81	16.69	3.38	4.84	8.74	2.67	**1.90**	8.81	3.62	2.12	**17.30**	4.61

 # Approach 1–4 denote the approaches presented in [1], [4], [12] and [13], respectively. T denotes the computational time, and the unit is seconds.

the second enlarged region. We used the blue ellipse to highlight the area of overenhancement. The global tones of the corrected images presented in Fig. 14(d) and (e) are consistent. However, there still exist some local color differences. In addition, the contrast of images presented in Fig. 14(d) and (e) is low. The existing approaches all fail to generate the pleasant results for this challenging dataset. However, our approach still works well, as shown in Fig. 14(f).

From the experimental results presented in Fig. 11–14, we observed that the proposed approach works well and outperforms the existing approaches. However, we also observed that there still exist small color differences between the images corrected by the proposed approach. This problem usually appears when the color differences between the input images are drastic and complex. The color consistency correction methods only can eliminate most of color differences, there usually are small color discrepancies that remained between adjacent images. In practical applications, the image blending methods [7]–[9] are usually applied to further smooth such small color differences.

E. Quantitative Evaluation

The quantitative evaluation is also conducted to convincingly illustrate the superiority of the proposed contrast-aware approach. We applied the two metrics presented in Section V-A to quantitatively evaluate the performance of all approaches, as shown in Table II. The proposed contrast-aware approach offers the best CD scores for Coast and Land datasets. For River and Village datasets, the CD score of our approach is only slightly lower than the score offered by Xia *et al.*'s approach [13]. However, the EME score of our approach is significantly higher than the score provided by Xia *et al.*'s approach [13]. In addition, we also found that the EME scores offered by the proposed

