

Jointly optimizing global and local color consistency for multiple image mosaicking

Li Li ^a, Menghan Xia ^b, Chi Liu ^a, Liang Li ^{c,*}, Hanyun Wang ^d, Jian Yao ^{a,*}

^a School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, PR China

^b Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

^c Shenzhen Polytechnic, Shenzhen 518055, PR China

^d School of Surveying and Mapping, Information Engineering University, Zhengzhou 450000, PR China



ARTICLE INFO

Keywords:

Color consistency
Image mosaicking
Linear model
Multiple images
Global and local

ABSTRACT

In multiple image mosaicking, color inconsistency is a common and severe problem that degrades the quality of composite images. To avoid the appearance of visible seams, we need to optimize the color consistency before mosaicking multiple images. To facilitate the global optimization framework, existing approaches mainly use less flexible global models, e.g., the linear or gamma model, to eliminate the color differences between multiple images. These models can effectively ensure that the global tones in multiple images are consistent. However, their ability to correct local color discrepancies is usually poor. In this paper, we propose a novel image color correction approach that can correct global and local color discrepancies simultaneously and preserve image gradient as much as possible. First, we design an effective and flexible color correction model to represent the color mapping function for each image. Instead of using the same global model to correct color discrepancies for all pixels in one image, we apply a series of local linear models for color correction. For different superpixel regions, different linear models are applied to model the mapping functions. Second, based on this model, a specific global cost function that considers both gradient preservation and color consistency is designed and solved. In addition, a global color constraint is fused into this cost function to ensure that the corrected images have the similar global tone. Thus, we can jointly optimize the global and local color consistency by minimizing this cost function. The experimental results on several challenging datasets captured by different sensors demonstrate that the proposed approach outperforms the state-of-the-art color correction approaches in both visual quality and quantitative metrics.

1. Introduction

Image mosaicking, which is used to merge a set of geometrically aligned images into a single composite image as seamlessly as possible, is an important and classical research topic in the remote sensing (Li et al., 2019b) and computer vision (Pandey and Pati, 2019) fields. Image mosaicking is a key technology for producing pleasant and seamless panoramic images and digital orthophoto maps (DOMs). There are two key problems in image mosaicking that need to be solved. The first problem is the geometrical misalignments caused by the image registration error. This problem can be solved well using image warping (Chen et al., 2018; Xiang et al., 2018) and optimal seamline detection (Lin et al., 2016; Li et al., 2019a) techniques. In addition, in most cases, especially for multitemporal satellite images, there are drastic photometric inconsistencies between images due to the atmospheric illumination variations, different exposure settings and different camera response functions. An image blending technique (Pérez et al.,

2003; Levin et al., 2004; Fang et al., 2019) can effectively smooth the small color differences along seamlines, but it cannot eliminate drastic color discrepancies. To avoid the appearance of color artifacts and visible seams on the last composite image, this color inconsistency must be eliminated before mosaicking multiple images. For the case of two images, many advanced color transfer approaches (Su et al., 2014; Arbelot et al., 2017; He et al., 2019; Niu et al., 2019) in the computer vision field and relative radiometric normalization methods (Canty et al., 2004; Li et al., 2017a; Syariz et al., 2019) in the remote sensing field have been proposed. These techniques can achieve pleasant color correction results, but they can hardly be used to correct the color discrepancies for multiple images. However, in most practical applications, we need to stitch multiple images to generate the final composite image. In this paper, our effort is focused on optimizing the color consistency for multiple images.

* Corresponding authors.

E-mail addresses: liliang@szpt.edu.cn (L. Li), jian.yao@whu.edu.cn (J. Yao).

Color correction for multiple images is a key issue for producing a composite image with a visually consistent tone. Many previous works have studied this issue. Generally, these approaches can be classified into path propagation-based and global optimization-based approaches. The path propagation-based approaches (Pan et al., 2010; Chen et al., 2014; Xie et al., 2018) first select one image as the reference image and then repeatedly transfer the color of the target images into the reference image along the optimal propagation path found from the adjacent graph. Pan et al. (2010) proposed a network-based color correction approach to eliminate the color differences between images. This approach first constructs a network using the area Voronoi diagrams with overlap and selects the center image as the reference. Then, the shortest propagation path of each target image is found from the network using the Dijkstra shortest path algorithm. At last, the color of each image is transferred to the same reference along the propagation path. Xie et al. (2018) proposed to automatically select a set of images with a maximum connected sub graph instead of one image as the reference image. Their approach transfers the color of the target image to the reference via a histogram extreme point matching algorithm (Li et al., 2017b). Then, a global consistency optimization approach is further applied to improve the initial color correction result generated by the path propagation. However, these path propagation-based approaches have two essential problems that are difficult to solve. First, automatic reference image selection still is an open issue in the color correction field (Ibrahim et al., 2016; Yu et al., 2017b), and we usually need to manually select one image from the input images as the best reference. Second, because the large accumulated error cannot be avoided in the path propagation approaches for images that are far away from the reference image, these approaches only work well for the case of a few images. Thus, the path propagation-based approach is not a good solution for the color correction of multiple images.

To solve the above-mentioned problems that exist in path propagation-based approaches, many global optimization-based approaches have been proposed. Instead of correcting images in sequence, they correct all images simultaneously. They formulate the color correction problem as a global energy optimization problem and solve the model parameters of all images at the same time. However, to facilitate the global optimization framework, existing optimization-based approaches mainly use less flexible models, e.g., the linear model (Brown and Lowe, 2007; Cresson and Saint-Geours, 2015; Shen et al., 2016; Yu et al., 2017a) or the gamma model (Xiong and Pulli, 2010; Park et al., 2016; Li et al., 2017a) to correct color discrepancies. Brown and Lowe (2007) first proposed to optimize the color consistency for multiple images by globally optimizing a cost function defined over all images. This approach applied the linear model to correct color discrepancies. It has been applied in panoramic image mosaicking. Xiong and Pulli (2010) combined the gamma and linear models to address the problem of luminance and color correction simultaneously. The gamma and linear models are applied in the luminance channel and chrominance channels, respectively. Shen et al. (2016) proposed to apply the linear model to adjust the color histograms of multiple images. This approach has been used for the texture mapping of three-dimensional models. Park et al. (2016) estimated the parameters of gamma models for each image via robust low-rank matrix factorization. However, both linear and gamma models cannot effectively approximate the drastic color discrepancies between multiple images, and they especially struggle to correct local color differences. In addition, the visual quality of images are not considered in their cost function. To solve these problems, Yu et al. (2017b) proposed a global-to-local color correction method that aims to eliminate the global and local color differences. They first solved the parameters of linear models for all images via a global optimization strategy. A local optimization strategy is further applied to eliminate the local color differences. Actually, the local optimization is a postprocessing technique that is tedious and not optimal. In addition, this technique still does not consider the visual quality in the cost function. Instead of using less flexible linear and gamma models, Xia

et al. (2019) applied a parametrized spline curve model (HaCohen et al., 2013; Xia et al., 2017; Hu et al., 2019) to model the color remapping function for each image. The optimal model parameters are solved by minimizing a specific cost function that considers both the visual quality of an individual image and the global color consistency. Compared with existing linear and gamma models, the curve model is more flexible, but it still cannot effectively correct local color discrepancies. However, because the reflectivity of different object materials in the scene is different, the color discrepancies between different regions of images are usually different (Fan et al., 2019). Almost all the global optimization-based approaches use the same color mapping function for all pixels in one image, and their ability to correct local color discrepancies is low. Therefore, color correction for multiple images still is an open issue.

To address the problems discussed above, we proposed a novel global optimization-based color correction approach. We creatively propose a flexible color correction model that can effectively approximate complex local color discrepancies between multiple images. This model is comprised of a series of linear models. Instead of using the same global model to correct the color for all pixels in one image, we correct it locally. For different local superpixel regions, different linear models are applied to correct the local color differences. To globally solve the model parameters for all images at the same time, a special cost function that considers both gradient preservation and color consistency is designed. Because the proposed model approximates the color discrepancies locally, to ensure that the corrected images have the similar global tone, a global color consistency constraint is also fused into this cost function. To the best of our knowledge, in the global optimization-based color correction approaches, we are the first one to consider global and local color differences between multiple images together using a series of local models.

The rest of this paper is organized as follows. Section 2 introduces our proposed approach in detail, including the superpixel segmentation and the construction and optimization of the color correction model. The experimental results on several sets of challenging datasets are presented in Section 3. The conclusions are given in Section 4.

2. Our proposed approach

The input of our approaches is a set of images. Let $\mathcal{I} = \{\mathbf{I}_i\}_{i=1}^N$ denote all input images, where N is the number of images. The input images should be geometrically aligned into the same coordinate so we can construct the adjacent image relationship before color correction. The goal of the proposed approach is to simultaneously eliminate the global and local color differences between multiple images while preserving the image gradient of each input image as much as possible. The whole workflow of our approach is presented in Fig. 1. As shown in Fig. 1, the proposed approach consists of three steps: superpixel segmentation, correction model construction and model parameter optimization.