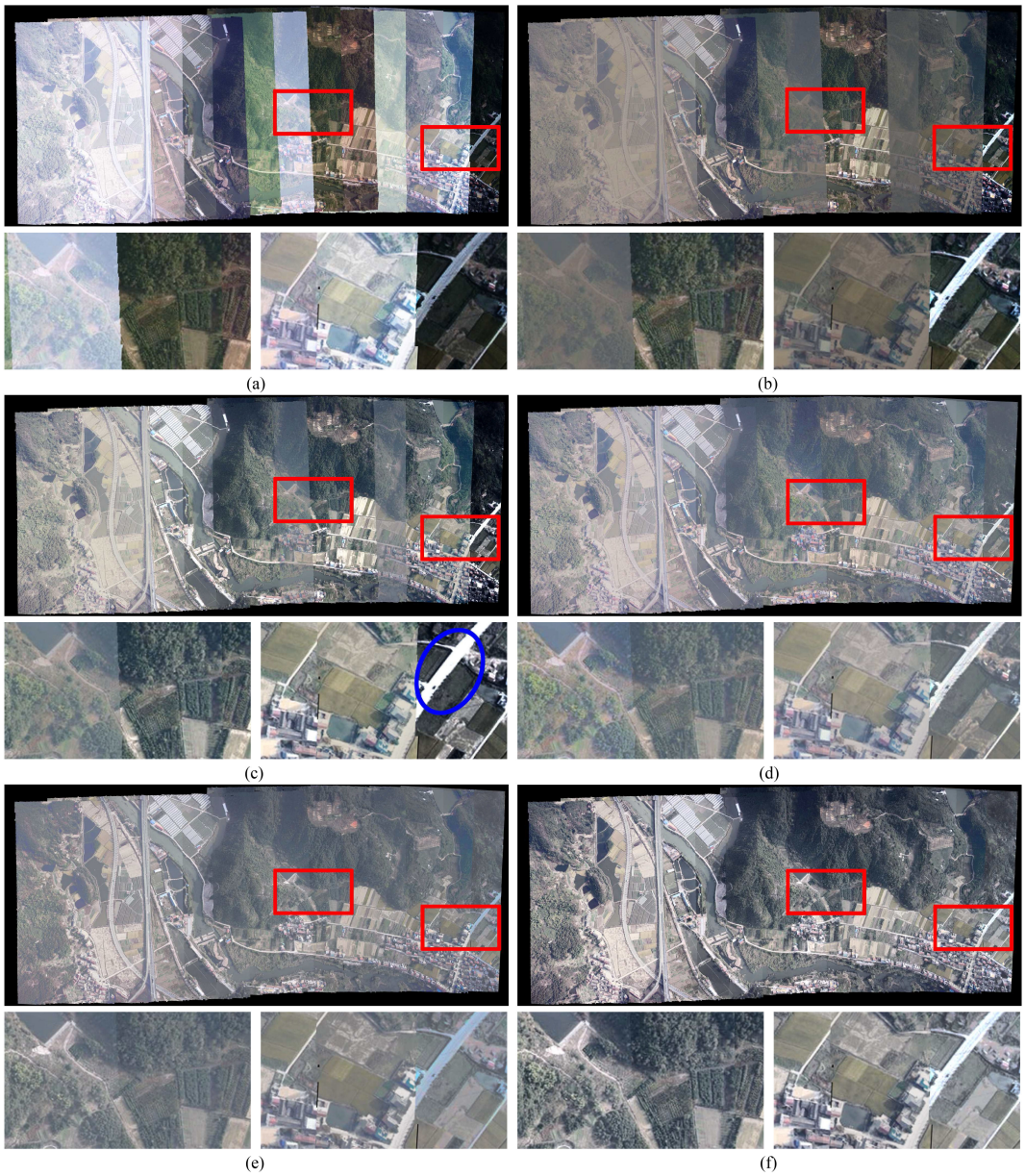

Fig. 14. Comparative experiment on Village dataset. The input images are shown in (a). (b)–(f) are the color consistency correction results generated by [1], [4], [12], [13] and our contrast-aware approach, respectively. (a) Input images. (b) Brown and Lowe's approach [1]. (c) Shen *et al.*'s approach [4]. (d) Yu *et al.*'s approach [12]. (e) Xia *et al.*'s approach [13]. (f) Our approach.

contrast-aware approach are significantly higher than other approaches in Coast, River, and Village datasets. For Land dataset, Shen *et al.*'s approach [4] offers the best score of EME. However, the problem of overenhancement appears in the images corrected by Shen *et al.*'s approach, as shown in Fig. 13(c). In addition, the CD score of our approach is significantly better than that of Shen *et al.*'s approach. Considering the CD and EME metrics together, the proposed contrast-aware approach is the best among these approaches.

In Table II, we also presented the computational time of these approaches. We observed that our approach is relatively time-consuming. The computational time of the proposed contrast-aware approach is longer than that of other approaches in all datasets. This is because the construction of the color histogram using the contextual and gradient information is time-consuming. Although the efficiency of our approach is lower than the existing approaches, the results generated by our approach are visually appealing and significantly better than the current approaches.

VI. Conclusion

In this work, we present a novel contrast-aware color consistency correction approach for multiple images. In the proposed approach, not only can the color differences between input images be effectively eliminated but also the contrast of input images can be enhanced. The contributions of our work are summarized as follows.

1) We creatively integrate the problems of color consistency optimization and image contrast enhancement into the same global energy optimization framework. Thus, we can generate the corrected images with consistent tones and visually appealing contrast even the contrast of input images is low.
2) Instead of directly penalizing the dynamic range of input images, we apply the original color information extracted from the color histogram to adaptively define our contrast term. In addition, to alleviate the problems of over- and underenhancement, we also propose to construct the color histogram using the contextual and gradient information.
3) The experimental results on several real and simulated datasets illustrate the superiority and effectiveness of our approach, and also illustrate that the proposed approach significantly outperforms the existing state-of-the-art color consistency correction approaches.

To the best of our knowledge, it is the first work that attempts to simultaneously eliminate the drastic color differences between multiple images and enhance the contrast of individual images using the original color histogram information. It provides a new idea to solve the visual quality degradation problem existed in the current color consistency correction approaches.

Appendix A
Solving Process of Our Energy Function

To solve color parameters for all images, we are supposed to write (8) in matrix form. To this end, for x falls in the scope of knots $\{v_p^i, v_{p+1}^i, v_{p+2}^i\}$, following formulation of spline model in [13], color remapping $f_i(x)$ can be written in matrix form as

$$f_i(x) = \frac{1}{2}\underbrace{\begin{bmatrix}0\cdots 0 & t_1 & t_2 & t_3 & 0\cdots 0\end{bmatrix}}_{M}\begin{bmatrix}\widetilde{v}_1^i\\ \widetilde{v}_2^i\\ \vdots\\ \widetilde{v}_M^i\end{bmatrix} \tag{15}$$

where $t_1 = 1 - 2t + t^2$, $t_2 = 1 + 2t - t^2$, and $t_3 = t^2$. $\{\widetilde{v}_k^i\}_{k=1}^M$ are parameters of f_i and t can be computed easily by solving

$$x = \frac{1}{2}\left[t_1\nu_p^i + t_2\nu_{p+1}^i + t_3\nu_{p+2}^i\right]. \tag{16}$$

Let $\boldsymbol{v}$ denote a vector of color parameters for all images

$$\boldsymbol{v} = [\widetilde{v}_1^1, \widetilde{v}_2^1, \ldots, \widetilde{v}_M^1, \ldots, \widetilde{v}_1^N, \widetilde{v}_2^N, \ldots, \widetilde{v}_M^N]^\top \tag{17}$$

which contains all the variables to be solved. Then we can write (15) as

$$f_i(x) = \frac{1}{2}\underbrace{\begin{bmatrix}0\cdots 0 & t_1 & t_2 & t_3 & 0\cdots 0\end{bmatrix}}_{M\times N}\boldsymbol{v}. \tag{18}$$

Learning from (18), (9) can be rewritten as

$$E_{\text{color}}(\mathbf{I}_i, \mathbf{I}_j) = (\mathrm{A}_{i,j}\boldsymbol{v})^\top \mathrm{A}_{i,j}\boldsymbol{v} \tag{19}$$

where $\mathrm{A}_{i,j}$ is $K_c \times (M \times N)$ sparse matrix and each row of $\mathrm{A}_{i,j}$ relates to one color correspondence between image $\mathbf{I}_i$ and image $\mathbf{I}_j$. Similarily, (10) can be rewritten as

$$E_{\text{regular}}(\mathbf{I}_i) = (\Gamma_i\boldsymbol{v} - \boldsymbol{v}_0^i)^\top(\Gamma_i\boldsymbol{v} - \boldsymbol{v}_0^i) \tag{20}$$

since $f_i(v_k^i) = \widetilde{v}_k^i$, where Γ_i is $(M \times N) \times (M \times N)$ sparse matrix with only six elements on the diagonal being 1 and $\boldsymbol{v}_0^i = [v_1^i, \ldots, v_M^i]^\top$. And (12) can be rewritten as

$$E_{\text{contrast}}(\mathbf{I}_i) = (\mathrm{B}_i\boldsymbol{v} - \boldsymbol{b}_i)^\top(\mathrm{B}_i\boldsymbol{v} - \boldsymbol{b}_i) \tag{21}$$

where B_i is $K_b \times (M \times N)$ sparse matrix and $\boldsymbol{b}_i = [\widetilde{b}_2^i, \widetilde{b}_2^i, \ldots, \widetilde{b}_{K_b}^i]^\top$. According to (19)–(21), we, thereby can derive the matrix form of (8)

$$E = (\mathrm{A}\boldsymbol{v})^\top\mathrm{A}\boldsymbol{v} + (\boldsymbol{v} - \boldsymbol{v}_0)^\top(\boldsymbol{v} - \boldsymbol{v}_0) + (\mathrm{B}\boldsymbol{v} - \boldsymbol{b})^\top(\mathrm{B}\boldsymbol{v} - \boldsymbol{b}) \tag{22}$$

where A, B, $\boldsymbol{v}_0$, and $\boldsymbol{b}$ can be obtained by continuously stacking matrix and vector respectively. Finally, rewriting hard constraints imposed on $\boldsymbol{v}$ in [13] as $C\boldsymbol{v} \preceq \boldsymbol{c}$, we can obtain constrained optimization formulation in matrix form

$$\begin{aligned}&\text{Minimize } \tfrac{1}{2}\boldsymbol{v}^\top(\mathrm{A}^\top\mathrm{A} + \mathrm{I} + \mathrm{B}^\top\mathrm{B})\boldsymbol{v} + (-\boldsymbol{v}_0 - \mathrm{B}^\top\boldsymbol{b})^\top\boldsymbol{v}\\ &\text{subject to } C\boldsymbol{v} \preceq \boldsymbol{c}\end{aligned} \tag{23}$$

where I is $(M \times N) \times (M \times N)$ identity matrix and terms unrelated to $\boldsymbol{v}$ are omitted. This is a typical quadratic programming problem which can be solved using QuadProg++[2].

References

[1] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, 2007.

[2] L. Li, J. Yao, R. Xie, M. Xia, and W. Zhang, "A unified framework for street-view panorama stitching," *Sensors*, vol. 17, no. 1, 2017, Art. no. 1.

[3] L. Li, M. Xia, C. Liu, L. Li, H. Wang, and J. Yao, "Jointly optimizing global and local color consistency for multiple image mosaicking," *ISPRS J. Photogrammetry Remote Sens.*, vol. 170, pp. 45–56, 2020.