2.1. Superpixel segmentation

In this paper, superpixels are considered as the basic processing units to optimize the color consistency for multiple images. One of the most widely used superpixel segmentation approaches is the simple linear iterative cluster (SLIC) algorithm (Achanta et al., 2012). It first samples K superpixel cluster centers at regular grid intervals. The mean superpixel size M is defined as $M = (W \times H)/K$, where W and H are the width and height of the input image, respectively. We can use M to control the number of superpixels. Then, the algorithm generates superpixels by clustering pixels in the 5D space defined by the L , a and b of the CIELAB color space as well as the x and y coordinates of the pixels. The distance measure that is applied to evaluate the similarity between the cluster center and the pixel is designed by combining the color distance and space distance. However, the SLIC algorithm can only generate superpixels for one image each time, and the segments

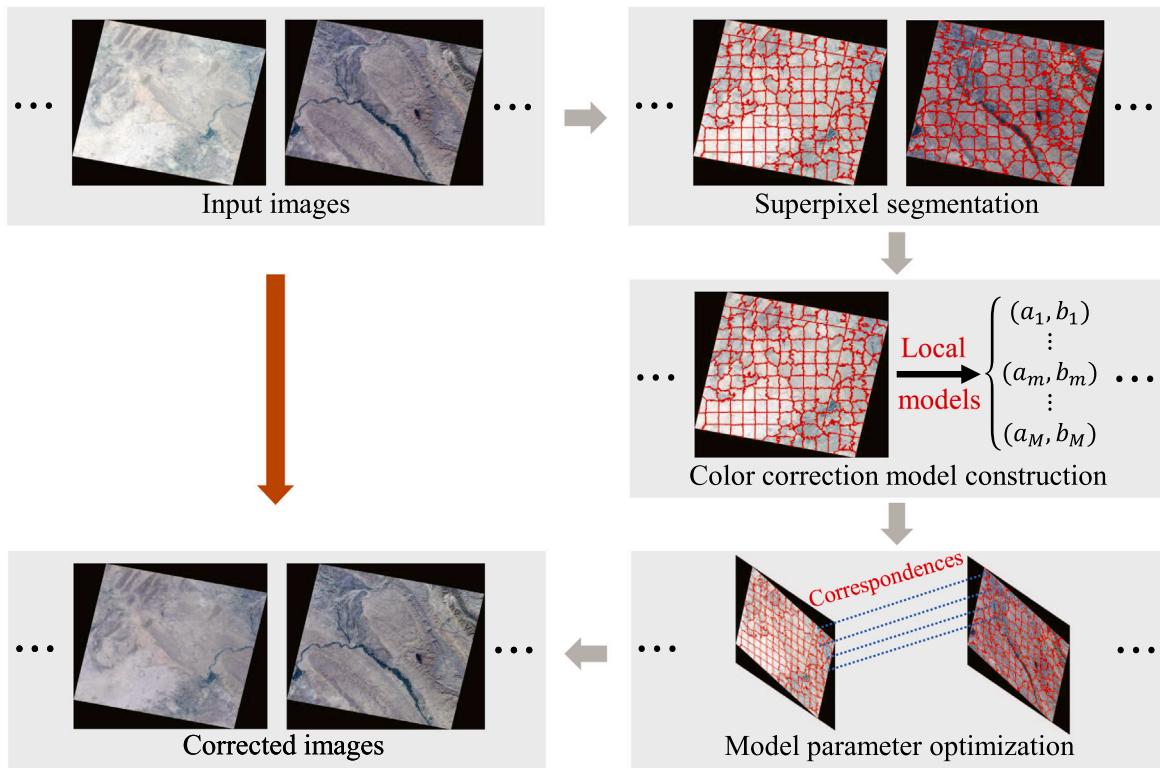


Fig. 1. The workflow of our proposed color correction approach.

of different images may be different. However, for each image pair, we need to ensure that the segments of two images are the same so we can directly extract the color correspondences. To solve this problem, we directly applied the modified SLIC algorithm presented in Li et al. (2018) to generate the superpixels for multiple images. In the modified SLIC algorithm, the value of the distance measurement in each image is calculated as with the traditional SLIC, and then the maximum value of all images is directly used as the final distance measurement. The information of all input images can be involved in the clustering process using this simple strategy. An example is shown in Fig. 2. There are two geometrically aligned images. From the results, it is apparent that the boundaries of the superpixels match the edges of the two images as expected.

2.2. Color correction model construction

Each color-rich image may contain many different objects. However, because the reflectivity of the different materials in the image is different, the color discrepancies between images are usually local. In our approach, instead of using the same global model to correct the color for all pixels in one image, we propose correcting it locally. Because the superpixels are an over-segmentation of an image, each superpixel belongs to the same physical world object. Therefore, we can assume that the pixels included in one superpixel share the same color mapping function, and we apply a linear model to represent this function, as shown in Fig. 3. The linear model is a widely used model in the color correction field. According to the definition of the linear model, for a pixel p in the input image \mathbf{I}_i , the intensity of p in the corrected image $\hat{\mathbf{I}}_i$ is calculated as follows:

$$\hat{\mathbf{I}}_i(p) = \mathbf{I}_i(p) \times a + b, \quad (1)$$

where a and b are the gain and bias of the linear model, respectively. $\mathbf{I}_i(p)$ and $\hat{\mathbf{I}}_i(p)$ represent the intensities of p in images \mathbf{I}_i and $\hat{\mathbf{I}}_i$, respectively.

We apply the method introduced in Section 2.1 to generate consistent superpixels for all input images \mathcal{I} . Let $S = \{S_k\}_{k=1}^K$ denote all superpixels, where K is the number of superpixels. Obviously, each superpixel S_k may be covered by one image or several images. In addition, each image \mathbf{I}_i contains many superpixels. Let $S_i^{\mathbf{I}_i} = \{S_k^{\mathbf{I}_i}\}_{k=1}^{K_i}$ denote all superpixels included in \mathbf{I}_i , where K_i denotes the number of corresponding superpixels. Obviously, $S_i^{\mathbf{I}_i}$ is a subset of S . For each superpixel $S_k^{\mathbf{I}_i}$ included in image \mathbf{I}_i , we apply $L_k^{\mathbf{I}_i} = (a_k^{\mathbf{I}_i}, b_k^{\mathbf{I}_i})$ to denote the gain and bias of the local linear model. For each image \mathbf{I}_i , the final color correction model is a set of local linear models, and it can be represented as $\mathcal{L}^{\mathbf{I}_i} = \{L_k^{\mathbf{I}_i}\}_{k=1}^{K_i}$. The proposed color correction model is simple, but it is flexible and effective when approximating complex color discrepancies.

2.3. Model parameter optimization

Compared with existing global models, the proposed local color correction model is more flexible. However, because our color model is comprised of a series of linear models, the number of parameters in our model is large. Thus, it is hard to solve the parameters for each image globally. To solve this problem, we will discuss how to design the global cost function over all images and how to solve it globally in the next section. We attempt to estimate the optimal $\mathcal{L}^{\mathbf{I}_i}$ for each image \mathbf{I}_i to eliminate the color discrepancies between multiple images and preserve the image gradient for each image. It should be noted that the local models are optimized in each channel of the RGB color space independently. For each channel, we designed the global cost function as:

$$\begin{aligned} E = & \sum_{k=1}^K \sum_{(\mathbf{I}_i, \mathbf{I}_j) \in \mathcal{I}(S_k)} E_{data}(S_k^{\mathbf{I}_i}, S_k^{\mathbf{I}_j}) + \sum_{\mathbf{I}_i \in \mathcal{I}} \sum_{m=1}^{K_i} \sum_{S_m^{\mathbf{I}_i} \in \mathcal{N}(S_k^{\mathbf{I}_i})} E_{smooth}(S_m^{\mathbf{I}_i}, S_n^{\mathbf{I}_i}) \\ & + \sum_{\mathbf{I}_i \in \mathcal{I}} \sum_{k=1}^{K_i} (E_{regular}(S_k^{\mathbf{I}_i}) + \lambda E_{gradient}(S_k^{\mathbf{I}_i}) + E_{global}(S_k^{\mathbf{I}_i})), \end{aligned} \quad (2)$$