[4] T. Shen, J. Wang, T. Fang, S. Zhu, and L. Quan, "Color correction for image-based modeling in the large," in *Proc. Asian Conf. Comput. Vis.*, 2016, pp. 392–407.

[5] J. Yang, L. Liu, J. Xu, Y. Wang, and F. Deng, "Efficient global color correction for large-scale multiple-view images in three-dimensional reconstruction," *ISPRS J. Photogrammetry Remote Sens.*, vol. 173, pp. 209–220, 2021.

[6] Y. Li, Y. Li, J. Yao, Y. Gong, and L. Li, "Global color consistency correction for large-scale images in 3D reconstruction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3074–3088, 2022.

[7] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 777–784, 2003.

[8] A. Levin, A. Zomet, S. Peleg, and Y. Weiss, "Seamless image stitching in the gradient domain," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 377–389.

[9] F. Fang, T. Wang, Y. Fang, and G. Zhang, "Fast color blending for seamless image stitching," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 7, pp. 1115–1119, Jul. 2019.

[10] X. Li, R. Feng, X. Guan, H. Shen, and L. Zhang, "Remote sensing image mosaicking: Achievements and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 8–22, Dec. 2019.

[11] A. Pandey and U. C. Pati, "Image mosaicing: A deeper insight," *Image Vis. Comput.*, vol. 89, pp. 236–257, 2019.

[12] L. Yu, Y. Zhang, M. Sun, X. Zhou, and C. Liu, "An auto-adapting global-to-local color balancing method for optical imagery mosaic," *ISPRS J. Photogrammetry Remote Sens.*, vol. 132, pp. 1–19, 2017.

[13] M. Xia, J. Yao, and Z. Gao, "A closed-form solution for multi-view color correction with gradient preservation," *ISPRS J. Photogrammetry Remote Sens.*, vol. 157, pp. 188–200, 2019.

[2]QuadProg++:https://github.com/liuq/QuadProgpp

[14] K. Liu, T. Ke, P. Tao, J. He, K. Xi, and K. Yang, "Robust radiometric normalization of multitemporal satellite images via block adjustment without master images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6029–6043, 2020.
[15] Z. Su, K. Zeng, L. Liu, B. Li, and X. Luo, "Corruptive artifacts suppression for example-based color transfer," *IEEE Trans. Multimedia*, vol. 16, no. 4, pp. 988–999, Jun. 2014.
[16] D. Wang, C. Zou, G. Li, C. Gao, Z. Su, and P. Tan, "L0 gradient-preserving color transfer," in *Computer Graphics Forum*, vol. 36, pp. 93–103, 2017.
[17] Y. Huang, S. Qiu, C. Wang, and C. Li, "Learning representations for high-dynamic-range image color transfer in a self-supervised way," *IEEE Trans. Multimedia*, vol. 23, pp. 176–188, 2020.
[18] J. Li, Q. Hu, and M. Ai, "Optimal illumination and color consistency for optical remote-sensing image mosaicking," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 1943–1947, Nov. 2017.
[19] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2006.
[20] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sep. 2009.
[21] X. Fu, J. Wang, D. Zeng, Y. Huang, and X. Ding, "Remote sensing image enhancement using regularized-histogram equalization and DCT," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2301–2305, Nov. 2015.
[22] X. Wu, X. Liu, K. Hiramatsu, and K. Kashino, "Contrast-accumulated histogram equalization for image enhancement," in *Proc. IEEE Int. Conf. Image Process.* 2017, pp. 3190–3194.
[23] K. Mayathevar, M. Veluchamy, and B. Subramani, "Fuzzy color histogram equalization with weighted distribution for image enhancement," *Optik*, vol. 216, 2020, Art. no. 164927.
[24] Z. Huang, Z. Wang, J. Zhang, Q. Li, and Y. Shi, "Image enhancement with the preservation of brightness and structures by employing contrast limited dynamic quadri-histogram equalization," *Optik*, vol. 226, 2021, Art. no. 165877.
[25] J. Pan, M. Wang, D. Li, and J. Li, "A network-based radiometric equalization approach for digital aerial orthoimages," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 401–405, Apr. 2010.
[26] C. Chen, Z. Chen, M. Li, Y. Liu, L. Cheng, and Y. Ren, "Parallel relative radiometric normalisation for remote sensing image mosaics," *Comput. Geosci.*, vol. 73, pp. 28–36, 2014.
[27] R. Xie, M. Xia, J. Yao, and L. Li, "Guided color consistency optimization for image mosaicking," *ISPRS J. Photogrammetry Remote Sens.*, vol. 135, pp. 43–59, 2018.
[28] M. He, J. Liao, D. Chen, L. Yuan, and P. V. Sander, "Progressive color transfer with dense semantic correspondences," *ACM Trans. Graph.*, vol. 38, no. 2, pp. 1–18, 2019.
[29] Y. Niu, X. Zheng, T. Zhao, and J. Chen, "Visually consistent color correction for stereoscopic images and videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 3, pp. 697–710, Mar. 2020.
[30] A. Ilie and G. Welch, "Ensuring color consistency across multiple cameras," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2005, pp. 1268–1275.
[31] R. Cresson and N. Saint-Geours, "Natural color satellite image mosaicking using quadratic programming in decorrelated color space," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 4151–4162, Aug. 2015.
[32] Y. Xiong and K. Pulli, "Color matching of image sequences with combined gamma and linear corrections," in *Proc. Int. Conf. ACM Multimedia*, 2010, pp. 261–270.
[33] J. Park, Y. W. Tai, S. N. Sinha, and I. So Kweon, "Efficient and robust color consistency for community photo collections," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 430–438.
[34] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski, "Optimizing color consistency in photo collections," *ACM Trans. Graph.*, vol. 32, no. 4, 2013, Art. no. 38.
[35] X. Zhang, R. Feng, X. Li, H. Shen, and Z. Yuan, "Block adjustment-based radiometric normalization by considering global and local differences," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 8002805.
[36] S. S. Agaian, B. Silver, and K. A. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 741–758, Mar. 2007.