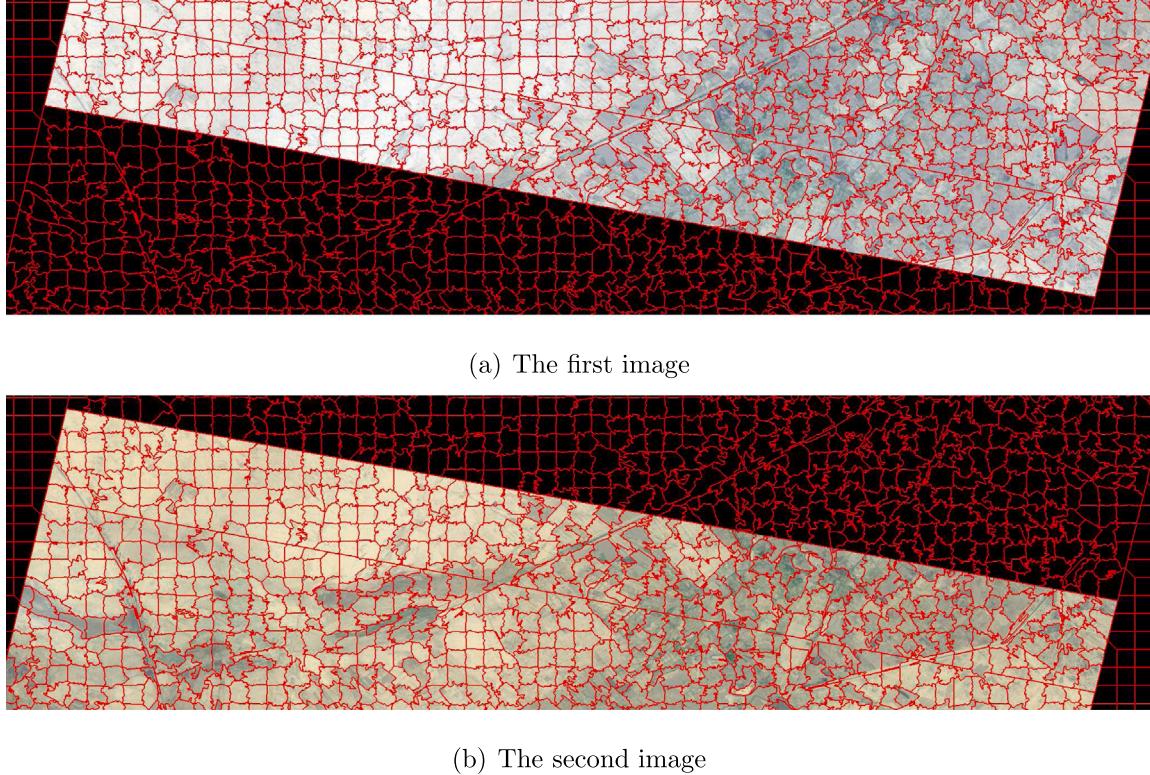


Fig. 2. A visual example of the superpixels generated by the modified SLIC algorithm for two geometrically aligned images.

where $\mathcal{I}(S_k)$ denotes the images that cover the superpixel S_k , and \mathbf{I}_i and \mathbf{I}_j are two overlapped images included in $\mathcal{I}(S_k)$. $\mathcal{N}(S_m^{\mathbf{I}_i})$ denotes the neighboring superpixels of $S_m^{\mathbf{I}_i}$ in image \mathbf{I}_i , and $S_n^{\mathbf{I}_i}$ is one of the neighboring superpixel of $S_m^{\mathbf{I}_i}$. λ is a weight that balances the influence of the gradient term. This function is comprised of data, smooth, regular, gradient and global terms. The data term, which is defined as the sum of the squared differences of the corrected colors between all overlapped superpixels, is applied to evaluate the color consistency between images. For a superpixel S_k , we assume that S_k is covered by image \mathbf{I}_i and image \mathbf{I}_j . The corresponding data term $E_{data}(S_k^{\mathbf{I}_i}, S_k^{\mathbf{I}_j})$ is calculated as follows:

$$E_{data}(S_k^{\mathbf{I}_i}, S_k^{\mathbf{I}_j}) = \|(\mu(S_k^{\mathbf{I}_i}) \times a_k^{\mathbf{I}_i} + b_k^{\mathbf{I}_i}) - (\mu(S_k^{\mathbf{I}_j}) \times a_k^{\mathbf{I}_j} + b_k^{\mathbf{I}_j})\|_2, \quad (3)$$

where $\mu(S_k^{\mathbf{I}_i})$ and $\mu(S_k^{\mathbf{I}_j})$ denote the mean intensities of the superpixel region S_k in images \mathbf{I}_i and \mathbf{I}_j , respectively. $\|\cdot\|_2$ denotes the L_2 norm. Obviously, the data term is applied to ensure that the local color differences between images are minimal.

Because we apply different linear models to correct the local color differences for different superpixel regions, the color inconsistency among the individual images may appear if the linear models of the neighboring superpixels are very different. To alleviate this problem, we design a smooth term to constrain the linear model's smoothness. We assume that the linear models of neighboring superpixels in the same image should be similar. We sample a set of intensity values in the range of [0, 255] at a fixed interval. Let $\mathcal{X} = \{x_r\}_{r=1}^R$ denote the sampled intensities, where R is the number of sampled intensities. In fact, our approach is insensitive to R . We empirically set $R = 3$. In our designed smooth term, we attempt to retain these sampled intensities that are close after two neighboring local color corrections using different linear models. In this way, we can constrain the two parameters of the linear model at the same time. For two neighboring superpixels $S_m^{\mathbf{I}_i}$ and $S_n^{\mathbf{I}_i}$ in an image \mathbf{I}_i , the smooth term $E_{smooth}(S_m^{\mathbf{I}_i}, S_n^{\mathbf{I}_i})$ is defined as follows:

$$E_{smooth}(S_m^{\mathbf{I}_i}, S_n^{\mathbf{I}_i}) = \sum_{x_r \in \mathcal{X}} \|(x_r \times a_m^{\mathbf{I}_i} + b_m^{\mathbf{I}_i}) - (x_r \times a_n^{\mathbf{I}_i} + b_n^{\mathbf{I}_i})\|_2. \quad (4)$$

However, since an optimal solution to this problem is that the parameters of all local linear models are equal to 0, we need to add a more regular term to ensure that the solution is valid. Here, we assume that the corrected image colors should be as close as possible to the original image colors. Therefore, for each superpixel $S_k^{\mathbf{I}_i}$ included in image \mathbf{I}_i , the regular term is defined as follows:

$$E_{regular}(S_k^{\mathbf{I}_i}) = \|(\mu(S_k^{\mathbf{I}_i}) \times a_k^{\mathbf{I}_i} + b_k^{\mathbf{I}_i}) - \mu(S_k^{\mathbf{I}_i})\|_2. \quad (5)$$

We can obtain a valid solution by minimizing the energy function comprised of data, smooth and regular terms. However, the corrected image may suffer from the problem of over-blurring, which will destroy the original image details and lower the sharpness of object edges. The image gradient is a key low-level feature that represents the image details. Hence, we design a gradient term in our energy function to preserve the image gradient and avoid the loss of image details. As we know, the image gradient is the first order differential of the intensity. Thus, if the gains of the linear models are equal to 1, the image gradient would be effectively preserved. Therefore, for each superpixel $S_k^{\mathbf{I}_i}$ in image \mathbf{I}_i , the gradient term $E_{gradient}(S_k^{\mathbf{I}_i})$ is defined as follows:

$$E_{gradient}(S_k^{\mathbf{I}_i}) = \|a_k^{\mathbf{I}_i} - 1\|_2. \quad (6)$$

Obviously, since the gradient term attempt to keep the gains close to 1, it will limit the flexibility of our proposed model. However, such a compromise is worthwhile because it will significantly improve the visual quality of corrected images. The influence of the gradient term will be visually and quantitatively illustrated in Section 3.2.

We can obtain a statistical global optimal solution by minimizing the energy function comprised of data, smooth, regular and gradient terms. However, because our proposed color correction model is comprised of a series of local linear models, it can effectively eliminate the local color discrepancies, but its ability to correct global color discrepancies is low. To solve this problem, we creatively fuse the global color consistency constraint into the energy function as a global term. For each image \mathbf{I}_i , we define an additional global linear model.

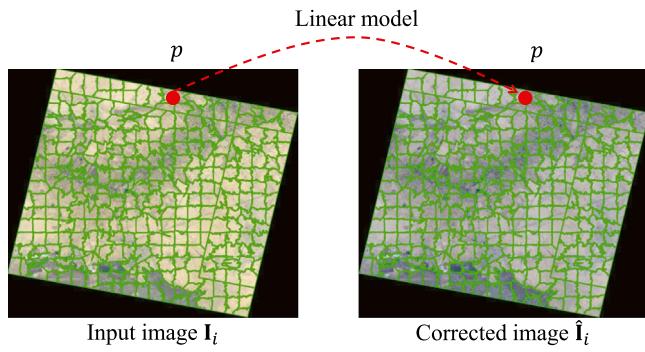


Fig. 3. Illustration of the proposed color correction model.

Let a^{I_i} and b^{I_i} denote the gain and bias of the global linear model. It should be noted that the global linear model is a virtual model. It is only applied to constrain the proposed local color models, and it will not be applied to correct the color for input images. We attempt to constrain the parameters of the local linear models in one image to be close to the parameters of the global model. In this way, we can effectively eliminate the global and the local color discrepancies at the same time. For each superpixel $S_k^{I_i}$ in image I_i , the global term $E_{global}(S_k^{I_i})$ is defined as follows:

$$E_{global}(S_k^{I_i}) = \|(\mu(S_k^{I_i}) \times a_k^{I_i} + b_k^{I_i}) - (\mu(S_k^{I_i}) \times a^{I_i} + b^{I_i})\|_2. \quad (7)$$

In Fig. 4, we present a visual color correction example. The original images are shown as a composite image in Fig. 4(a). There are severe color discrepancies between the input images. Fig. 4(b) presents the color correction result without the use of the proposed global constraint. Although the color differences in the overlapped regions are almost invisible and the image details are effectively preserved, the global tones between different corrected images are still different. The two red boxes in Fig. 4(b) denote two different tones. Fig. 4(c) shows the correction result with the use of the global constraint, and the global color discrepancies have been eliminated as expected. The composite image has the same global tone.

The energy function E is a quadratic objective function, which can be easily solved in closed form by setting the derivative to 0. This function can be formulated as a set of linear equations, and the least-square solution can be easily obtained using any sparse linear solver. We directly adjust the color for each image based on the solved local color correction models.

2.4. Application extension

In our global energy function, we design a smooth term to alleviate the local color inconsistency between neighboring superpixels. However, it cannot completely avoid the local inconsistency problem. Therefore, if it is necessary, the proposed approach can be improved by adding a postprocessing after global energy optimization. Based on the solved local model parameters, we can interpolate the linear model parameters for each pixel. Then, the color of original images can be smoothly adjusted pixel by pixel. In this way, the local inconsistency problem can be almost completely solved. Of course, the computational cost will increase. We think that the smooth constraint defined in Eq. (4) can alleviate the local inconsistency problem well. And it should be noted that the above-mentioned postprocessing strategy is not performed in our experiments. In this section, we just want to illustrate how to improve our approach if the users thought that it is necessary to completely solve the local inconsistency problem in their practical applications.

Table 1

The details of the five datasets used in our experiments.

Dataset	Image Size	Image number	Mean overlap	Platform
Land	2758 × 2290	8	20%	ZY-3
Coast	2270 × 1939	12	20%	ZY-3
City	900 × 940	9	30%	SuperView-1
Village	800 × 600	130	60%	UAV
Kitchen	800 × 1200	15	50%	SLR camera

3. Experimental results and discussion

We conducted our experiments on five datasets captured by different sensors from different areas, including three satellite datasets, one unmanned aerial vehicle (UAV) dataset and one panoramic image dataset. For each dataset, the included images have been geometrically aligned into the same coordinate system. The key information of the used five datasets is presented in Table 1. It should be noted that the images included in these five datasets have been down-sampled to reduce the computational cost and the requirement of computer memory. The Land, Coast and City datasets consist of multi-temporal satellite images. In these datasets, there are severe color discrepancies between the original images. The Village dataset is comprised of hundreds of images captured by a UAV camera. The Kitchen dataset consists of a series of images acquired for 360° panoramic image production. To largely increase the difficulty of color correction, the original colors of the images included in the Kitchen and Village datasets have been adjusted randomly. In our experiments, the Land dataset is used to illustrate the influences of the parameters, and the remaining four datasets are used to conduct the comparative experiments. All experiments were implemented with C++ on Windows and tested on a computer with an Intel Core i7-8700 3.2 GHz CPU and 16 GB of RAM memory.

3.1. Evaluation metrics

As we know, it would be better if the ground truth could be applied to quantitatively evaluate the experiments. However, it is infeasible to our problem, because it is difficult to define what is the ground truth of an image sequence showing inconsistent colors. The color consistency across images is a relative measure instead of an absolute measure. In addition, we still cannot use the satellite image with low spatial resolution but high coverage as the ground truth for our satellite image datasets. Because they are captured at different times by different sensors, there are large global tone differences between two kinds of images.

Therefore, in this paper, we apply two metrics presented in Xia et al. (2019) to quantitatively evaluate the color correction results generated by different approaches. The metrics consist of the color distance (CD), which is applied to evaluate the color discrepancies between overlapped corrected images; and the gradient loss (GL), which is applied to evaluate the gradient variations between the input images and corrected images. They are defined as follows:

$$CD = \sum_{I_i \cap I_j \neq \emptyset} w_{ij} \frac{\Delta H(\hat{I}_{ij}, \hat{I}_{ji})}{N_b} \quad (8)$$

$$GL = \frac{1}{N} \sum_{i=1}^N \frac{\Delta G(I_i, \hat{I}_i)}{N_p},$$

where I_i and I_j are two overlapped images. \hat{I}_i and \hat{I}_j denote the corresponding corrected images. \hat{I}_{ij} represents the region of \hat{I}_i overlapped with \hat{I}_j . Similarly, \hat{I}_{ji} represents the region of \hat{I}_j overlapped with \hat{I}_i . w_{ij} is the normalized weight and is set proportional to the area of \hat{I}_{ij} ($\sum w_{ij} = 1$). $\Delta H(\bullet)$ computes the difference between the color histograms extracted from I_{ij} and \hat{I}_{ij} by bins, and N_b is the bin number of the color histogram. $\Delta G(\bullet)$ computes the difference between the gradient orientation maps extracted from I_i and \hat{I}_i by pixels, and N_p is the number of valid pixels in \hat{I}_i . Smaller values of CD and GL indicate a higher quality color correction result.

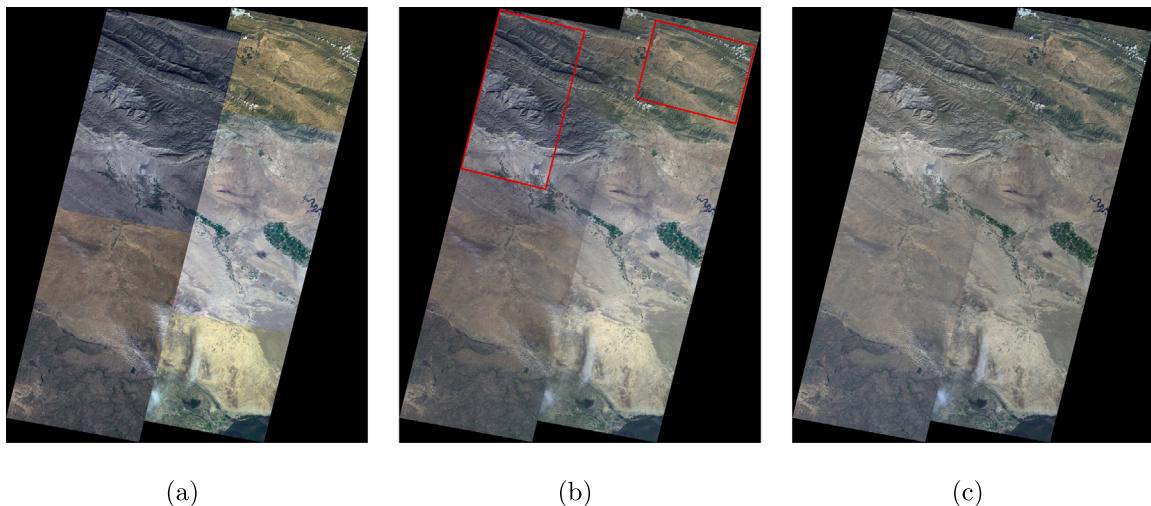


Fig. 4. Illustration of the influence of the proposed global constraint. (a) The source composite image, and (b)–(c) the color correction results without (b) and with (c) the use of the global constraint. The two red boxes are used to highlight the different tones.

3.2. Parameter determination

In our approach, there is a key parameter (the balance weight λ) that needs to be tuned. The weight λ is applied to balance the color consistency and gradient preservation. In this experiment, we used the Land dataset to quantitatively and visually illustrate how this parameter influences the color correction result of our approach, and the results are shown in Fig. 5(a) and Fig. 6. The curves of the color distance and gradient loss of our color correction results with different λ are presented in Fig. 5(a). We found that as the value of λ increases, the gradient loss decreases and the color distance increases, as expected. This is because the additional gradient term would limit the flexibility of our proposed model. We also found that for all datasets applied in our approach, $\lambda = 10$ works well, and the color consistency and gradient preservation can be well balanced. We suggest that users tune the value of λ according to their requirements in the range of [5, 20].

In addition, the mean superpixel size M will also influence the performance of our approach. The mean superpixel size M is applied to determine the number of superpixels. We fixed the value of λ as 10 and then quantitatively measure the quality of the color correction results generated using different M . The quality and time curves of our approach with different M are presented in Fig. 5(b). From the curves, we observed that as the value of M increases, both the quality and the computational time decrease. This is because the flexibility of our proposed model and the number of model parameters that need to be solved in our approach decrease as the number of local linear models decreases. The mean superpixel size M largely depends on the image sizes of the input dataset. However, we found that $M = 4000$ performs well for all datasets applied in our approach, and the efficiency and effectiveness can be well balanced. This is because the image sizes of all datasets are similar.

In Fig. 6, we visually illustrate why we need to add the gradient term into our energy function. When $\lambda = 0$, namely, the gradient term is not used, although the color discrepancies between input images can be eliminated effectively, the correction result suffers from severe over-blurring, as shown in Fig. 6(a). As the weight λ increases, increasingly more image details are preserved as expected, as shown in Fig. 6(b) and (c). When $\lambda = 10$, we observed that the image details have been well preserved and the color discrepancies have been effectively eliminated.