Yinxuan Li received the B.E. degrees in photogrammetry and remote sensing, in 2014, from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, where she is currently working toward the Ph.D. degree in photogrammetry and remote sensing with the School of Remote Sensing and Information Engineering.

Her research interests include color correction for multiple images, texture mapping for 3-D models, image-based 3-D reconstruction, 2-D image processing and analysis, and machine learning.

Li Li received the B.E., M.S., and Ph.D. degrees in photogrammetry and remote sensing from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2013, 2016, and 2019, respectively.

He is currently an Associate Researcher with the School of Remote Sensing and Information Engineering, Wuhan University. His research interests include image mosaicking, texture mapping, LiDAR data processing, and machine learning.

Jian Yao received the B.Sc. degree in automation from Xiamen University, Xiamen, China, in 1997, the M.Sc. degree in computer science from Wuhan University, Wuhan, China, and the Ph.D. degree in electronic engineering from The Chinese University of Hong Kong, Hong Kong, in 2006.

He was a Research Assistant with Shenzhen R&D Centre of City University of Hong Kong, Hong Kong, from 2001 to 2002; a Postdoctoral Fellow with Computer Vision Group, IDIAP Research Institute, Martigny, Switzerland, from 2006 to 2008; a Research Grantholder with the Institute for the Protection and Security of the Citizen, European Commission Joint Research Centre, Ispra, Italy, from 2009 to 2011; and a Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Beijing, China, from 2011 to 2012. Since April 2012, he has been a Hubei "Chutian Scholar" Distinguished Professor with the School of Remote Sensing and Information Engineering, Wuhan University, where he is currently the Director of Computer Vision and Remote Sensing Lab. He has authored or coauthored more than 100 papers in international journals and proceedings of major conferences and holds more than 30 patents. His research interests include computer vision, image processing, deep learning, LiDAR data processing, robotics, SLAM, etc.

Menghan Xia received the B.Eng. degree in remote sensing science and techniques and the master's degree in pattern recognition and intelligent system from Wuhan University, Wuhan, China, in 2014 and 2017, and the Ph.D. degrees in computer science from The Chinese University of Hong Kong, Hong Kong, Hong, Kong, in 2021.

He is currently a Researcher with Tencent AI Lab, Shenzhen, China. His research interests include computer vision, visual media restoration and generation.

Hanyun Wang received the Ph.D. degree in electronic science and technology from National University of Defense Technology, Changsha, China, in 2015.

He is currently an Associate Professor with the School of Surveying and Mapping, Information Engineering University, Zhengzhou, China. He was a visiting Ph.D. student with Xiamen University, Xiamen, China, from 2011 to 2014. He has authored over 30 articles in journals and conferences, such as IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE Conference on Computer Vision and Pattern Recognition, and *ISPRS Journal of Photogrammetry and Remote Sensing*. His research interests include 3-D computer vision, especially on point cloud registration, 3-D object detection and scene understanding.