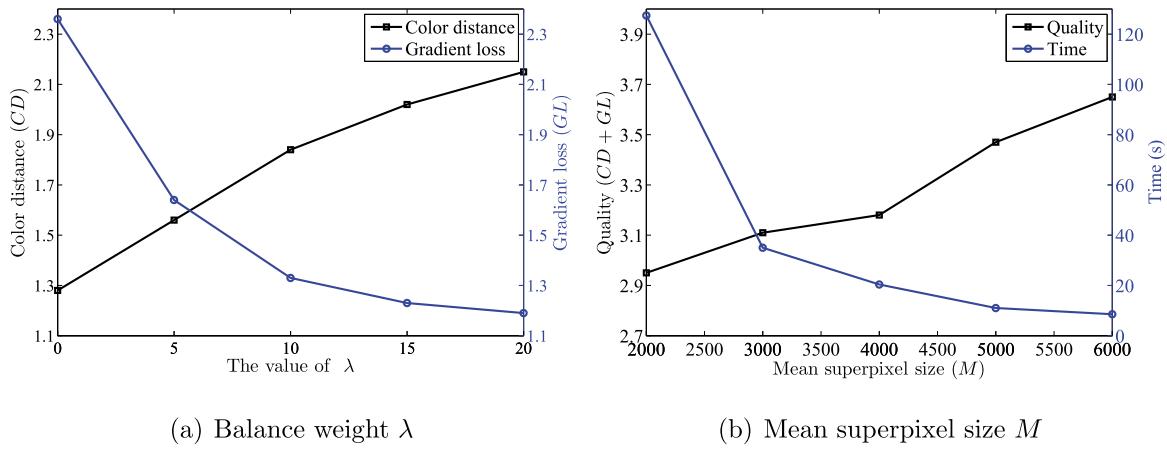
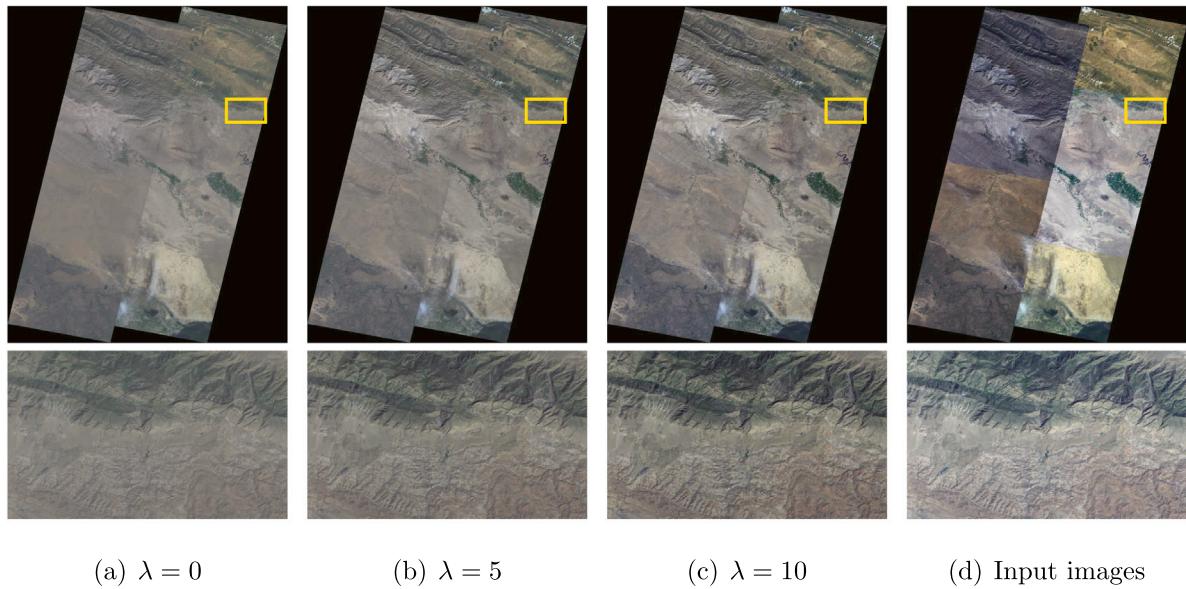
3.3. Comparative experiments

In this section, we compared our approach with four existing color correction approaches (Brown and Lowe, 2007; Xiong and Pulli, 2010;

Shen et al., 2016; Yu et al., 2017a) using the Coast, City, Village and Kitchen datasets. In the next experiments, we named these four approaches as Brown and Lown's approach, Xiong and Pulli's approach, Shen et al.'s approach and Yu et al.'s approach, respectively. A brief introduction of these approaches can be found in Section 1. In our approach, the numbers of superpixels in the Coast, City, Village and Kitchen datasets are 7243, 1121, 2525 and 1218, respectively.

We first tested our approach on two datasets captured by different satellites, as shown in Figs. 7 and 8. The first is the Coast dataset consisting of 12 multi-temporal images captured by the Chinese ZY-3 satellite. The input images are shown in Fig. 7(a) as a composite image. There are severe color differences between the input images because they are acquired at different times with variational atmosphere illuminations. In Fig. 7(b), we observed that the result of Brown and Lowe's approach has a hazy appearance, especially in the right strip. In addition, it also has visible color discrepancies that remained between the images, as shown in the second enlarged region. Fig. 7(c) presents the color correction result of Xiong and Pulli et al.'s approach, and we observed that it has visible color discrepancies that remained across image strips, as shown in the first and second enlarged regions. The result of Shen et al.'s approach is presented in Fig. 7(d). We found that the result of their approach has the highest contrast. However, some regions are too bright, as shown in the first enlarged region. The main reason is that the images corrected by their approach have been stretched linearly after the color correction. However, it still has some color discrepancies that remained between neighboring images, as shown in the second enlarged region. Fig. 7(e) presents the color correction result of Yu et al.'s approach. Although Yu et al.'s approach offers a result with consistent color, it suffers from over-blurring in a local region, as shown in the first enlarged region. From the comparative results of the Coast dataset presented in Fig. 7, we observed that the color correction result generated by our approach is the best, as shown in Fig. 7(f).

The second dataset used is the City dataset, which is comprised of 9 images captured by the Chinese SuperView-1 satellite. In this dataset, the color discrepancies between input images are relatively small, as shown in Fig. 8(a). Our approach and Yu et al.'s approach perform well on this dataset in general, and the color discrepancies between input images have been eliminated effectively. However, for the first enlarged region generated by Yu et al.'s approach, we found that the right image is slightly brighter than the left image. For the second enlarged region generated by Yu et al.'s approach, we observed that the color difference between bottom image and upper image is larger than the result generated by our approach. But, in the second enlarged region,

(a) Balance weight λ (b) Mean superpixel size M Fig. 5. Illustration of the influence of the balance weight λ and the mean superpixel size M .(a) $\lambda = 0$ (b) $\lambda = 5$ (c) $\lambda = 10$

(d) Input images

Fig. 6. Visual illustration of the influence of the balance weight λ .

we also found that the right image is brighter than the left image in the result generated by our approach. Yu et al.'s result is more consistent in terms of brightness between these two images. Although the local and global color differences are considered in our approach together, we still cannot ensure that every local region is very good in color consistency. The goal of our approach is to ensure that the whole color differences between multiple images are small. The results of Brown and Lowe's approach and Xiong and Pulli's approach still have some tonal differences, especially in the regions of the mountain and lake (the second enlarged region), as shown in Fig. 8(b) and (c). Fig. 8(d) presents the color correction result of Shen et al.'s approach. It offers the highest contrast but some regions are overly brightened. Especially in the region of the rightmost image, the color of the mountain is too bright.

Next, we tested our approach on the dataset captured by a UAV, as shown in Fig. 9. We evaluated the performance of our approach on the Village dataset, which is comprised of more than a hundred UAV images. For this challenging dataset, the results of Brown and Lowe's approach, Xiong and Pulli's approach, Shen et al.'s approach and Yu et al.'s approach all suffer from some serious problems. The result of Brown and Lowe's approach has a consistent global tone, but there still are some tonal differences between neighboring images, as shown in the first and second enlarged regions of Fig. 9(b). Xiong

and Pulli's approach generates the worst color correction result for this dataset, and there still are severe color differences between the corrected images, similar to the input images. Shen et al.'s approach performs well in optimizing the color consistency, and only small color discrepancies appeared in the corrected result, as shown in the second enlarged region. However, the result of Shen et al.'s approach suffers from severe inconsistent illumination, and some regions are too dark (the first enlarged region). Compared with the above approaches, Yu et al.'s approach offers a better result, but it still has some color discrepancies, as shown in the first and second enlarged regions. Similar to the results of satellite images, our approach offers the most visually pleasant result that is significantly better than the results generated by the rest of the approaches.

Finally, in Fig. 10, we tested the proposed approach on the Kitchen dataset. It is easy to observe that the results presented in Fig. 10(b) and (c) still have severe color differences. Compared with the results presented in (b) and (c), the result of Shen et al.'s approach is better, but it suffers from low global contrast, and still has some visible color differences. As shown in Fig. 10(e) and (f), we observed that the color correction result of Yu et al.'s approach is similar to our approach. However, in some local regions (two enlarged local areas), the result of Yu et al.'s approach also has some color differences. For this dataset, our approach is significantly better than Brown and Lowe's approach,

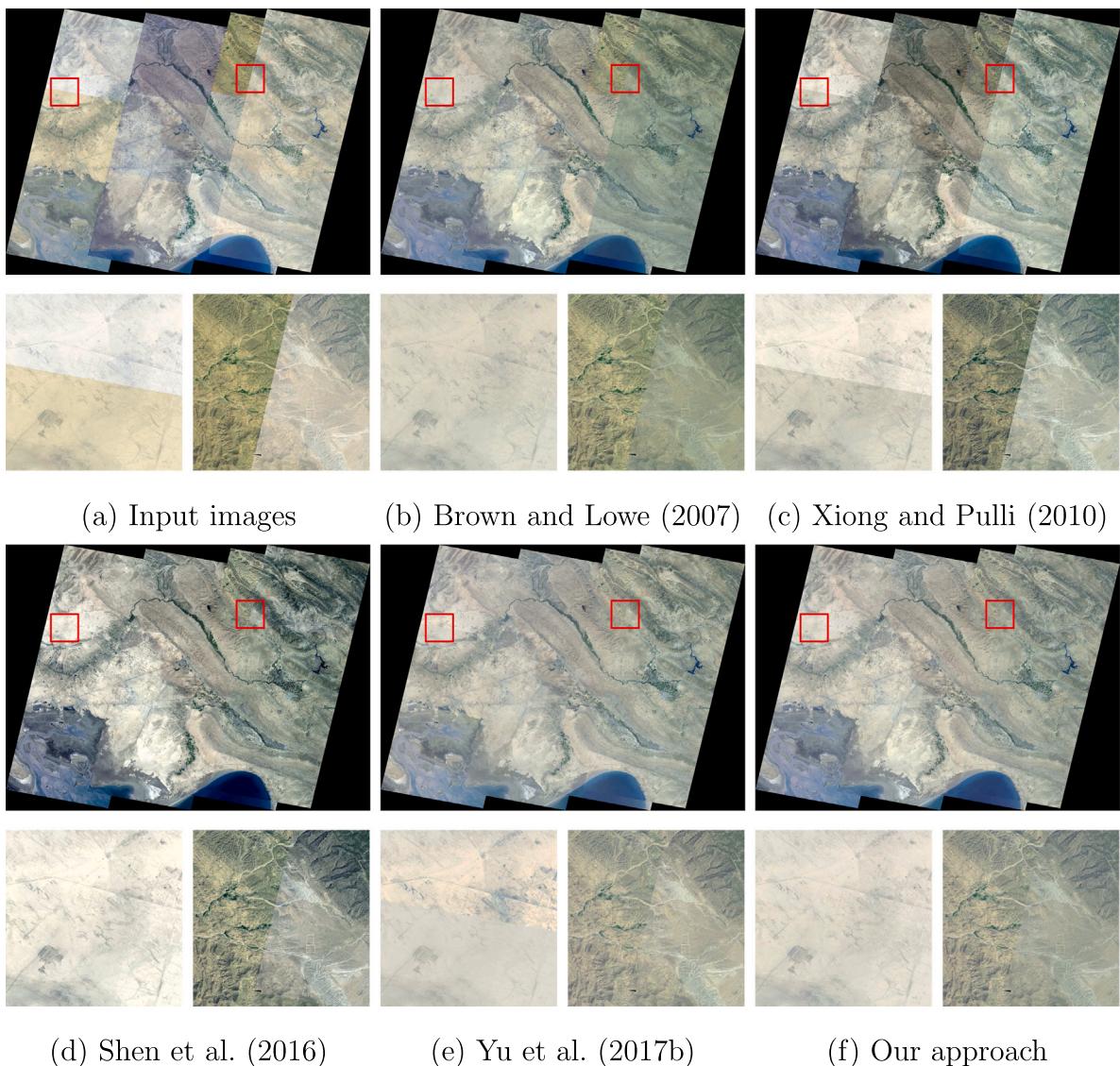


Fig. 7. The color correction results of the Coast dataset. (a) is the input images. (b)–(f) are the results of Brown and Lown’s approach, Xiong and Pulli’s approach, Shen et al.’s approach, Yu et al.’s approach and our proposed approach, respectively.

Xiong and Pulli’s approach, and Shen et al.’s approach and is slightly superior to Yu et al.’s approach.

3.4. Quantitative comparison

In addition to the visual comparison, to convincingly illustrate the superiority of our approach, we also performed the quantitative evaluation for all approaches using the metrics defined in Section 3.1. The quality evaluation results are presented in Table 2. From this table, we observed that our approach offers the best *CD* scores for all datasets, and these scores are significantly better than the scores of the rest of the four approaches. Especially, for the City dataset, our approach offers the best *CD* and *GL* scores. However, for the Coast, Village and Kitchen datasets, we found that the *GL* scores offered by our approach are not the best. The main reason is that the color information of the input images in these datasets has been greatly changed to eliminate the severe color discrepancies, and some gradient information is inevitably sacrificed. Although Xiong and Pulli’s approach offers the best *GL* scores for these three datasets, the *CD* scores of this approach are the worst. Considering the *CD* and *GL* scores together, the proposed approach outperforms the rest of the compared approaches.

In addition, we reported the computational times of all approaches in Table 2. In Table 2, the computational time is denoted as *T*, and the unit is seconds. We found that the computational times of Brown and Lown’s approach, Xiong and Pulli’s approach, and Shen et al.’s approach are at the same level. These three approaches are efficient because all of those approaches apply a less flexible model with few parameters to correct the image color. We also found that the computational time of our approach is shorter than that of Yu et al.’s approach except for the Village dataset. This is mainly because the number of images in the Village dataset is large and the number of parameters in our model increases drastically.

4. Conclusion

In this paper, to generate a large-scale composite image with visually consistent color, we proposed an effective global optimization-based color correction approach. The key advantage of our approach is that we can effectively and simultaneously eliminate the global and local color discrepancies between multiple images while preserving the image gradient of individual images based on our flexible color correction model. The key contributions of this paper can be summarized as follows:

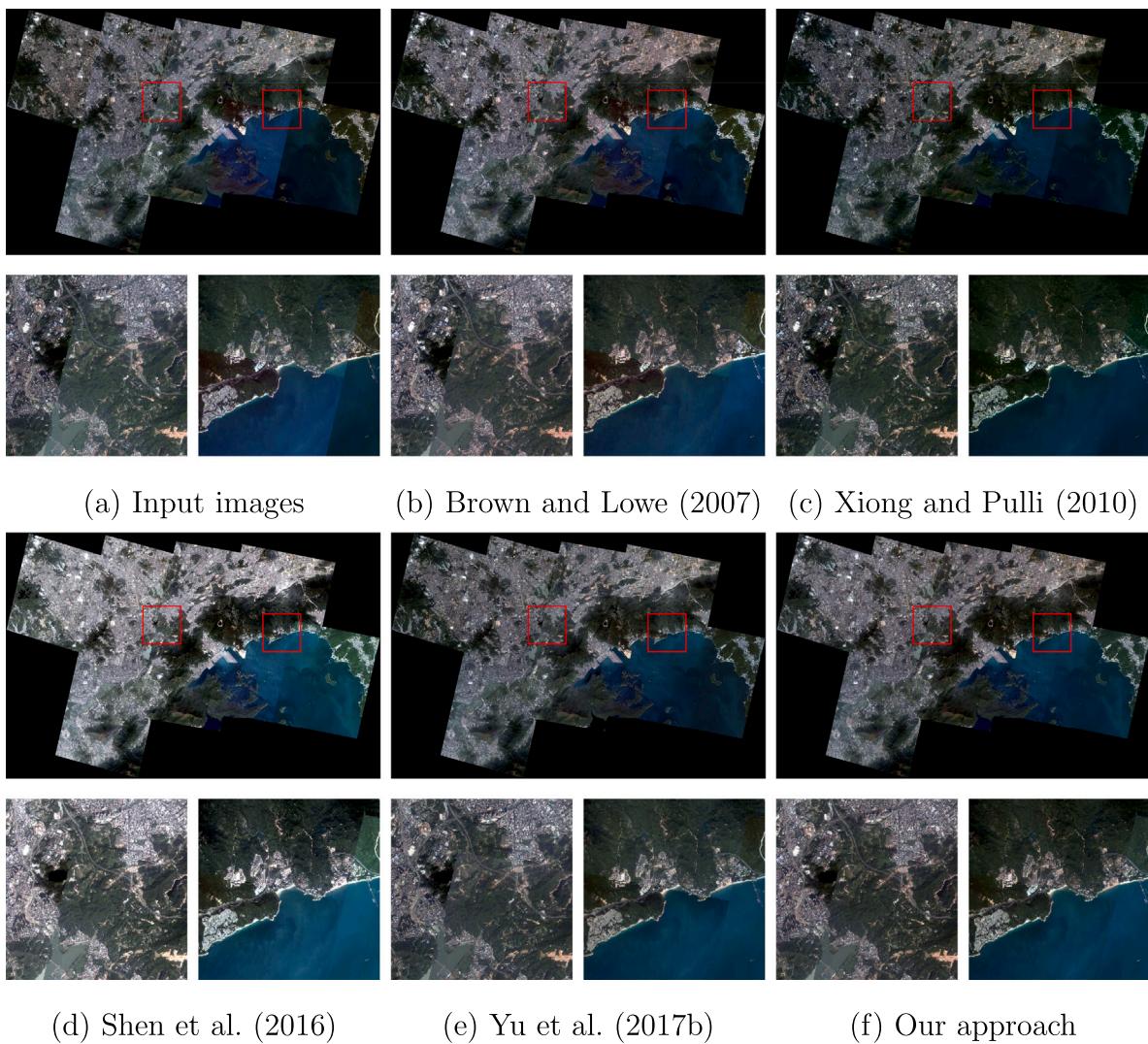


Fig. 8. The color correction results of the City dataset. (a) is the input images. (b)–(f) are the results of Brown and Lown's approach, Xiong and Pulli's approach, Shen et al.'s approach, Yu et al.'s approach and our proposed approach, respectively.

Table 2

The quantitative quality assessment of the color correction results generated by different approaches.

Dataset	Input			Approach 1			Approach 2			Approach 3			Approach 4			Our approach		
	CD	GL	T (s)	CD	GL	T (s)	CD	GL	T (s)	CD	GL	T (s)	CD	GL	T (s)	CD	GL	T (s)
Coast	13.85	0.00	–	5.22	1.45	6.73	10.85	1.32	6.96	3.95	1.41	7.95	3.39	1.56	31.01	2.05	1.41	17.79
City	9.24	0.00	–	6.21	1.75	1.39	5.68	2.11	1.18	4.22	1.74	1.34	2.67	1.73	5.02	2.44	1.72	2.93
Village	19.54	0.00	–	9.07	1.57	34.01	16.06	1.36	37.90	6.60	1.49	42.27	6.54	1.59	46.89	4.63	1.46	68.27
Kitchen	20.55	0.00	–	8.16	4.46	3.51	15.92	4.19	5.71	5.15	4.66	6.06	6.07	4.33	16.81	2.99	4.55	8.76

Approach 1 denotes the Brown and Lown's approach. Approach 2 denotes the Xiong and Pulli's approach. Approach 3 denotes the Shen et al.'s approach. Approach 4 denote the Yu et al.'s approach.

- Instead of using the same global model to correct the color discrepancies of all pixels in one image, we proposed a novel flexible model that is comprised of a series of local linear models to locally correct the color discrepancies. This model is flexible enough to approximate complicated color discrepancies. Notably, this model can effectively express the local color discrepancies.
- To globally and simultaneously solve the model parameters for all input images, we designed a special cost function based on the

proposed model. We can effectively eliminate the global and local color discrepancies between multiple images while preserving the image gradient of each individual images at the same time by minimizing this energy function.

The experimental results on several challenging datasets captured by different sensors also demonstrate the superiority of our method over the state-of-the-art approaches in both the visual quality and quantitative metrics.

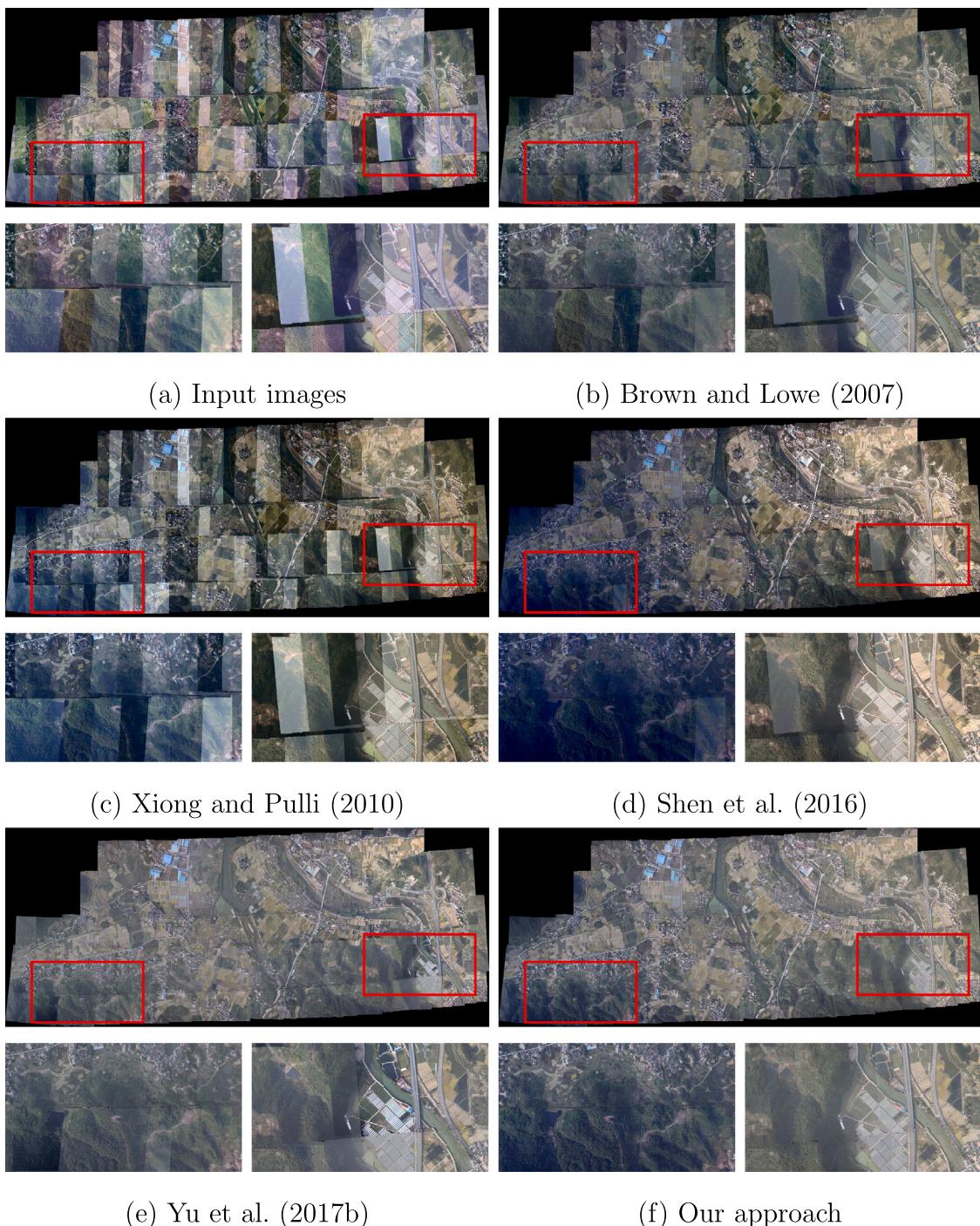


Fig. 9. The color correction results of the Village dataset. (a) is the input images. (b)–(f) are the results of Brown and Lown's approach, Xiong and Pulli's approach, Shen et al.'s approach, Yu et al.'s approach and our proposed approach, respectively.

Nevertheless, the proposed approach has some disadvantages. First, although the gradient preservation has been considered in the energy function, the corrected result still loses some image details, and thus a more effective gradient preservation strategy should be proposed. Second, the proposed approach performs well in eliminating color discrepancies, but the computational cost of our approach is high because the number of parameters that need to be solved in our approach is large. Last, since the original spectral signatures may be destroyed after color correction, the proposed approach cannot be applied in some remote sensing applications in which the original spectral signatures are necessary.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors want to thank the editor and three anonymous reviewers for their valuable comments and suggestions to improve the manuscript. This work was partially supported by the National Key Research and Development Program of China (Project No.

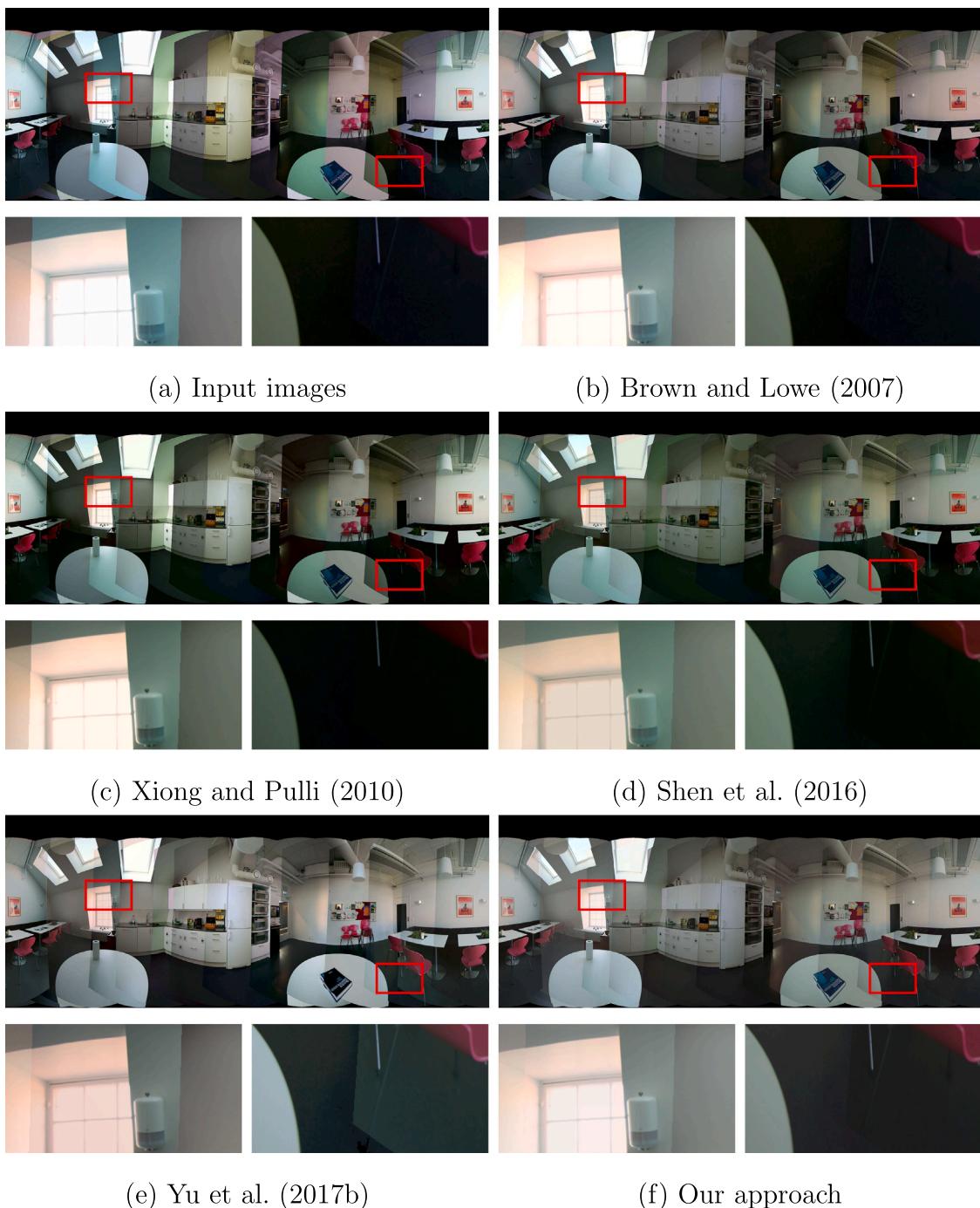


Fig. 10. The color correction results of the Kitchen dataset. (a) is the input images. (b)–(f) are the results of Brown and Lown's approach, Xiong and Pulli's approach, Shen et al.'s approach, Yu et al.'s approach and our proposed approach, respectively.

2017YFB1302400) and the National Natural Science Foundation of China (Project No. 41571436).

References

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (11), 2274–2282.
Arbelot, B., Vergne, R., Hurtut, T., Thollot, J., 2017. Local texture-based color transfer and colorization. *Comput. Graph.* 62, 15–27.

- Brown, M., Lowe, D.G., 2007. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vis.* 74 (1), 59–73.
Canty, M.J., Nielsen, A.A., Schmidt, M., 2004. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sens. Environ.* 91 (3–4), 441–451.
Chen, C., Chen, Z., Li, M., Liu, Y., Cheng, L., Ren, Y., 2014. Parallel relative radiometric normalisation for remote sensing image mosaics. *Comput. Geosci.* 73, 28–36.
Chen, K., Tu, J., Yao, J., Li, J., 2018. Generalized content-preserving warp: Direct photometric alignment beyond color consistency. *IEEE Access* 6, 69835–69849.
Cresson, R., Saint-Geours, N., 2015. Natural color satellite image mosaicking using quadratic programming in decorrelated color space. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8 (8), 4151–4162.

- Fan, Y., Liu, P., Niu, Y., 2019. Deep residual optimization for stereoscopic image color correction. In: International Symposium on Parallel Architectures, Algorithms and Programming. Springer.
- Fang, F., Wang, T., Fang, Y., Zhang, G., 2019. Fast color blending for seamless image stitching. *IEEE Geosci. Remote Sens. Lett.* 16 (7), 1115–1119.
- HaCohen, Y., Shechtman, E., Goldman, D.B., Lischinski, D., 2013. Optimizing color consistency in photo collections. *ACM Trans. Graph.* 32 (4), 1–10.
- He, M., Liao, J., Chen, D., Yuan, L., Sander, P.V., 2019. Progressive color transfer with dense semantic correspondences. *ACM Trans. Graph.* 38 (2), 1–18.
- Hu, H., Wu, B., Chen, L., 2019. Color balancing and geometrical registration of high-resolution planetary imagery for improved orthographic image mosaicking. *Planet. Space Sci.* 178, 104719.
- Ibrahim, M.T., Hafiz, R., Khan, M.M., Cho, Y., 2016. Automatic selection of color reference image for panoramic stitching. *Multimedia Syst.* 22 (3), 379–392.
- Levin, A., Zomet, A., Peleg, S., Weiss, Y., 2004. Seamless image stitching in the gradient domain. In: European Conference on Computer Vision.
- Li, X., Feng, R., Guan, X., Shen, H., Zhang, L., 2019b. Remote sensing image mosaicking: Achievements and challenges. *IEEE Geosci. Remote Sens. Mag.* 7 (4), 8–22.
- Li, J., Hu, Q., Ai, M., 2017a. Optimal illumination and color consistency for optical remote-sensing image mosaicking. *IEEE Geosci. Remote Sens. Lett.* 14 (11), 1943–1947.
- Li, L., Tu, J., Gong, Y., Yao, J., Li, J., 2019a. Seamline network generation based on foreground segmentation for orthoimage mosaicking. *ISPRS J. Photogramm. Remote Sens.* 148, 41–53.
- Li, L., Yao, J., Shi, S., Yuan, S., Zhang, Y., Li, J., 2018. Superpixel-based optimal seamline detection in the gradient domain via graph cuts for orthoimage mosaicking. *Int. J. Remote Sens.* 39 (12), 3908–3925.
- Li, L., Yao, J., Xie, R., Xia, M., Zhang, W., 2017b. A unified framework for street-view panorama stitching. *Sensors* 17 (1), 1.
- Lin, K., Jiang, N., Cheong, L.F., Do, M., Lu, J., 2016. Seagull: Seam-guided local alignment for parallax-tolerant image stitching. In: European Conference on Computer Vision.
- Niu, Y., Zheng, X., Zhao, T., Chen, J., 2019. Visually consistent color correction for stereoscopic images and videos. *IEEE Trans. Circuits Syst. Video Technol.* 30 (3), 697–710.
- Pan, J., Wang, M., Li, D., Li, J., 2010. A network-based radiometric equalization approach for digital aerial orthoimages. *IEEE Geosci. Remote Sens. Lett.* 7 (2), 401–405.
- Pandey, A., Pati, U.C., 2019. Image mosaicing: A deeper insight. *Image Vis. Comput.* 89, 236–257.
- Park, J., Tai, Y.W., Sinha, S.N., So Kweon, I., 2016. Efficient and robust color consistency for community photo collections. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 430–438.
- Pérez, P., Gangnet, M., Blake, A., 2003. Poisson image editing. *ACM Trans. Graph.* 22 (3), 777–784.
- Shen, T., Wang, J., Fang, T., Zhu, S., Quan, L., 2016. Color correction for image-based modeling in the large. In: Asian Conference on Computer Vision. Springer, pp. 392–407.
- Su, Z., Zeng, K., Liu, L., Li, B., Luo, X., 2014. Corruptive artifacts suppression for example-based color transfer. *IEEE Trans. Multimed.* 16 (4), 988–999.
- Syariz, M.A., Lin, B.Y., Denaro, L.G., Jaelani, L.M., Van Nguyen, M., Lin, C.H., 2019. Spectral-consistent relative radiometric normalization for multitemporal landsat 8 imagery. *ISPRS J. Photogramm. Remote Sens.* 147, 56–64.
- Xia, M., Yao, J., Gao, Z., 2019. A closed-form solution for multi-view color correction with gradient preservation. *ISPRS J. Photogramm. Remote Sens.* 157, 188–200.
- Xia, M., Yao, J., Xie, R., Zhang, M., Xiao, J., 2017. Color consistency correction based on remapping optimization for image stitching. In: IEEE International Conference on Computer Vision Workshops.
- Xiang, T.Z., Xia, G.S., Bai, X., Zhang, L., 2018. Image stitching by line-guided local warping with global similarity constraint. *Pattern Recognit.* 83, 481–497.
- Xie, R., Xia, M., Yao, J., Li, L., 2018. Guided color consistency optimization for image mosaicking. *ISPRS J. Photogramm. Remote Sens.* 135, 43–59.
- Xiong, Y., Pulli, K., 2010. Color matching of image sequences with combined gamma and linear corrections. In: International Conference on ACM Multimedia.
- Yu, L., Zhang, Y., Sun, M., Lu, Y., 2017a. Automatic reference image selection for color balancing in remote sensing imagery mosaic. *IEEE Geosci. Remote Sens. Lett.* 14 (5), 729–733.
- Yu, L., Zhang, Y., Sun, M., Zhou, X., Liu, C., 2017b. An auto-adapting global-to-local color balancing method for optical imagery mosaic. *ISPRS J. Photogramm. Remote Sens.* 132, 1–19.